Can In-Silico Computational Virtual Screening of Fragment Libraries Model In-Vitro Biophysical Screening Methods with Similar Results to Identify Novel RNA-Binding Chemical Matter for Treatment of Myotonic Dystrophy?

Can In-Silico Computational Virtual Screening of Fragment Libraries Model In-Vitro Biophysical Screening Methods with Similar Results to Identify Novel RNA-Binding Chemical Matter for Treatment of Myotonic Dystrophy?

Ameya D. Madduri

A Thesis in the Field of Biotechnology

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2023

Abstract

Myotonic Dystrophy is an autosomal dominant trinucleotide repeat expansion disorder caused by a CTG repeat in the 3' untranslated region of the DMPK gene. These repeats become pathogenic when transcribed into RNA as they form ribonuclear foci comprised of auto complementary CUG hairpin structures, which in turn bind MBNL1, a key RNA splicing regulator. DM1 patients suffer from multisystemic muscle wasting, or myopathy, in muscled areas like arms, forearms, hands, ankles, jaws, tongue, and neck flexors (Ozimski et al., 2020). Utilizing small molecules to target CUG RNA hairpins is a potential tractable mechanism to prevent downstream mis-splicing thereby improving disease phenotype in patients. No published work, on novel small molecule drug discovery or drug repurposing, has shown potent CUG RNA binding with downstream in-vivo splicing rescue. This study shows early hit finding for small molecule binding to CUG RNA hairpins can be achieved using both biophysical methods and computational screening. Utilizing both approaches in conjunction, a statistically significant number (3.2%) of overlapping hits can be selected for further experimental validation.

Acknowledgments

I would like to thank and acknowledge the numerous individuals who provided guidance and support for the scientific work presented in this thesis. First and foremost, I would like to thank my family- Savitha and Hari Madduri, Laasya Madduri for their support of my educational endeavors. Next, I would like to express my gratitude and sincere thanks to the past and current members of the Chemistry/Biochemistry team at Syros Pharmaceuticals without whom this work would not have been accomplished-

Kent Baker, Jhuma Das, Diane Hamann, Claudio Chuaqui, Jason Marineau, Derek LaPlaca, Modi Sathappa, Kristin Hamman.

I would also like to thank James Morris for his guidance and support of this work as my thesis advisor.

Table of Contents

List of Tables

List of Figures

Chapter I.

Introduction to Genetic Diseases and Myotonic Dystrophy

Genetic disorders are commonly described as a disease which occurs due to a gene mutation at the DNA level. These DNA mutations may cause downstream RNA transcription errors and perturbed protein translation. Recent genome mapping has shown certain mutations can increase risk for developing a genetic disorder either with symptoms at birth or developed over time ("Genetic disorders: What are they types, symptoms & causes", 2021). Genetic disorders can be categorized into three main sections- chromosomal, complex or multifactorial, and single-gene or monogenic. A chromosomal disorder is when the structure of the chromosomes is disrupted which manifests as patients having translocated or duplicated chromosomal material. Examples of this are found in diseases such as Down syndrome (Trisomy 21), Fragile X syndrome, and Klinefelter syndrome. Multifactorial or complex disorders stem from both gene mutations at the DNA -level as well as environmental influences such as chemical exposure, diet, medication use, or tobacco or alcohol usage. These include diseases such as spina bifida, diabetes, coronary heart disease, autism spectrum disorders, and cancer. Finally, single-gene or monogenic disease arise from a single gene mutation ("Genetic disorders: What are they types, symptoms & causes", 2021). Common monogenic diseases are Cystic fibrosis, Sickle cell disease, Tay-Sachs disease, and Duchenne muscular dystrophy ("Genetic disorders: What are they types, symptoms & causes", 2021).

Each genetic disorder has unique disease biology and phenotypic symptoms based on the organ system that is impacted by the genome mutation. Some common symptoms of genetic disease, which are disease agnostic, are poor growth or short stature, muscle weakness or stiffness, eating and digestive issues, developmental delays including speech impediments or delayed social skills, cognitive deficits, or behavioral disturbances ("Genetic disorders: What are they types, symptoms & causes", 2021).

Identification of genetic diseases can be done at several stages. Genetic counseling can help educate on risks and steps needed to mitigate those risks if there is a known gene mutation and family history of disease. For example, if there is sufficient concern about family history, a carrier blood test can be done pre-conception. Carrier testing will indicate if either maternal or paternal genetic material carry a mutation for the genetic disorder in question. Prenatal screening is done at the pregnancy stage wherein a blood sample taken from the pregnant patient can indicate the likelihood of the unborn child having a specific known genetic disorder. Prenatal diagnostic testing is also done at the pregnancy stage wherein amniocentesis can indicate the unborn child's risk level for genetic disorders. Finally at the post-birth stage, a newborn screening blood test can indicate if the infant has a genetic disorder ("Genetic disorders: What are they types, symptoms & causes", 2021).

Like symptom specificity to any disorder, there are both generalized and unique treatments for genetic disorders. Some of the most common treatments include blood transfusions, organ transplants, occupational and speech therapy, counseling on nutrition and supplements, specialized treatments like surgery and radiation, and drug therapies

with various mechanisms of action ("Genetic disorders: What are they types, symptoms & causes", 2021).

Diving deeper into monogenic disease, Chial et. al share in their 2008 Nature review that the human genome contains 20,000-25,000 genes. In single-gene diseases, a mutation in just one of these genes is responsible for a systemic disease. Single-gene diseases typically run in families and can be dominant or recessive, and autosomal or sex-linked. One such group of diseases are muscular dystrophies. NYU Langone Health reports that there are over 30 types of muscular dystrophies caused by various known genetic mutations. Some disease types are mild and progress slowly over a lifetime without significant debilitation, while others cause rapid muscle wasting, physical disability, and/or shortened lifespan ("Types of muscular dystrophy", n.d). Within the set of 30 known muscular dystrophies, there are several well-characterized diseases such as Becker muscular dystrophy (BMD), congenital muscular dystrophies (CMD), Duchenne muscular dystrophy (DMD), Emery-Dreifuss muscular dystrophy (EDMD), Limb-girdle muscular dystrophy (LGMD), Myotonic dystrophy (DM), and Oculopharyngeal muscular dystrophy (OPMD).

Two of the most common types of muscular dystrophies are Duchenne and Becker muscular dystrophies which are both caused by mutations to the dystrophin (DMD) protein. In Duchenne muscular dystrophy, functional dystrophin is absent in muscle ("Duchenne and Becker muscular dystrophy", n.d). In Becker muscular dystrophy, there is some dystrophin present, but not enough for normal muscle function ("Duchenne and Becker muscular dystrophy", n.d). More deeply, Duchenne muscular dystrophy is a progressive, muscle-wasting disease caused by a mutation in the DMD

protein that prevents the production of the muscle isoform of dystrophin Dp427m (Duan, et al., 2021). Frameshift mutations or nonsense mutations cause premature truncation of protein translation, thereby leading to non-functional and unstable dystrophin (Duan et al., 2021). Both Duchenne and Becker muscular dystrophies are most often diagnosed in childhood or early adolescence. By comparison, another type of muscular dystrophy, Myotonic muscular dystrophy, the most common form of muscular dystrophy, is most often diagnosed in adults ("Types of muscular dystrophy", n.d).

At the broadest level, Myotonic dystrophy (DM) is an autosomal dominant muscular dystrophy which results in both skeletal and cardiac muscle complications (Kim et al., 2019). Type 1 DM is a nucleotide repeat expansion disorder caused by a CTG trinucleotide expansion in the 3' untranslated region of the DMPK (DM1 Protein Kinase) gene while type 2 (DM2) arises from a CCTG tetranucleotide expansion in intron 1 of the CNBP gene (Kim et al., 2019). Research in DM1 has found that there is a correlation between the number of these repeats and the age of disease onset (Kim et al., 2019). Those with 50-1,000 repeats typically have adult onset of DM1 and those with greater than 1,000 repeats typically have congenital or early onset disease (Kim et al., 2019). By comparison, in DM2 no data currently suggests correlation between repeat number and disease onset. All DM2 patients are diagnosed as adult-onset without incidences of a congenital form.

Taking a deeper look at the molecular mechanism of DM1, the disease is caused by an expansion of CTG repeats in the 3' untranslated region (UTR) of the DMPK gene. These become pathogenic when transcribed into RNA wherein they form ribonuclear foci comprised of auto complementary CUG hairpin structures which bind proteins (Ozimski

et al., 2020). Specifically, the CUG hairpin structures bind and sequester muscleblind-like (MBNLs) and CUGBP Elav-like family member (CELF1) proteins which are key regulators of RNA splicing, polyadenylation, and stability (Ozimski et al., 2020). Hence, most DM1 patients suffer from multisystemic muscle wasting, or myopathy, in muscled areas like arms, forearms, hands, ankles, jaws, tongue, and neck flexors (Ozimski et al., 2020).

Ozimski et al. aptly share in their 2021 review paper that there are currently no interventional or therapeutic treatment options for DM1 patients. In Pascual-Gilabert et al.'s 2021 review paper, they share three main potential avenues for interventional therapies- small molecules, oligonucleotide therapies, and gene therapies. There are currently 10 Myotonic dystrophy small molecule drug candidates in phase II and III clinical trials. Nine of these drug compounds are repurposed from other indications and show efficacy in myotonia symptom reduction, but do not directly target the CUG hairpin RNA as a mechanism of action.

Most DM1 research and therapies have focused on the MBNL and CELF1 proteins, attempting to reverse their functional disruption (Ozimski et al., 2020). Drug discovery have been done to develop new therapeutic drugs that directly targets the CUG RNA repeats. Academic chemical biology groups have published extensively on development of small molecule compounds that bind to the CUG hairpin with splicing rescue in-vivo. Two of the key academics in this space are Steve Zimmerman and Andrew Berglund.

Steven Zimmerman's lab has worked to synthesize and evaluate structurally diverse small molecules that bind either CTG DNA repeats or CUG RNA repeats found

in the DM1 codon repeat expansion stretches of the genome. Specifically, the work shows both macrocycles and triaminoatrizine-acridine ring compounds can bind to the repeats. Nyugyen et al show in their 2015 paper a set of triaminoatrizine-acridine compounds that have μM K_d's, measured by Isothermal Titration Calorimetry (ITC), to a (CUG)12 repeat oligomer. Biological data for these compounds show some reduction of the DM1 disease phenotype. It also indicates that further work will be needed to develop a novel nanomolar CUG repeat inhibitor to reduce hairpin binding of MBNL1 thereby rescuing mis-splicing affects seen in the disease physiology. X-ray crystallography work described in a 2020 suggests that the triaminoatrizine-acridine compound binds U: U mismatches by intercalation. This again suggests that there is further scope to develop additional small molecule inhibitors that have multiple binding modes, rather than just intercalation and additionally focus on physiologically relevant CUG repeat binding instead of artificially introduced U: U mismatches on the CUG hairpin.

Andrew Berglund's lab has published work on disruption of the MBNL1/CUG4 interaction using an FDA approved small molecule drug called pentamidine. While approved for use in treatment of Pneumocystis carinii infections, it also showed in-vivo splicing rescue in a DM1 HeLa cell model with 960 interrupted CUG repeats (Coonrod et al., 2013). Pentamidine also improved mis-splicing defects of endogenous transcripts in a DM1 mouse model expressing 220 CUG repeats (Coonrod et al., 2013). Further research showed that the pentamidine itself does not interact with MBNL1 but instead results in a drop in CUG transcript levels. Suggesting that pentamidine decreases transcription of the CUG RNA or increases the rate of degradation (Coonrod et al., 2013).

The small molecule work from the Zimmerman and Berglund labs along with other DM1-focused chemical biology labs shows clear biological tractability for targeting the toxic CUG RNA hairpins as a method to prevent downstream mis-splicing. However, no de-novo small molecule discovery or drug repurposing has shown CUG RNA binding potent enough for downstream in-vivo splicing rescue. The scope of the published and clinical work leaves this major gap to address. This thesis describes early small molecule discovery to develop novel chemical matter to address this gap. There is clear biological target tractability and a clear unmet patient need for such drugs.

The following section details an overview of fragment-based drug discovery and explains the theory of both surface plasmon resonance and molecular docking. Furthermore, it outlines the key research aims and hypotheses to be tested in this study.

Fragment Based Drug Discovery

Fragment-based drug discovery (FBDD) is a method to develop small molecule inhibitors starting from weakly binding low molecular weight chemical fragments (Li, 2022). Literature precedence indicates that targets can range from oligonucleotides, proteins, or protein-protein interactions. FBDD utilizes biophysical and/or biochemical methods to detect fragment binding which can in-turn can be optimized into drug leads (Kirsch, et al, 2019). Fragments are advantageous starting points for drug discovery as they are non-complex structures which allow for direct target engagement through binding in the active site, smaller allosteric pockets, or in-between a protein-protein interaction. Typical fragments obey what literature calls the “rule-of-three” or RO3 wherein fragments are ≤300 Da in molecular weight, have ≤3 hydrogen bond donors,

have ≤3 hydrogen bond acceptors, and the logP is ≤3. (Kirsch, 2019). An additional benefit of FBDD is that fragments are typically small enough to avoid steric clashes. While initial fragment hits typically have weak affinity, in the μM - mM range, they offer opportunity for enhanced potency through fragment linking, merging, and other structure-growing strategies.

FBDD can be executed using several different characterization techniques including X-ray crystallography, Nuclear Magnetic Resonance (NMR), Surface Plasmon Resonance (SPR), Interferometry, Isothermal Titration Calorimetry (ITC), Mass Spectrometry, and Computational Methods, as outlined in Erlanson's 2011 review of fragment-based drug discovery. X-ray crystallography and NMR provide structural insight into how fragments bind intended targets. SPR and Interferometry are biophysical methods which measure fragment binding to an intended target by detecting a change in refractive index caused by a shift in light. ITC is a method which measures the heat loss and resulting entropy or enthalpy changes due to a fragment binding to a target. Mass Spectrometry is an analytical method which measures the mass-to-charge ratio (m/z) and can determine the mass of an analyte. Computational methods include the utilization of software to model molecular interactions. Each method results in the generation of different types of data to help characterize binding interactions between fragments and the target of interest. Each method also has unique challenges including target flexibility, experimental development and optimization time, throughput, and sample consumption to name a few. In this study, SPR and computational methods were elected for fragment characterization and evaluation.

Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) is one of the primary biophysical methods used for the screening of FBDD fragment libraries due to its low protein consumption and label-free methodology (Shepherd et al., 2014). SPR biosensor interaction analysis is employed to screen, confirm, and optimize binding of fragments as it provides accurate information on the affinity and kinetics of molecular interactions (Shepherd et al., 2014). In a typical SPR experiment, a target or protein of interest is immobilized onto a flexible dextran matrix surface while the analyte is flowed over the ligand-bound surface using microfluidics. Binding interaction results in an increase in mass on the sensor surface resulting in a change in the refractive index of the medium in the vicinity of the sensor surface. The analyte induced change in the refractive index is measured in resonance units (R.U.). Signal changes due to the association and disassociation of complexes are monitored in a non-invasive manner, continuously, and in real-time, the results of which are reported in the form of a sensorgram.

Molecular Docking

Molecular docking is a computer-based method for virtually screening fragments in drug discovery (Kirsch 2019, Luo et al 2019). This approach allows for the virtual docking of millions of compounds with diverse chemical scaffolds filtering the list of virtual hits to <1000. Allowing for manageable follow-up with experimental biophysical and biochemical methods. Screening can be either ligand-based or target-based depending on where there is more available knowledge. Ligand-based screening utilizes diverse fragment libraries to determine the essential molecular features and functional groups needed for target binding (Kirsch, 2019). Target-based or structure-based virtual

screening requires knowledge of the target structure where fragments are computationally docked into the structure to predict how and where chemical matter binds (Kirsch, 2019). Determining hits can be done using molecular docking wherein first the preferred binding mode of the ligand are determined and then a scoring function is applied to estimate the binding energy to help select the favorable binding modes (Luo et al. 2019).

Key Research Questions and Specific Aims

1. Can a biophysical SPR-based screen identify specific 30-mer CUG RNA hairpin binding fragments?
    i. This aim of this project sought to understand if any fragments in the Enamine Golden fragment library binds specifically to the 30-mer RNA DM1 related hairpin oligomer. The RNA hairpin structure has been implicated in sequestration of MBNL1 thereby causing the mis-splicing disease phenotype. A 30-mer RNA oligo was selected as it contains sufficient repeats to form a hairpin structure and is still small enough to be efficiently immobilized onto an SPR chip surface without concern for surface conformation changes of the oligonucleotide. To generate this data, an SPR biophysical screen of the Enamine Golden fragment library was performed to determine which fragments bind to the 30-mer RNA hairpin oligo target. The Enamine Golden fragment library is an internally available library with good diversity and structural similarity to published fragment compound cores. The primary screen was done at two concentrations followed by a secondary 10-point dose response screen to confirm hit binding.

2. Can a computational docking of the same fragment library generate a hit list of fragments with energetically favorable binding interactions to a 30-mer RNA hairpin?
    i. The second aim of this project sought to understand if any fragments in the Enamine Golden fragment library computationally identified to bind to the RNA hairpin structure. To answer this question, we proposed initiating a virtual molecular docking screen. The high-resolution X-ray crystallography structures of several RNA hairpin oligos published in the RCSB PDB thus providing a good docking target for this screen.

The generated data will confirm or disprove the hypothesis that an in-silico computational virtual screen of a fragment library can identify a statistically significant number of overlapping novel RNA binding chemical matter hits as compared to an in-vitro biophysical screen utilizing the same fragment library.

Chapter II.

Materials and Methods

This section details how the fragment screens were designed and run.

Surface Plasmon Resonance Screen

The SPR biophysical screen was performed on the Biacore 8K instrument. The instrument utilizes the phenomenon of Surface Plasmon Resonance (SPR) to detect and measure binding interactions. In a typical SPR-based experiment, a ligand of interest is immobilized onto a flexible dextran matrix while the analyte is injected and flowed over the ligand-bound surface. A binding interaction results in an increase in mass on the sensor surface resulting in a change of the refractive index in the vicinity of the sensor surface. Changes in the refractive index are recorded as resonance units (R.U.). Signal changes due to association and dissociation of complexes are monitored in real-time, continuously, and in a non-invasive manner, the results of which are reported in the form of a time resolved sensorgram.

For this screen, the biotinylated 30-mer RNA target was immobilized to an 8-channel chip neutravidin surface. Biotinylated surface capture chemistry was selected as the optimal surface-capture chemistry based upon previously collected internal data. The high affinity biotin neutravidin interaction gives an exceptionally stable surface for screening. The fragment library was screened at two concentrations over the immobilized RNA surface to assess binding interactions between the fragments and the

immobilized RNA target. The screening methodology was adapted from the methods outlined by Gianetti in the 2011 publication of high-throughput fragment screening.

SPR Method Development

The method development was performed as a series of experiments to determine assay conditions. The assay parameters to develop and validate were buffer conditions, DMSO tolerance of the RNA, surface density needed to observe binding response of the fragment, determination and evaluation of a positive control compound, and surface stability of the immobilized RNA target across the experiment duration. Thermal stability or melting temperature of the CUG30 was determined using intrinsic fluorescence of the oligomer. The thermal stability experiments were used to determine oligomer biophysical stability in the optimized buffer and DMSO conditions.

Parameter 1: Buffer stability and buffer excipient concentration determination

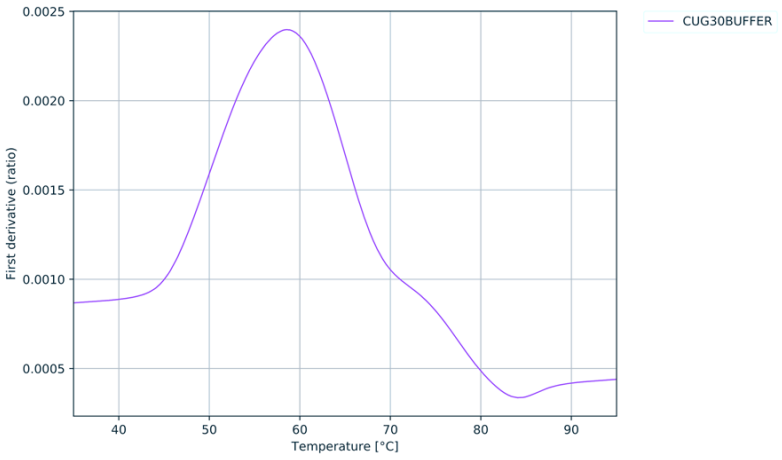


Figure 1. CUG30 RNA Thermal Stability Determination.

The CUG30 RNA oligomer shows a thermal melting temperature (Tm) of 58 °C. Assay buffer: 10 mM HEPES pH7.5, 50 mM NaCl, 0.005% Polysorbate 20.

Parameter 2: DMSO stability and tolerability

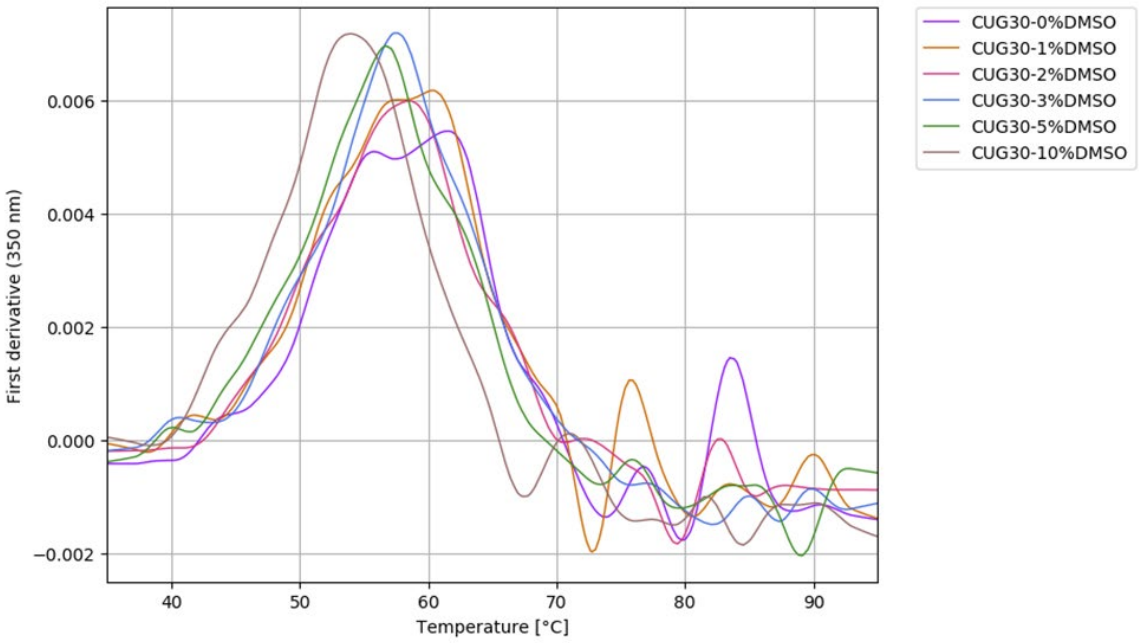


Figure 2. CUG30 RNA DMSO Tolerance Determination.

Data indicates that while thermal melt curves indicate destabilization of small populations at different temperatures, the major Tm peaks between 0-5% DMSO indicate oligo stability upwards of 5% DMSO. Therefore, 5% DMSO was selected for the screening conditions as it showed retained activity of the RNA while allowing for fragment solubility. Therefore, the final fragment screen buffer was determined to be 10 mM HEPES pH 7.50, 50 mM NaCl, 0.005% Polysorbate 20, 5% DMSO.

Parameter 3: Surface density of RNA required to observe fragment binding

Surface density was determined by using the R_{max} determination equation:

$$R_{max} = \left(\frac{(R_{ligand})(M_{r,analyte})(Valency_{ligand})}{M_{r,ligand}}\right)$$

Table 1. CUG30 RNA Immobilization Level Determination.

Parameter	Value
RNA Molecular Weight	9897 kDa
Ligand Molecular Weight	250 Da (average of library)
Target R_{max} Value	50 RU
Valency	Hypothesize 1:1 Binding

$$50 = \left(\frac{(R_{ligand})(9897)(1)}{250}\right) = 1200 \text{ RU}$$

Therefore, immobilize 1200 RU of RNA onto the chip surface.

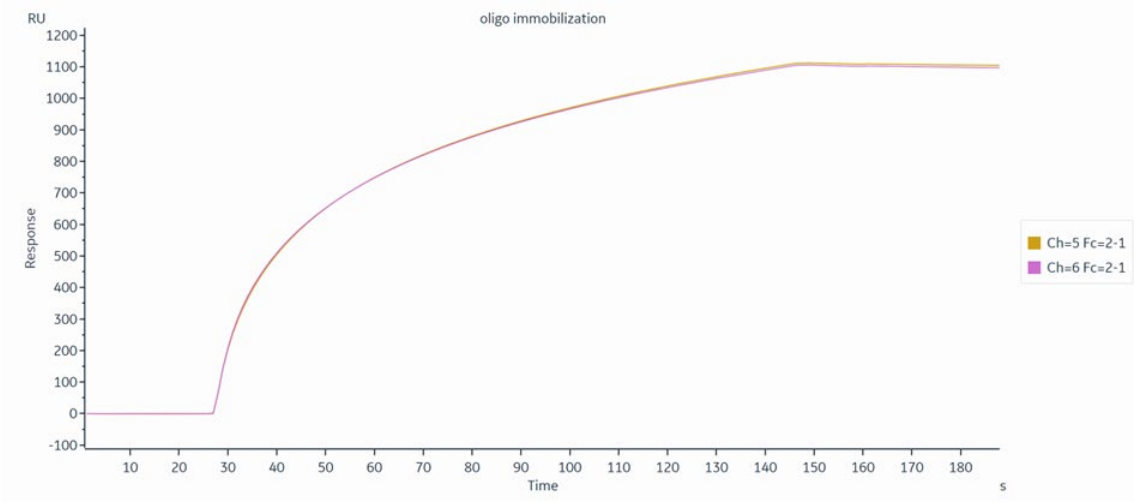


Figure 3. CUG30 RNA Successfully Immobilized at R_{ligand}= 1100 RU.

Parameter 4: Determination and evaluation of a positive control compound

- A positive control compound was necessary in fragment screens to ensure assay consistency, reproducibility, and comparability across different instrument runs (Gianetti et al 2011).

- Parameters to determine in positive control compounds: soluble, dose-dependent binding, consistency, and repeatability across the screen duration, fast on fast off kinetics preferred, no reference binding, and cross-validated as a binding in an orthogonal method. These compounds were identified as potential controls using previously collected internal data.

Compound 1 (SY2988)

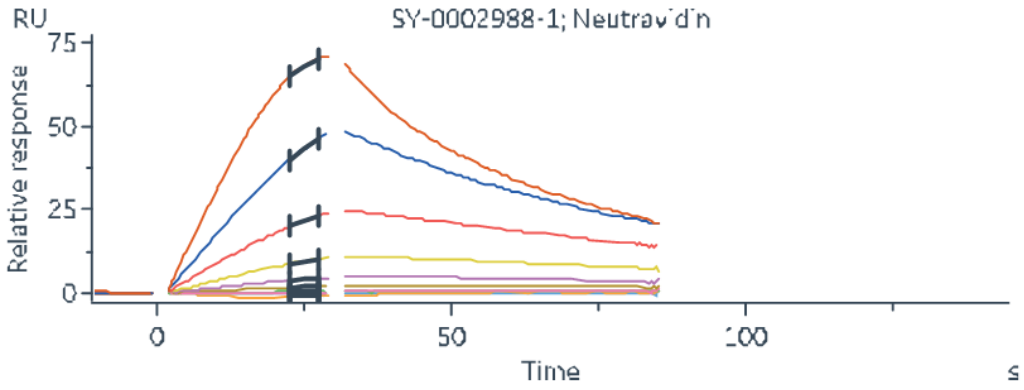


Figure 4. Dose Dependence Determination of Control Compound SY2988.

Table 2. Evaluation Data of Control Compound SY2988.

Evaluation Parameter	Result
Response – Beginning of Run (RU)	40
Affinity Beginning of Run (nM)	190
Response End of Run (RU)	30
Affinity End of Run (nM)	176
Soluble	Yes
Fast Off-rate Ranking *	4

Compound 2 (SY3170)

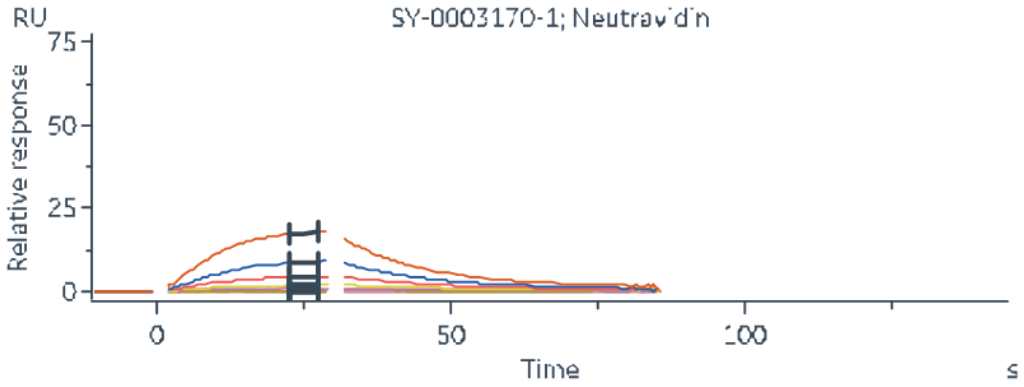


Figure 5. Dose Dependence Determination of Control Compound SY3170.

Table 3. Evaluation Data of Control Compound SY3170.

Evaluation Parameter	Result
Response – Beginning of Run (RU)	10
Affinity Beginning of Run (nM)	1151
Response End of Run (RU)	5
Affinity End of Run (nM)	1325
Soluble	Yes
Fast Off-rate Ranking *	3

Compound 3 (SY2611)

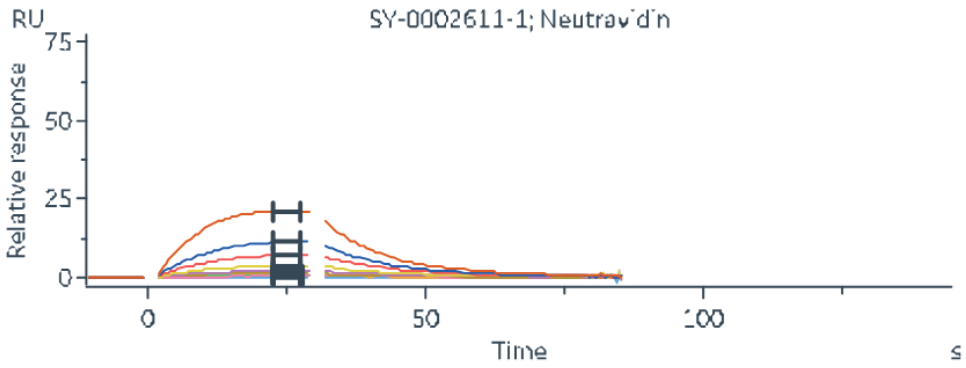


Figure 6. Dose Dependence Determination of Control Compound SY2611.

Table 4. Evaluation Data of Control Compound SY2611.

Evaluation Parameter	Result
Response – Beginning of Run (RU)	15
Affinity Beginning of Run (nM)	503
Response End of Run (RU)	10
Affinity End of Run (nM)	878
Soluble	Yes
Fast Off-rate Ranking *	2

Compound 4 (SY2546)

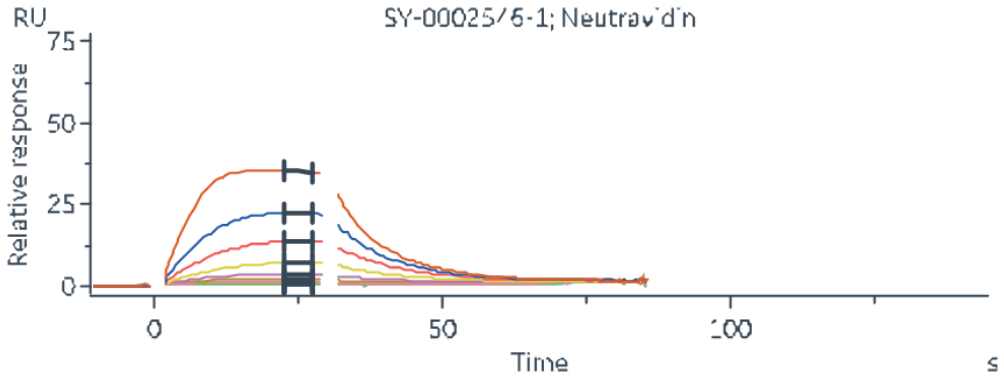


Figure 7. Dose Dependence Determination of Control Compound SY2546.

Table 5. Evaluation Data of Control Compound SY2546.

Evaluation Parameter	Result
Response – Beginning of Run (RU)	30
Affinity Beginning of Run (nM)	143
Response End of Run (RU)	20
Affinity End of Run (nM)	183
Soluble	Yes
Fast Off-rate Ranking *	1

Control 4 (SY2546) was selected as the compound to use as the positive control as it exhibited solubility, had the fastest off rate kinetics compared to other compounds, no reference binding, dose dependent binding, and additional internal data supporting binding in an orthogonal method.

Parameter 5: Compound binding consistency and determination and RNA surface stability determination

- Determination of compound binding consistency and RNA surface stability across the duration of the run was completed
- Compound 4 (SY2546) was run in a dose dependence experiment repeatedly for the duration of the experiment length to ensure surface stability

All parameters of the assay development resulted in the development and validation of a robust screening assay.

Executing on SPR Fragment Screen

The SPR fragment screen was performed in three steps in the lab.

1. Buffer Preparation

    The screen was run in the Biacore 8K across four methods spanning a duration of 18 hours for the 50 µM concentration. The 500 µM screen was also run across four methods spanning a duration of 18 hours each. Each method ran 576 unique fragments in single point. Each run required 2200 mL of assay buffer, 900 mL of DI H_20, and 500 mL of 50% DMSO.

2. Surface Preparation

The SPR chip surface was prepared in three steps- chip hydration, neutravidin immobilization, neutravidin conditioning, and RNA immobilization.

A CM5 chip was allowed to come to room temperature from 4°C and then placed in the Biacore 8K instrument (8K). The chip was then hydrated with 8 consecutive injections of buffer at 100 µL/min at 10 second intervals with each injection being followed by a 5 second instrument stabilization wait period. The buffer injections were followed by 8 injections of NaOH performed with the same parameters as the buffer—100 µL/min, 10 sec injection and 5 second stabilization period.

Neutravidin immobilization was performed after the chip hydration. The neutravidin immobilization on the CM5 chip was done via amine coupling as per manufacturer's instructions to reach immobilization levels of 10-12K RU's.

Table 6. Assay Reagents

Reagent Name	**Reagent Concentration/ Volume Supplied**	**Reagent Source**	**Catalog Numbers**
HBSP+ Buffer	10X	GE Healthcare (Cytiva Life Sciences)	BR-1006-71
Acetate pH **5.0** (Sodium Acetate pH **5.0**)	10 mM	GE Healthcare (Cytiva Life Sciences)	BR 100351
CM5 Chip	-	GE Healthcare (Cytiva Life Sciences)	29-1496-03
Neutravidin (1 mg/mL)	10 mg	ThermoFisher Scientific	31000
Sodium Chloride (NaCl)	5 M	Sigma Aldrich	S6546
Hydrochloric Acid (HCl)	100 µM	Fisher Scientific	7647-01-0,7732-18-5
Sodium Hydroxide (NaOH)	50 mM	GE Healthcare (Cytiva Life Sciences)	BR100358
Isopropanol (2-Propanol)	99.9%	Sigma Aldrich	190764
EDC (1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide Hydrochloride)	750 mg	GE Healthcare (Cytiva Life Sciences)	Amine Coupling Kit BR-1000-50
NHS (N-Hydroxysuccinimide)	115 mg		
Ethanolamine hydrochloride-NaOH pH 8.5	1 M		

Table 7. Preparation and Storage of Stock Solutions

Reagent Name	Volume	Component Name	Component Quantity	Final Stock Component Concentration
Neutravidin, stock solution *Store in 20 µL aliquots at -80°C*	10 mL	Neutravidin 10 mM HEPES pH 7.4	10 mg 10 mL	1 mg/mL
NHS	10 mL	NHS Deionized water	115 mg 10 mL	11.5 mg/mL
EDC	10 mL	EDC Deionized water	750 mg 10 mL	75 mg/mL

Table 8. Preparation of Working Solutions for Assay

Reagent Name	Volume	Component Name	Component Quantity	Working Component Concentration	Notes
HBSP+ Buffer (1X)	1 L	10X HBSP+ Buffer 100mM HEPES, etc	-	10mM HEPES, etc	Do a 1:10 dilution of stock buffer into DI H_20
Ethanolamine	140 µL	-	-	1 M	Use reagent as is from manufacturer
Neutravidin*	240 µL	1 mg/mL Neutravidin 10mM Acetate pH 4.5	15 µL 225 µL	62.5 µg/mL	Prepare by combing 225 µL Acetate pH4.5 buffer + 15 µL stock Neutravidin
EDC	90 µL	75 mg/mL	75 mg/mL	75 mg/mL	Use one of the previously prepared single-use aliquots from -80°C
NHS	90 µL	11.5 mg/mL	11.5 mg/mL	11.5 mg/mL	Use one of the previously prepared single-use

					aliquots from -80°C
Stock Wash Solution	10 mL	NaCl NaOH H_20	1.17 g 1 mL Up to 10 mL	2 M 100 mM	
Conditioning Solution 1	10 mL	NaCl HCl H_20	0.58 g 100 µL Up to 10 mL	1 M 10 mM	
Conditioning Solution 2	10 mL	Stock Wash Solution H_20	5 mL 5 mL (Mix in 1:1 ratio)	1 M NaCl 50 mM NaOH	
Isopropanol Wash Solution	10 mL	Stock Wash Solution Isopropanol	5 mL 5 mL	1 M NaCl 50 mM NaOH 50% Isopropanol	

**Experimental Note: Mix Neutravidin+Acetate immediately prior to beginning each immobilization cycle as lower immobilization levels had previously been observed when solution was prepared more than two minutes prior to the start of the immobilization protocol. This is due to Neutravidin being unstable at a low pH.*

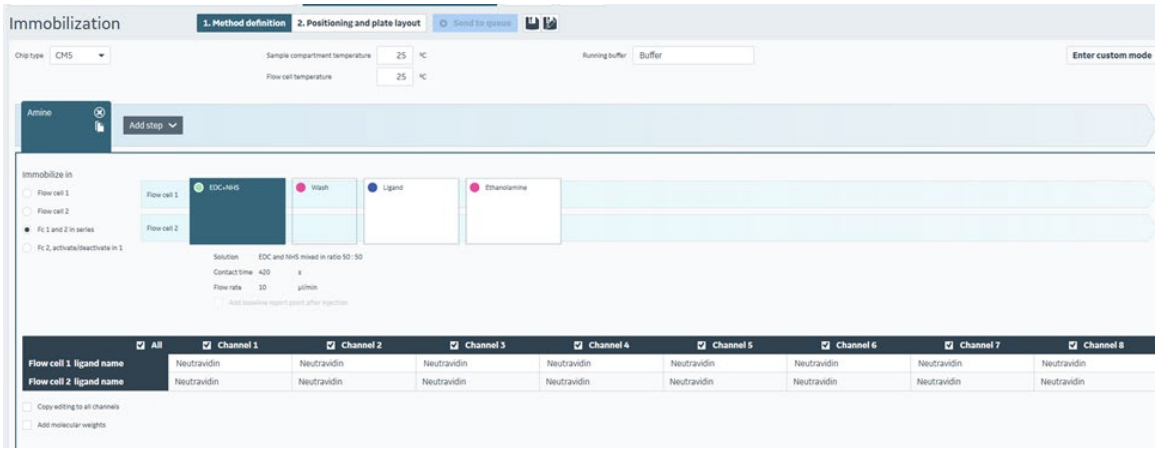


Figure 8. Neutravidin Immobilization Protocol Parameters

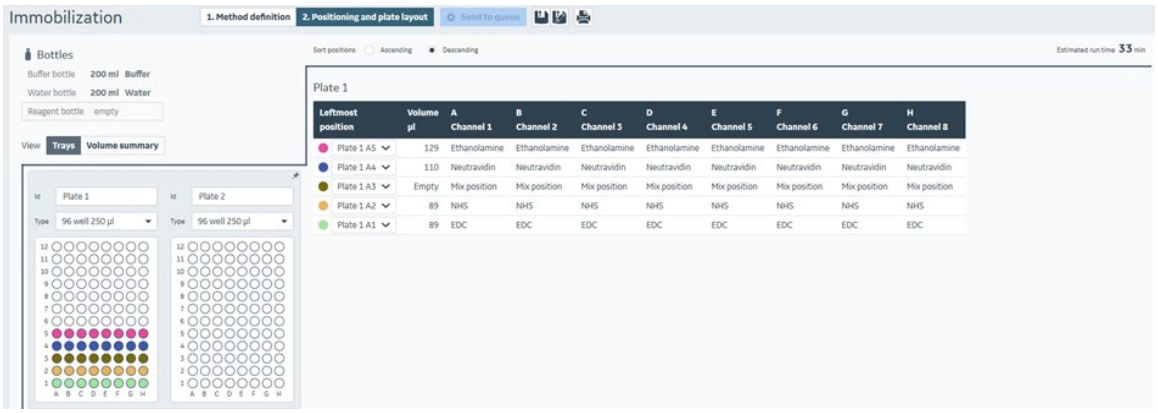


Figure 9. Neutravidin Immobilization Protocol Parameters- Locations and Volumes for CM5 Chip

After successful completion of the immobilization method, the neutravidin condition method was run to further condition the surface and remove any aggregated complexes in the flow system.

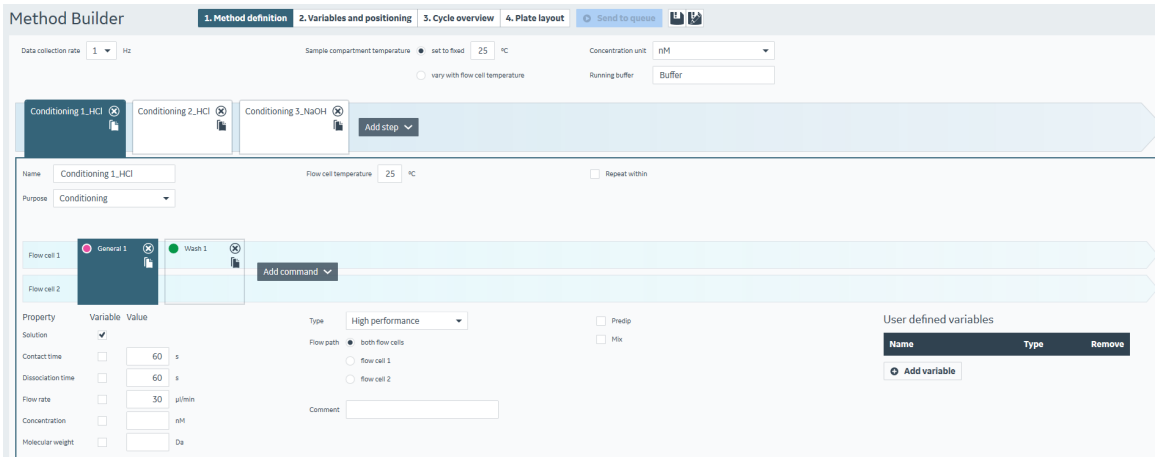


Figure 10. Neutravidin Conditioning Method Protocol Parameters

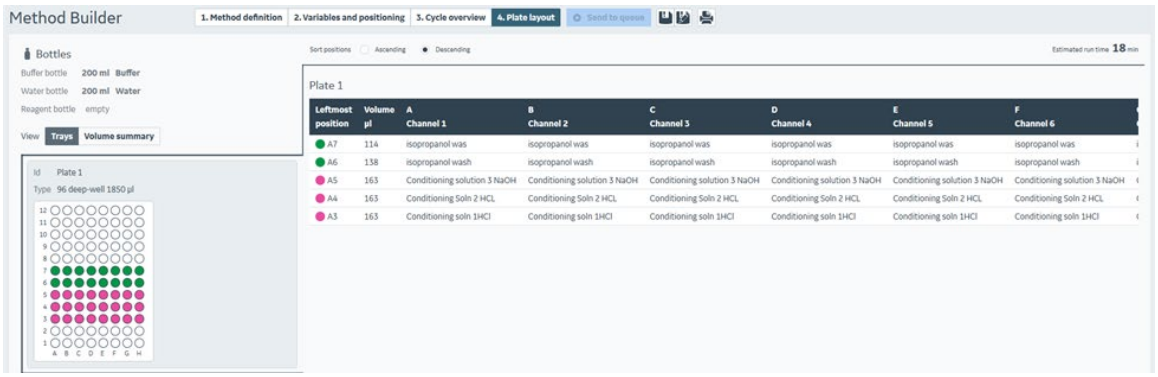


Figure 11. Neutravidin Conditioning Protocol Parameters- Locations and Volumes for CM5 Chip

After the neutravidin surface was conditioned, the RNA was immobilized onto the surface using biotinylated surface capture chemistry.

The RNA immobilization method had three sub portions:

- Initially all flow cells were injected with NaOH for 60 seconds at a 100 µL/min flow rate followed by a 60 second dissociation.

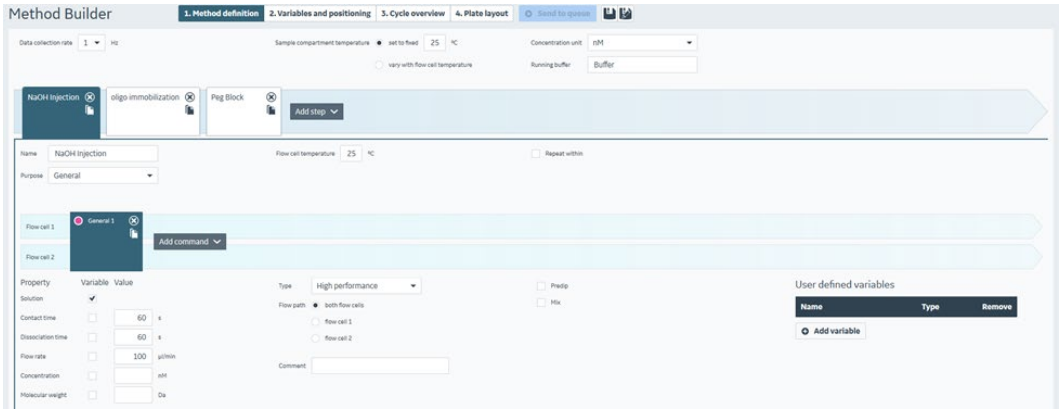


Figure 12. CUG30 RNA Oligomer Immobilization Method Protocol Parameters for NaOH.

- The oligomer stock was then injected onto the active surface for 120 seconds at a flow rate of 100 µL/min.

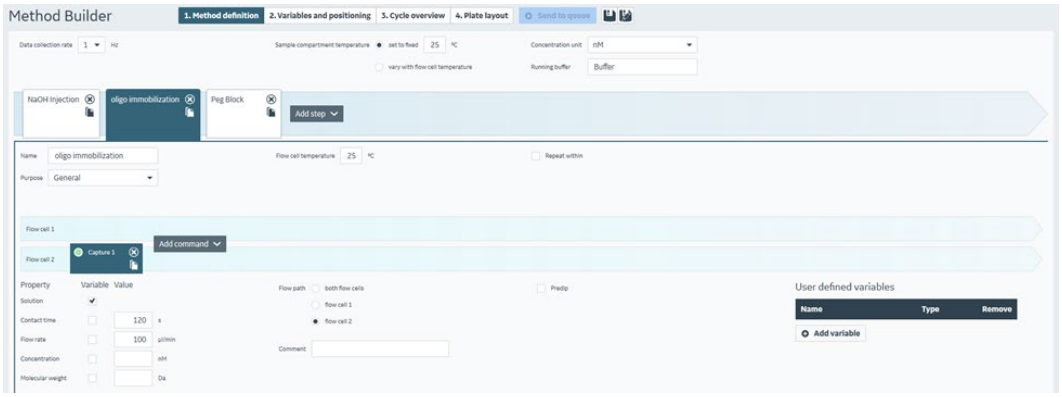


Figure 13. CUG30 RNA Oligomer Immobilization Method Protocol Parameters for RNA.

- Finally a PEG solution was injected over the surface for 60 seconds at a 100 µL/min flow rate followed by a 60 second dissociation.

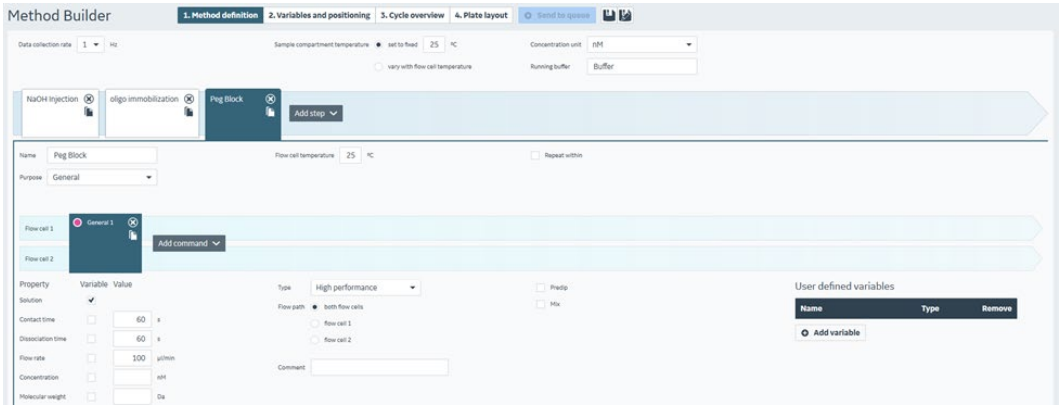


Figure 14. CUG30 RNA Oligomer Immobilization Method Protocol Parameters for PEG Block

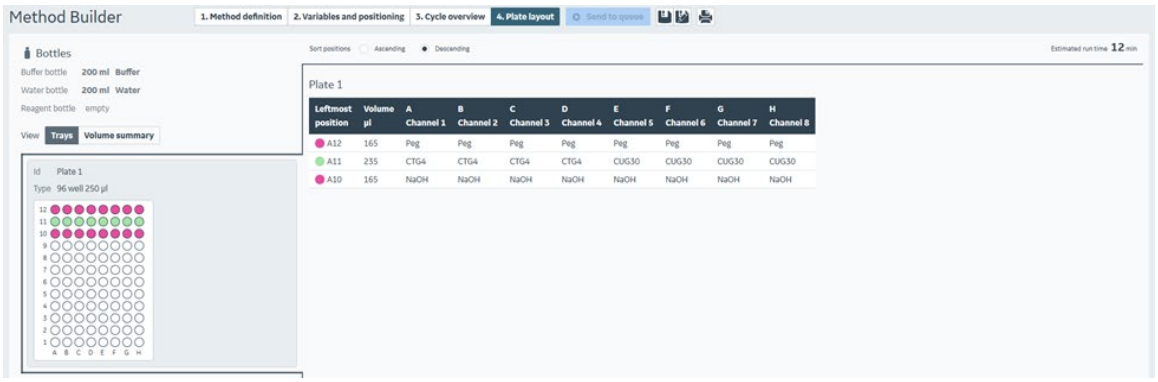


Figure 15. CUG30 RNA Oligomer Immobilization Method Protocol Parameters-Location and Volumes for CM5 Chip

Following successful RNA immobilization at necessary surface density, the buffer in the instrument was replaced with the assay buffer containing 5% DMSO and the system was primed. The system was then allowed to equilibrate for 60 minutes with the assay buffer prior to the start of the screen.

3. Ligand Preparation

   The fragment library of ligands were transferred to assay plates using the Echo acoustic dispenser. The 50 µM concentration was plated and run first. A protocol was created in the Echo Plate Reformat software to do a single point dispense of compound into each well of a 384-well assay plate. Buffer was added to the plate uniformly using the Certus liquid dispenser to reach the target ligand concentration, 100 µL of available sample, and 5% DMSO. Assay plates for the 500 µM screen were prepared similarly as were plates containing the controls and solvent corrections.

4. Dose Response Reconfirmation Study

   Buffer preparation and surface preparation were done exactly as the screening conditions. Fragments were tested in 10-point dose response with a top

concentration 1.8 μM and in 1.5-fold dilutions. All liquid handling was completed the same way between screening and the dose response confirmatory studies.

Computational Chemistry Virtual Screen

Virtual screening of the fragment library was completed using the Maestro software package produced by Schrodinger. The methods to complete the virtual screen were as follows:

1. Identify structure to use in docking simulation using RCS-PDB. This work utilized the structure of 1ZEV (Mooers et al., 2005)
2. Identify ligand library to utilize in docking simulation. The scope of this work utilized the internally available fragment library with associated virtual structures
3. Preparation of the ligand library
    i. Generate 3D structures
    ii. Add necessary Hydrogens to ligand structures to allow for Hydrogen bonding interactions to be accounted for in virtual screen
    iii. Set max ligand size to “under 500 Da” as the library is fragments which range in size from 150-300 Da with the average being 250 Da. This also leaves potential for medicinal chemistry iteration to add R groups to core molecule while keeping overall molecular weight under <500 Da
    iv. Allow for polar solvent accessibility

v. Select docking conditions for a neutral pH +/- 2

vi. Select enantiomers which have both R and S activity

vii. Deselect any enantiomers which have >3 chiral centers

4. Preparation of target

i. Oligomer must be prepared to allow for maximum number of interactions to occur

ii. Hydrogen optimization assignment must be completed

iii. Minimize and delete waters to allow for hydrogen atoms to be in the lowest density orientations

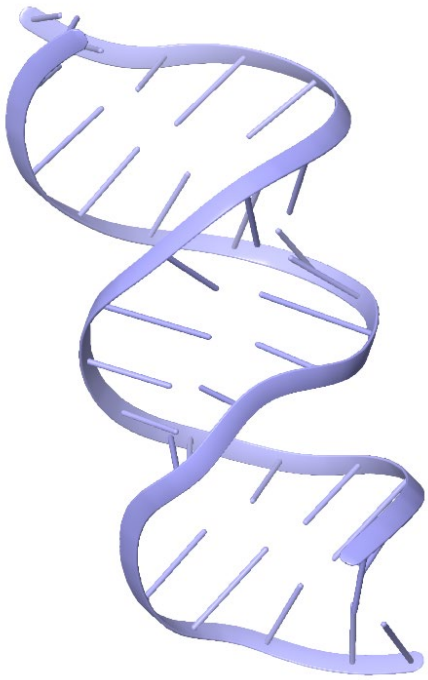


Figure 16. 1ZEV Structure After Cleaning and Docking Preparation

5. Generate grid

i. Grid was generated to constrain docking region to be between residues G9 and G10 on both chain A and B with a 10-angstrom dock length for the 1ZEV structure.

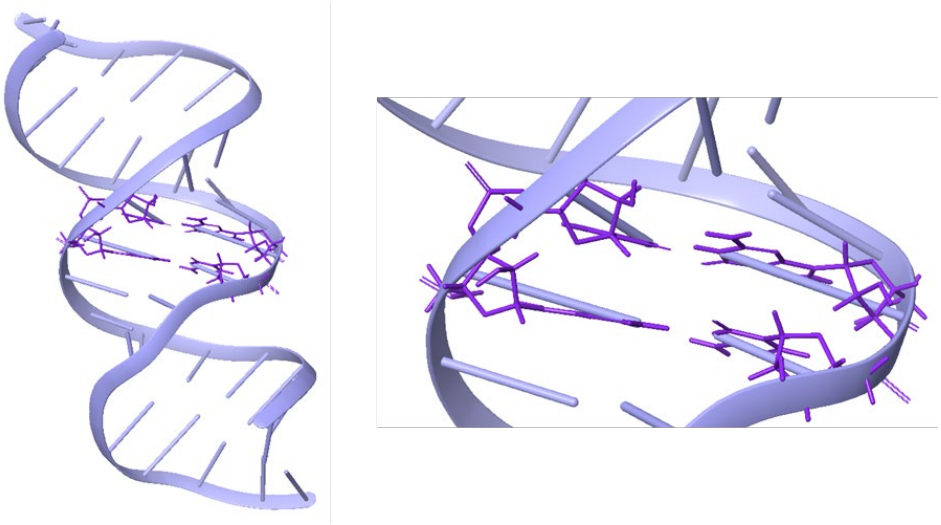


Figure 17. Residue Determination for Grid Generation

Chapter III.

Screen Execution

This section details how the screens were executed and the data analysis.

SPR Fragment Screen Data

1794 fragments were screened at 50 µM and 500 µM to capture a wide range of potency of the chemical matter. Fragments were first screened at 50 µM to characterize direct RNA binding, insolubility, and reference binding. Insoluble and reference binding fragments were removed from further experiments. All remaining fragments were then screened at 500 µM to characterize direct RNA binding, insolubility, and reference binding. Insoluble and reference binding fragments were removed from further analysis.

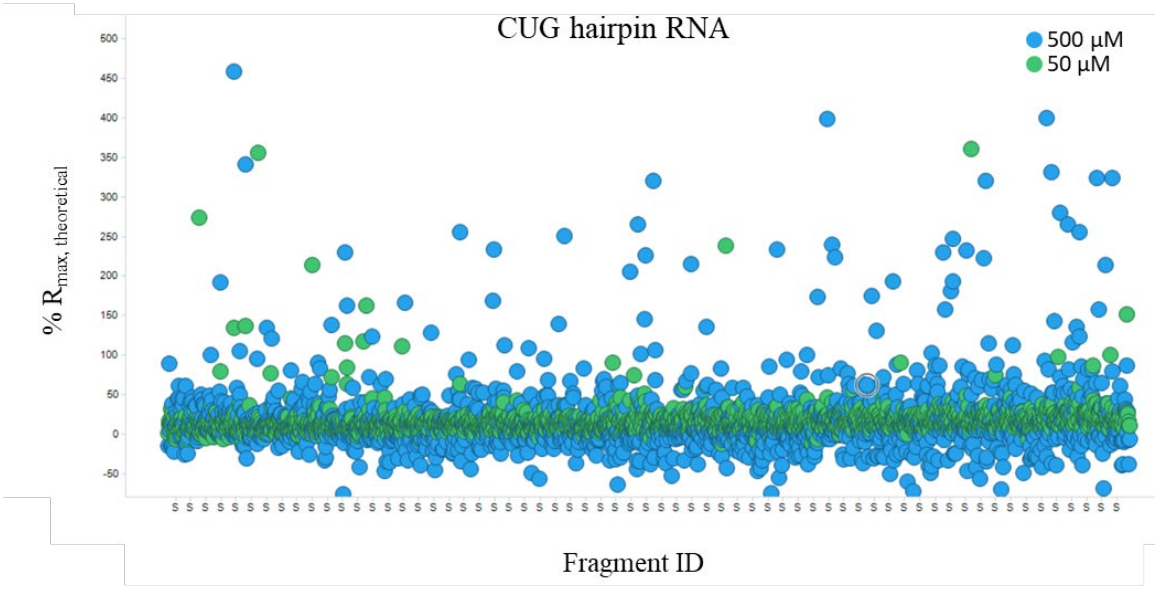
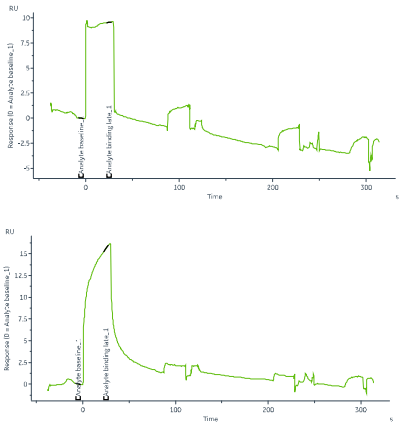
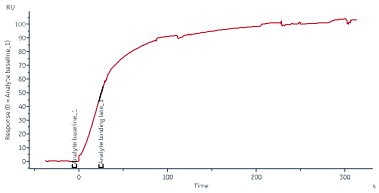
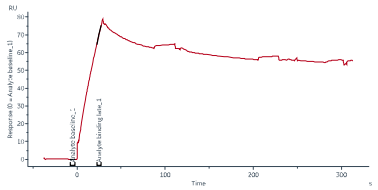
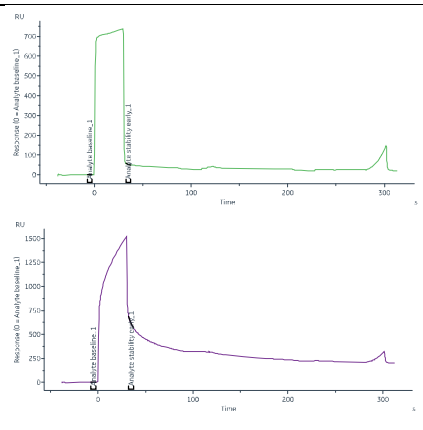
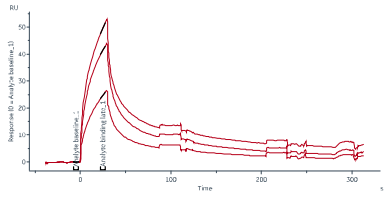


Figure 18. Binding Scatterplot for all Fragments at 50 µM and 500 µM Concentrations.

The fragment screen binding data was classified into five distinct types of fragment binding.

Table 9. Fragment Classification for Fragment Prioritization

Classification	Definition	SPR Sensorgram
Binder	Response is within expected Rmax range, fragment dissociates completely from surface, baseline is restored.	
Insoluble	Fragment is not soluble and does not dissociate from surface, baseline is not restored.	
Lack of Dissociation	Fragment binds to surface and does not dissociate form surface, baseline is not restored.	
Reference Binder	Fragment binds to the Neutravidin reference surface and exhibits target non-specific binding. Baseline may or may not be restored depending on fragment interaction with Neutravidin	

Other	At or slightly above expected Rmax, slow dissociation observed, baseline is restored.	

Each fragment bin was given a relative prioritization ranking to guide follow-up experiments. Fragments binned as “binders” were given first priority. Second priority was given to fragments in the “other” bin. Fragments falling in the “insoluble”, “lack of dissociation”, and “reference binder” bins were deprioritized for further experiments.

Fragments in the “binder” bin were then further evaluated for hit determination. Hits were defined as fragments exhibiting binding response above three standard deviations above the mean response and which exhibited an $R_{max,\ theoretical} > 30\%$. Hits were further categorized based on qualitative assessment of the sensorgram shape as either “fast on fast off” or “complex binding”. “Fast on fast off” fragment sensorgrams showed a near horizontal curve in the association phase and an immediate decline in response to baseline at the end of the association phase. The observed response in these types of sensorgrams is concentration dependent and is indicative of the curves quicky reaching equilibrium. Fragments that exhibited “complex binding” showed curvature in the association phase of the sensorgram and an off rate while dissociating from the surface after the injection end.

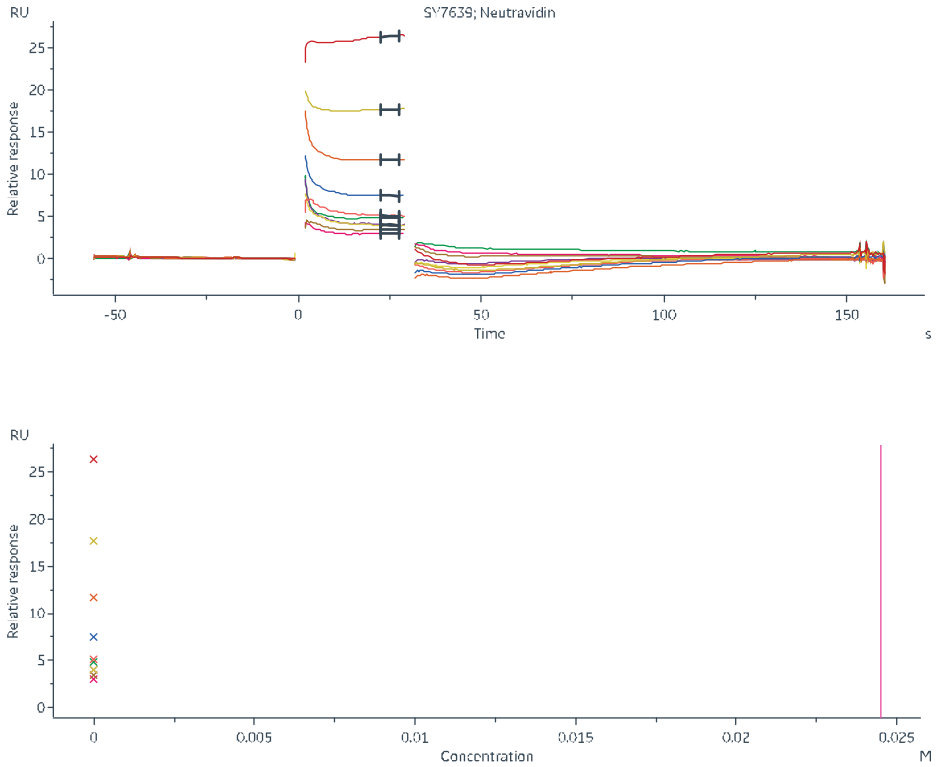
In the 50 µM concentration portion of the screen, 120 fragments showed direct RNA binding giving a 6.68% hit rate. At the 500 µM concentration of the screen, the data showed 300 fragments directly binding to the RNA hairpin thus giving a 17.1% hit rate. Given the known flexible nature of surface-bound oligonucleotides and known non-

specificity of fragments, the hit rates at both screening concentrations were considered acceptable for a single point screen. Gianetti et al publish that between 3-15% hit rates on average is an acceptable range.

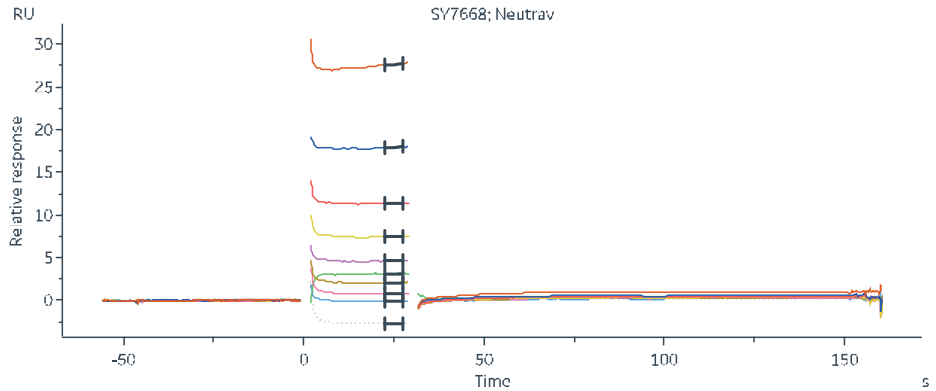
37 fragments bound the RNA hairpin oligo at both concentrations. 83 fragments bound the RNA oligo at only the 50 μM concentration and 263 fragments bound the RNA at the 500 μM concentration. To confirm these single point data, the identified fragments were tested in 10-point dose dependence confirmatory studies. The first confirmatory study was done with the same fragment stocks used in the single point screen. The second confirmatory study was done with resynthesized fragment stocks. The purpose of the two-step confirmatory studies were to filter out false positives that bound at the single concentrations and any artifactual binding due to impurities in the original fragment stock.

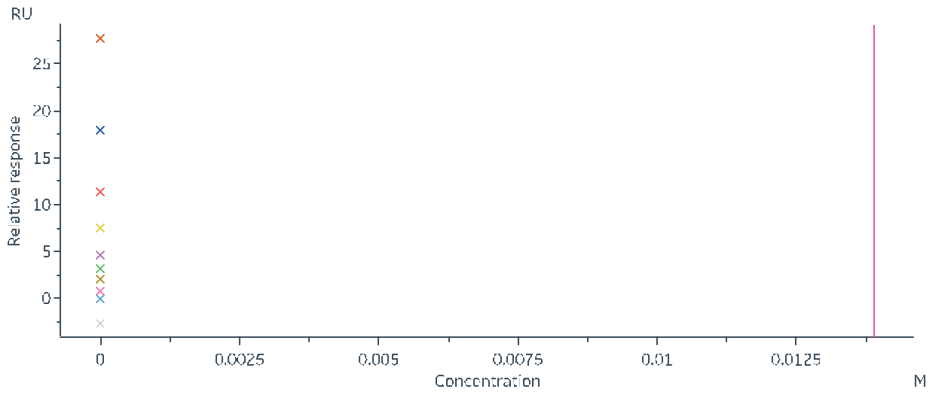
The data generated from the first confirmatory study showed 43 fragments exhibiting dose-dependent binding with fast on fast off kinetics observed, and 27 fragments exhibiting complex binding. The data generated from the second confirmatory study showed 24 fragments exhibiting dose-dependent binding with fast on fast off kinetics and 7 fragments exhibiting complex binding. This resulted in a total of 31 confirmed fragment hits from the SPR-based fragment screen. The data for these confirmed hits is shown below.

Fragment 1: SY7639

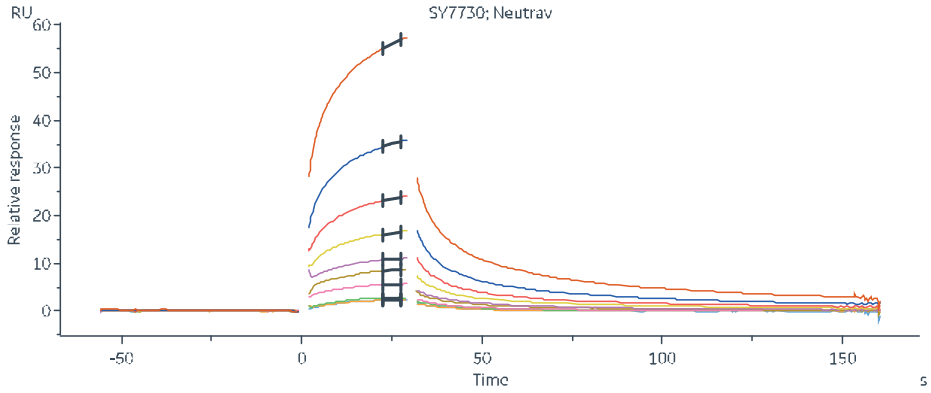


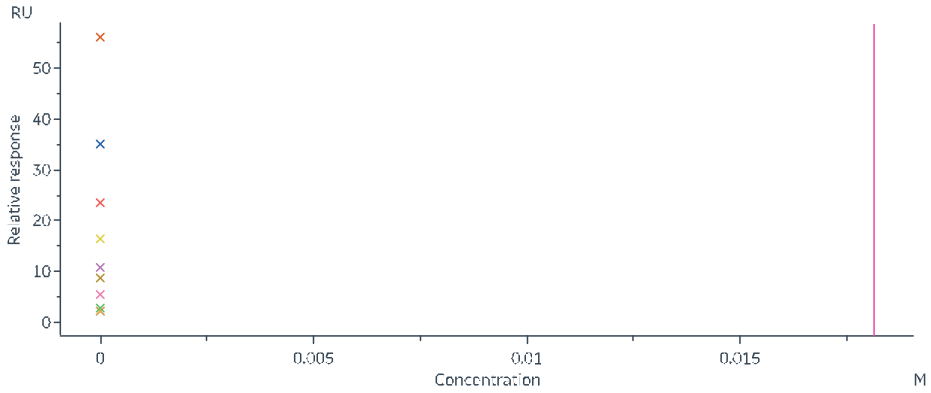
Fragment 2: SY7668



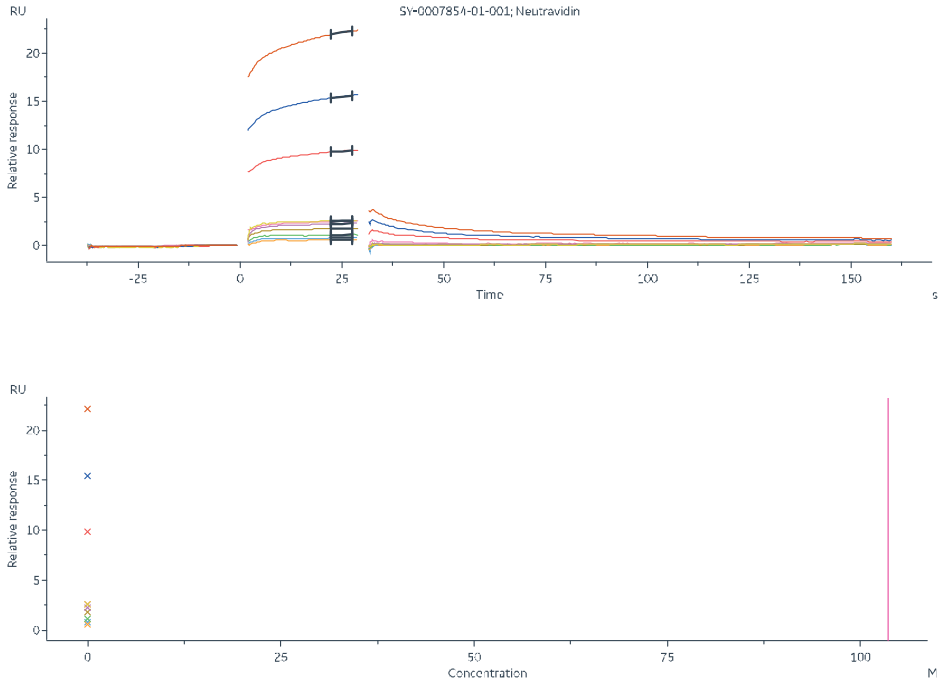


Fragment 3: SY7730

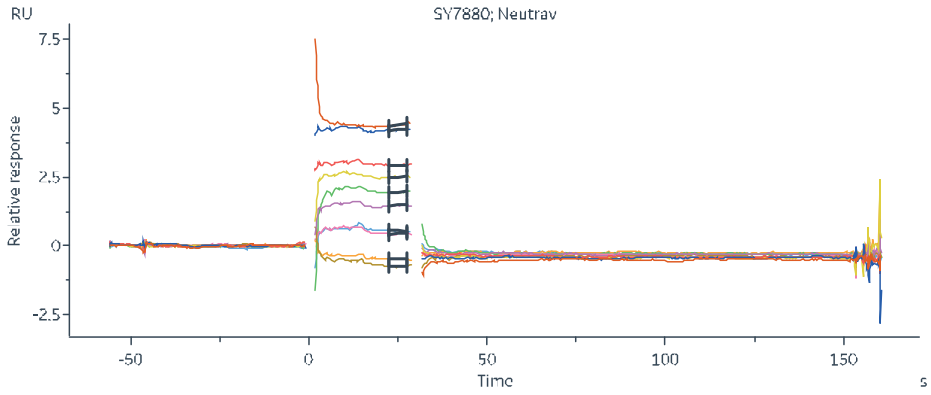


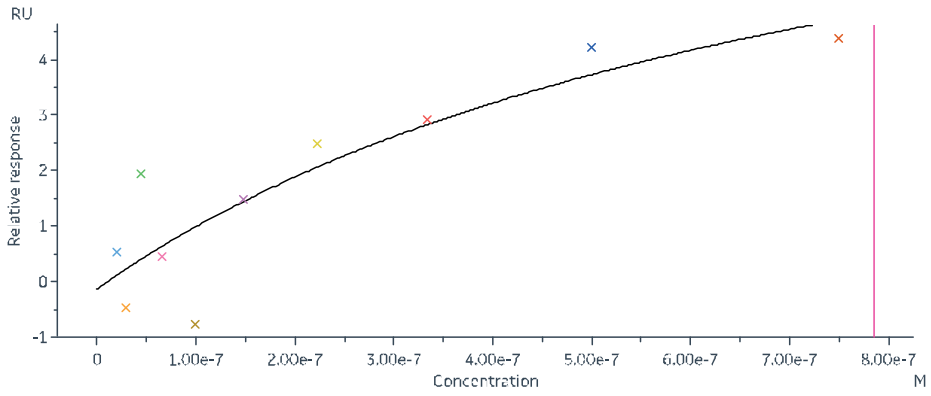


Fragment 4: SY7854

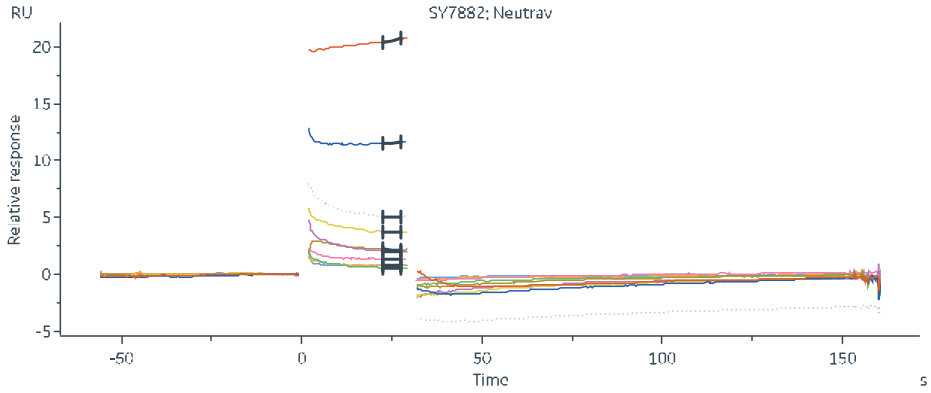


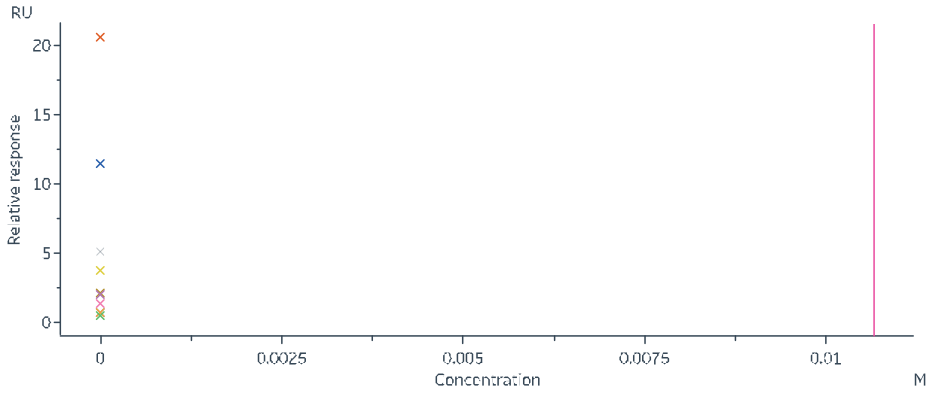
Fragment 5: SY7880



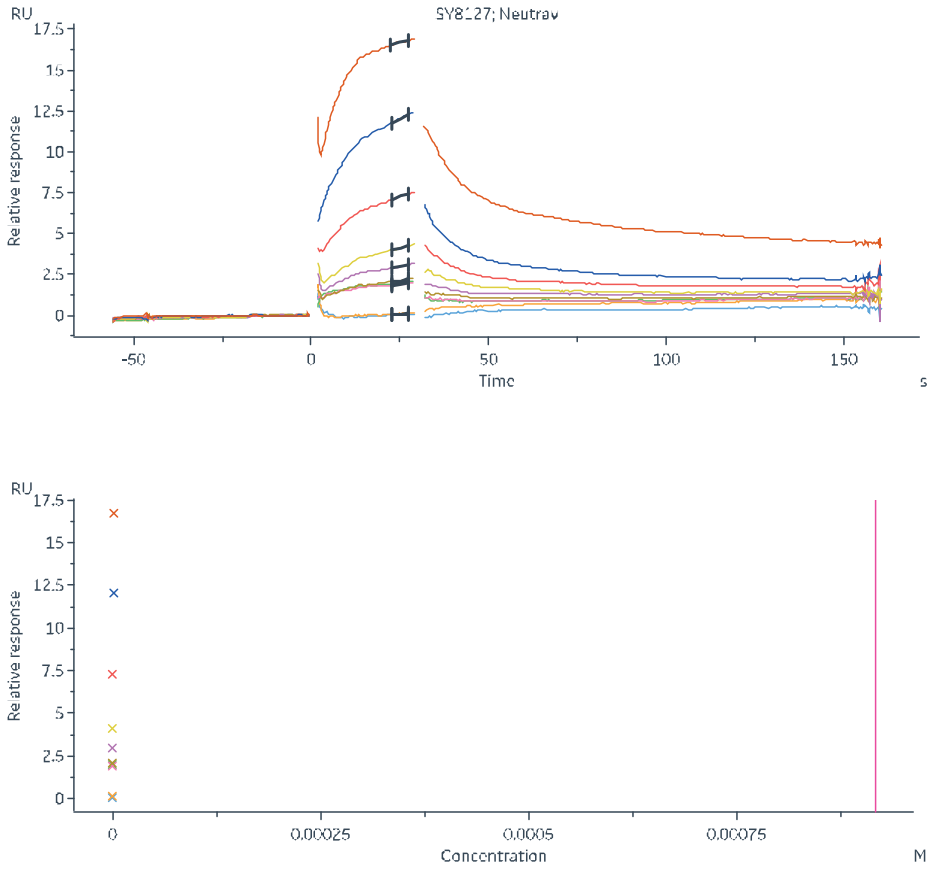


Fragment 6: SY7882

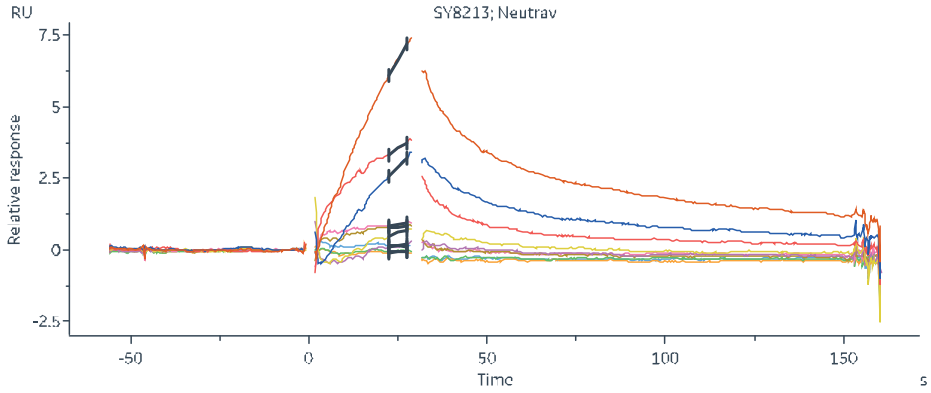


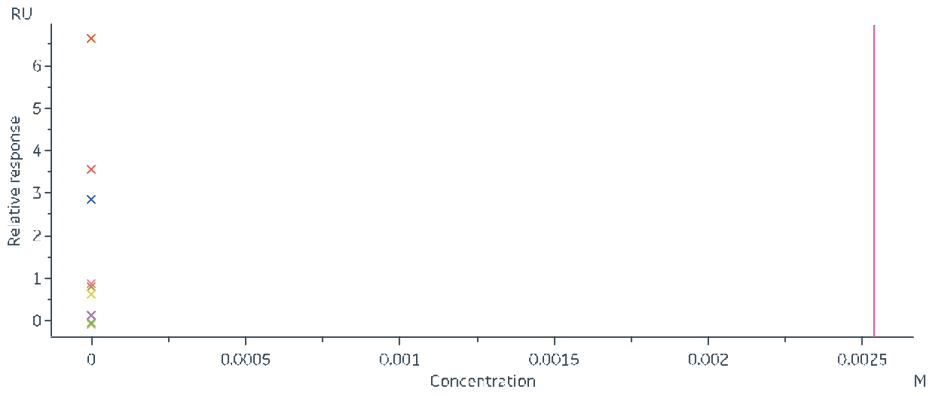


Fragment 7: SY8127

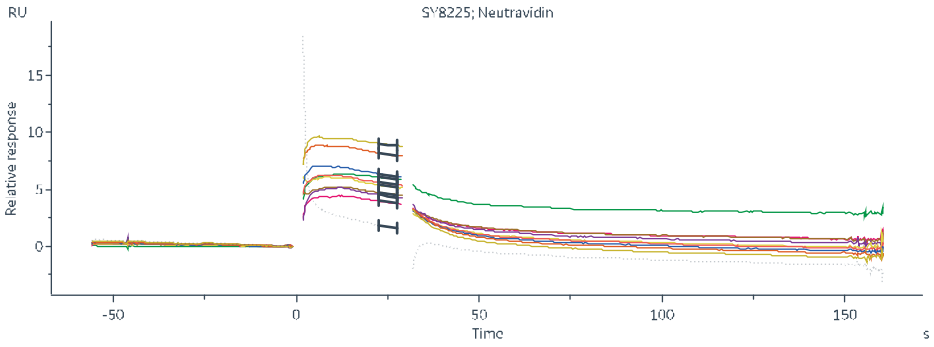


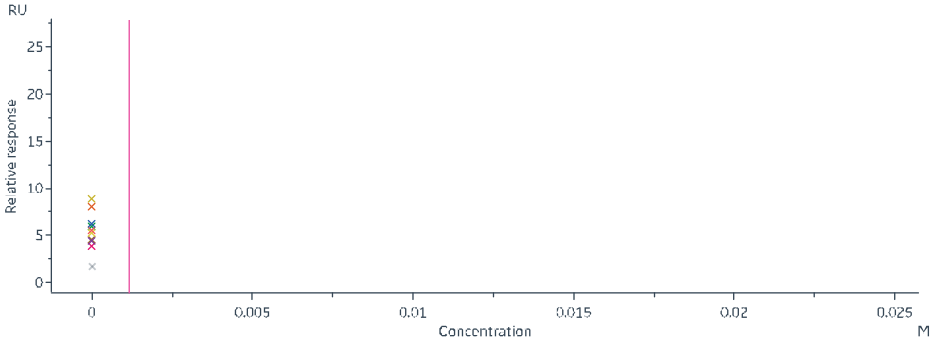
Fragment 8: SY8123



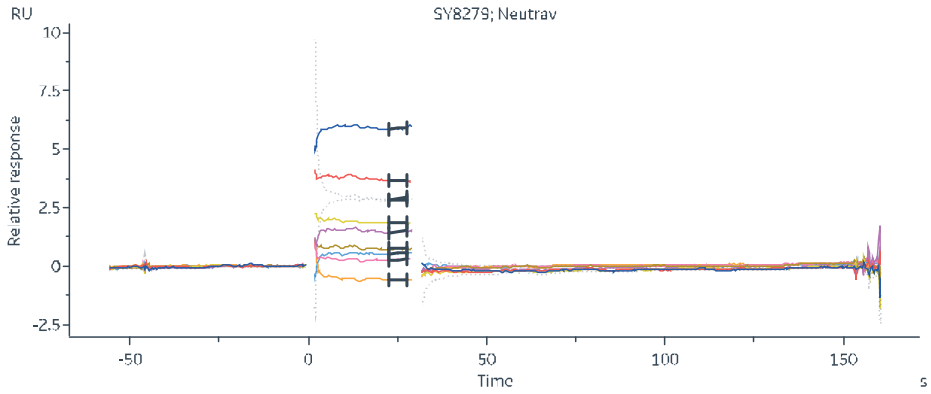


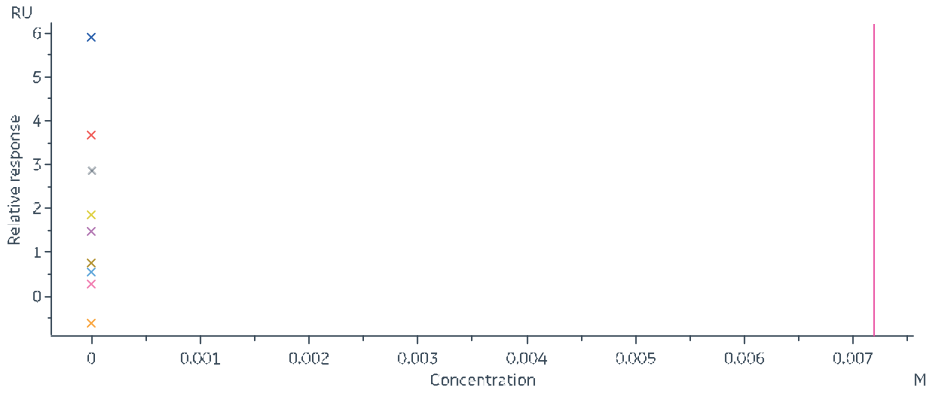
Fragment 9: SY8225



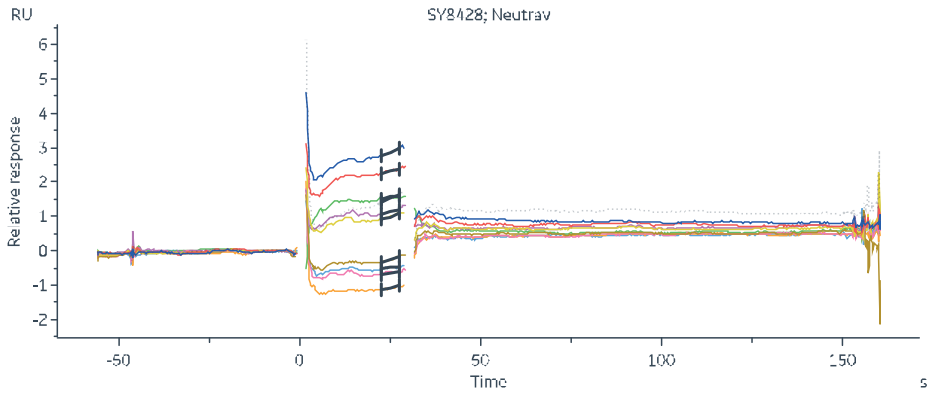


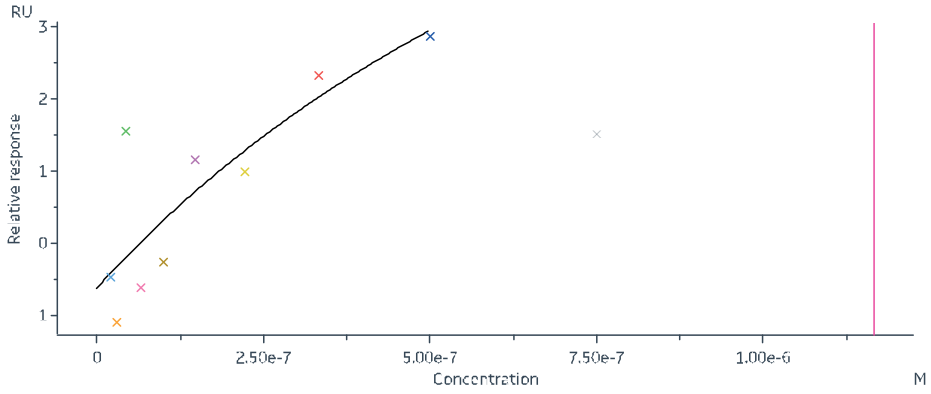
Fragment 10: SY8279



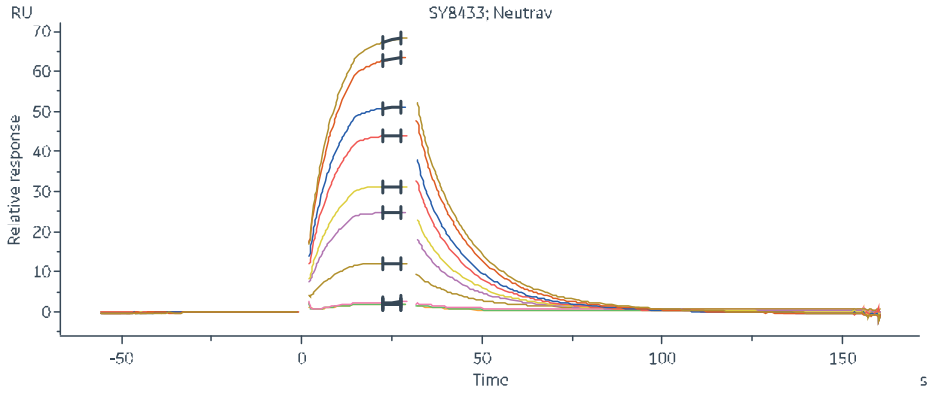


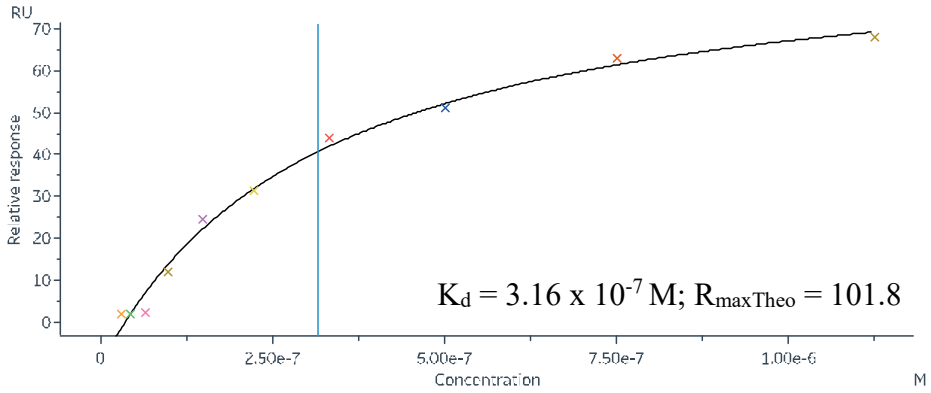
Fragment 11: SY8428



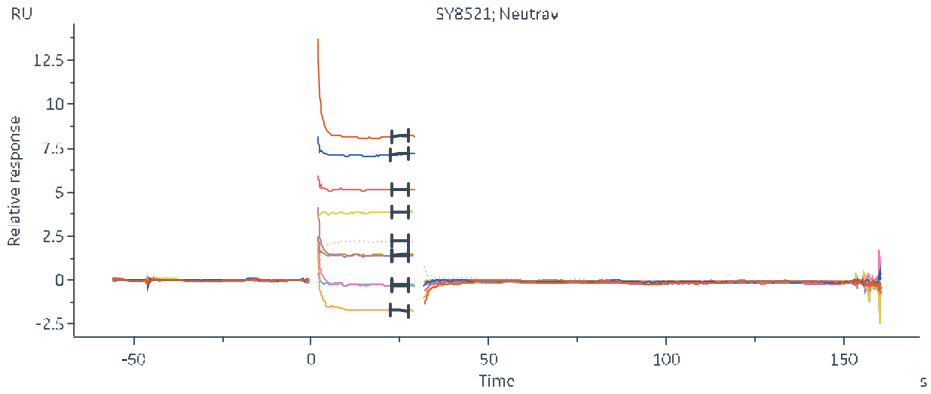


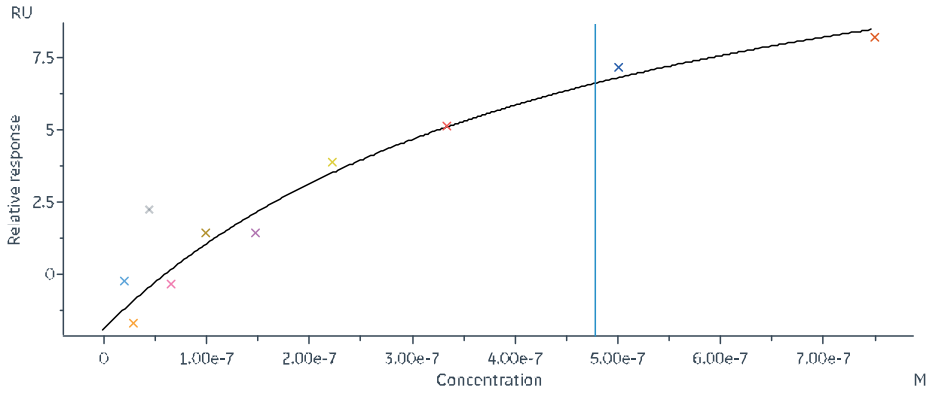
Fragment 12: SY8433



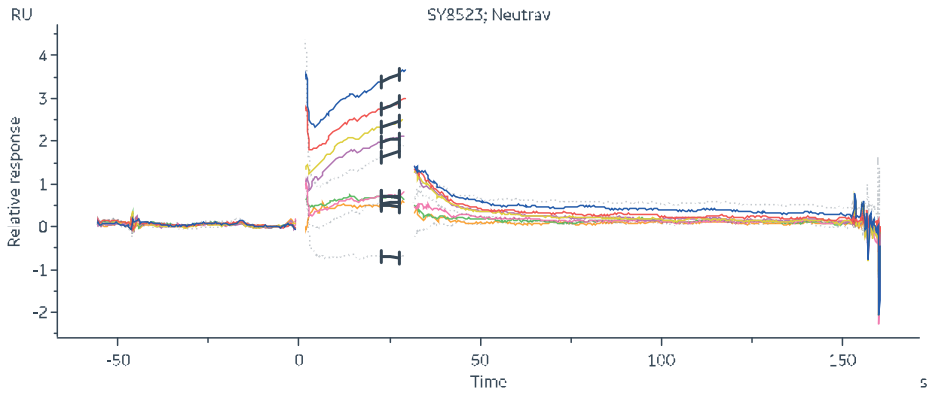


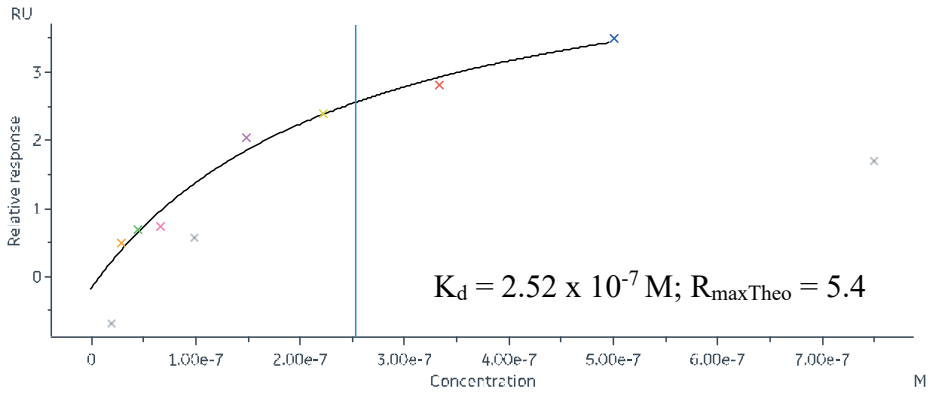
Fragment 13: SY8521



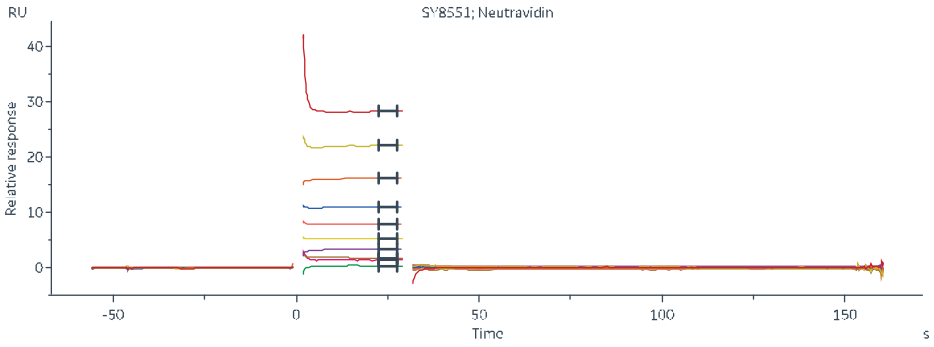


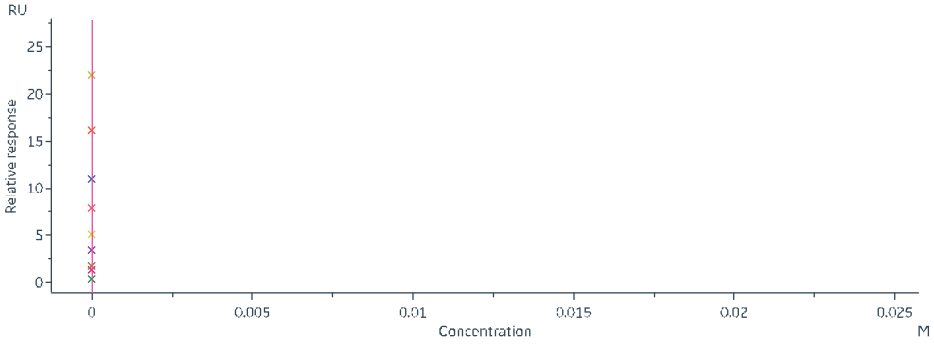
Fragment 14: SY8523



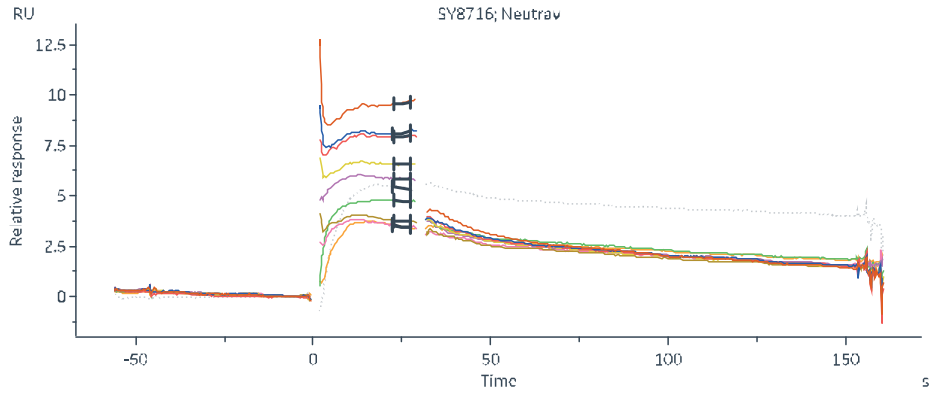


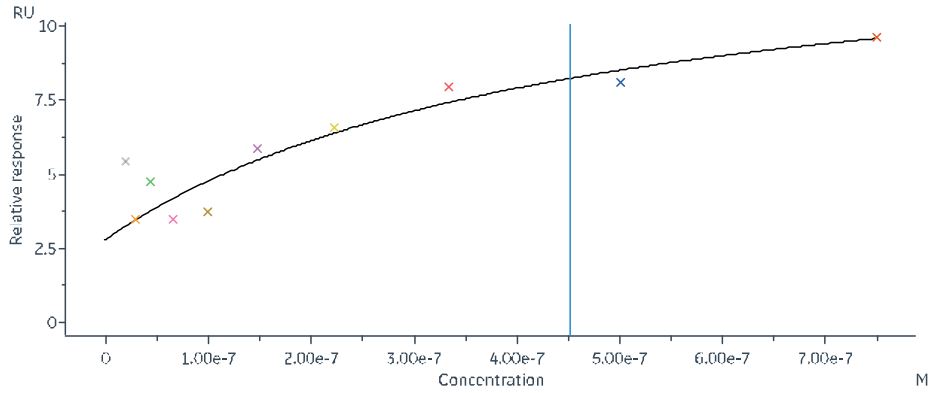
Fragment 15: SY8551



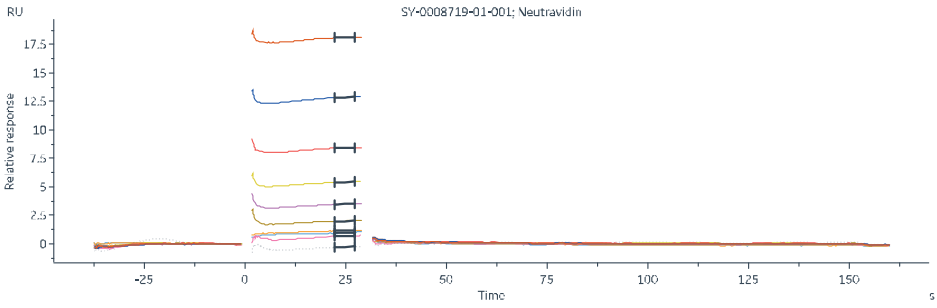


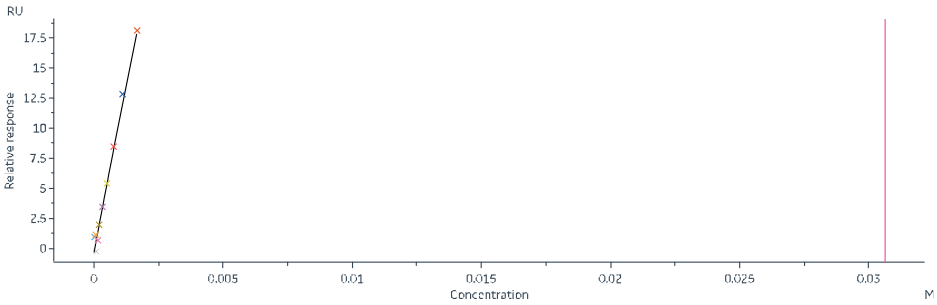
Fragment 16: SY8716



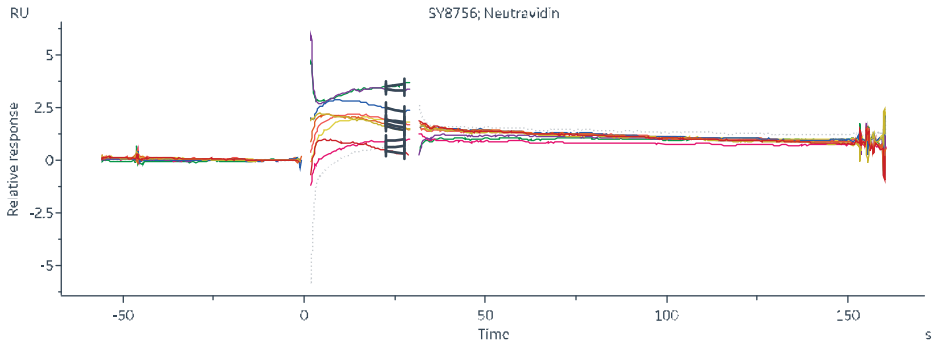


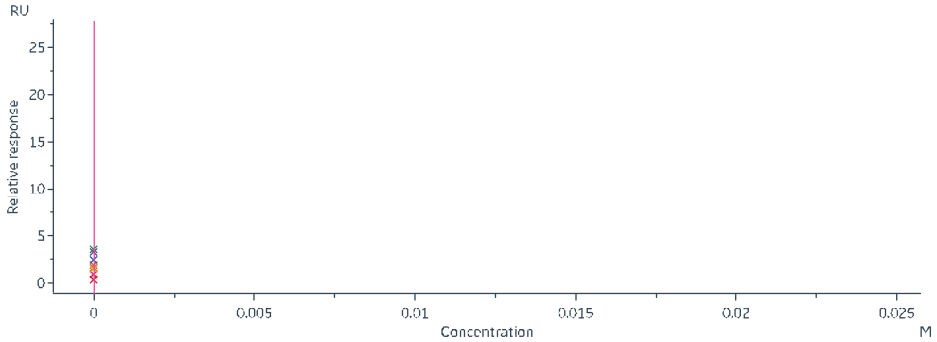
Fragment 17: SY8719



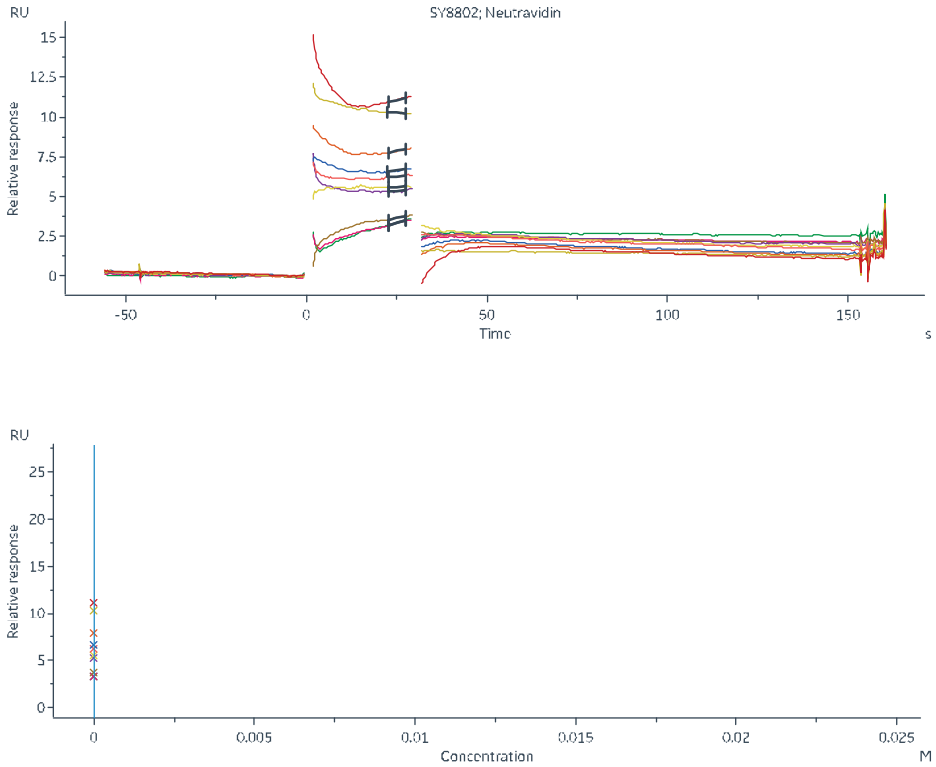


Fragment 18: SY8765

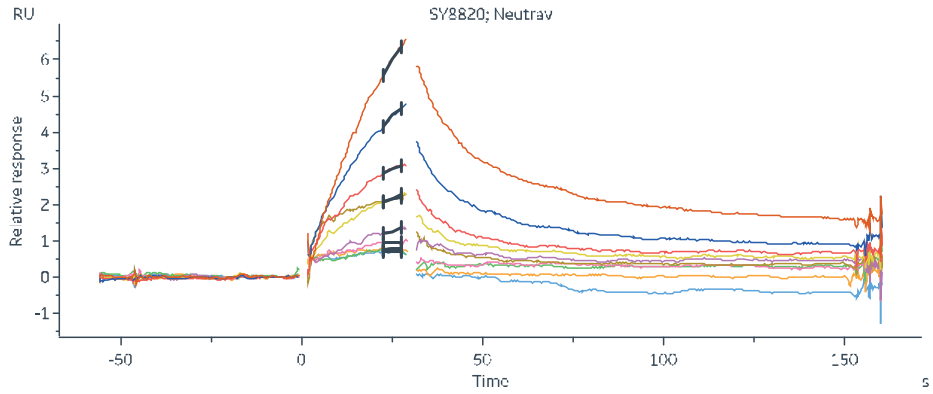


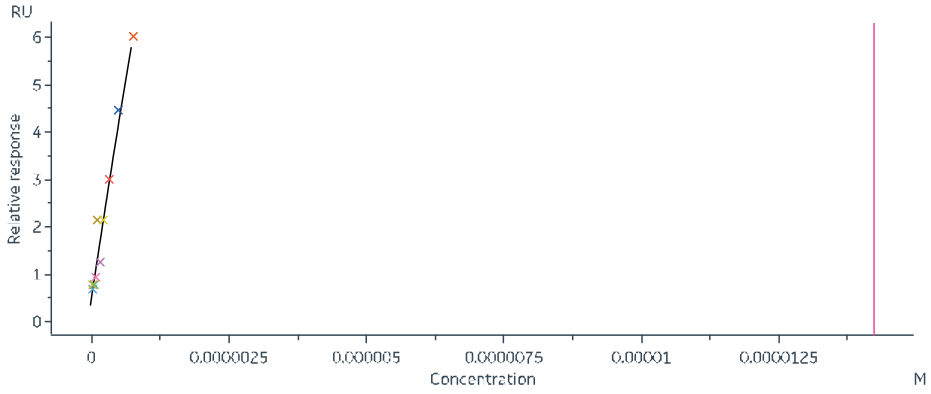


Fragment 19: SY8802

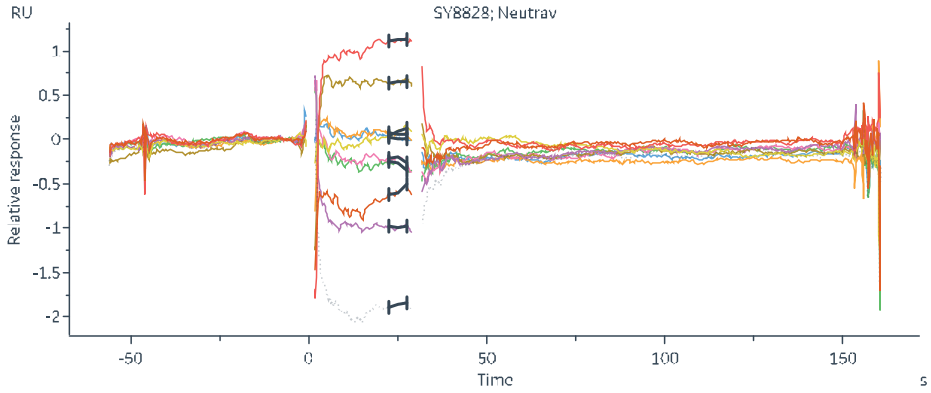


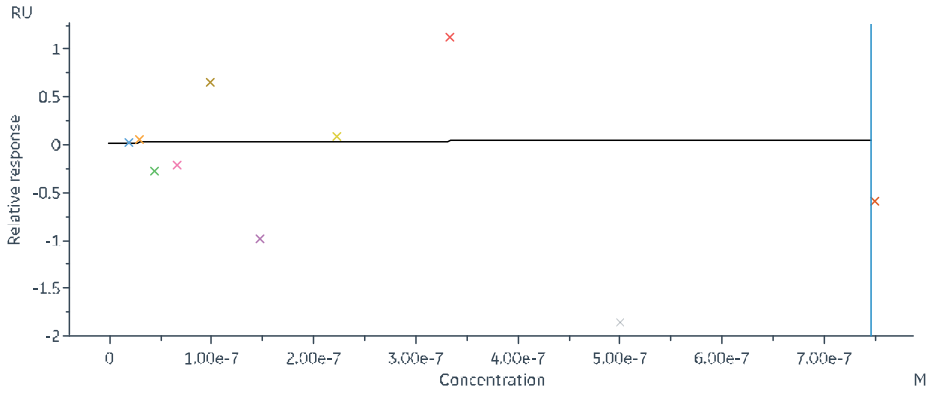
Fragment 20: SY8820



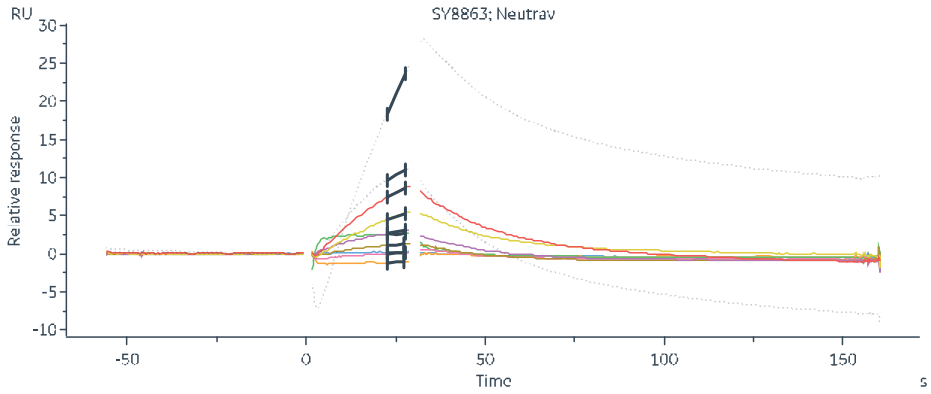


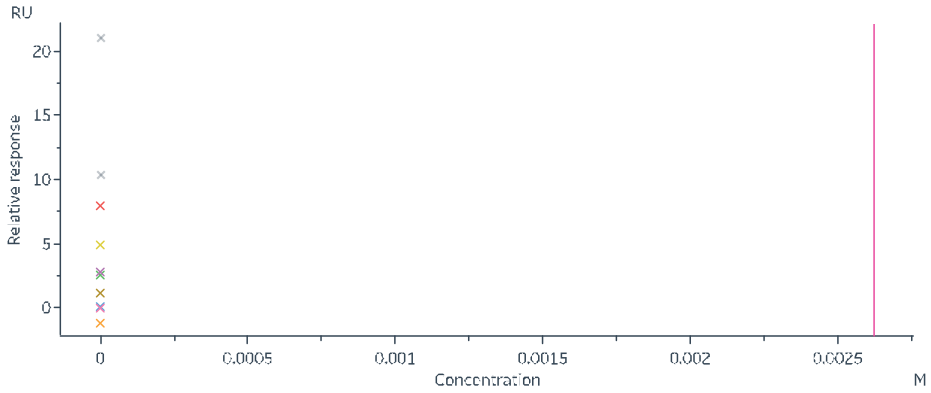
Fragment 21: SY8828



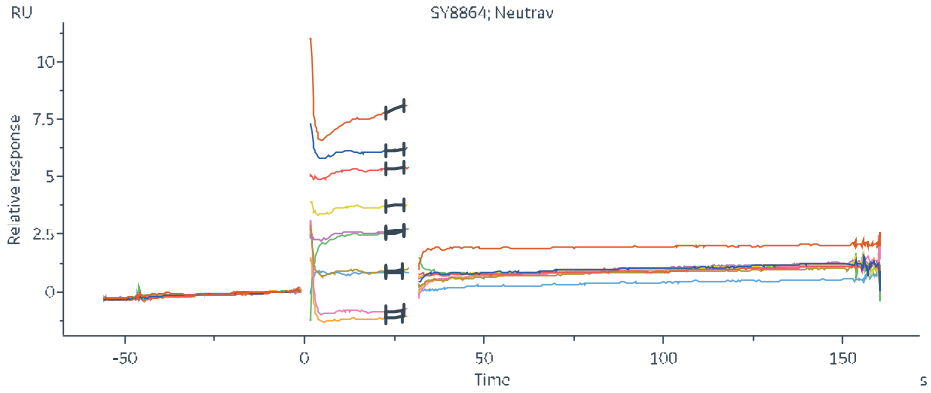


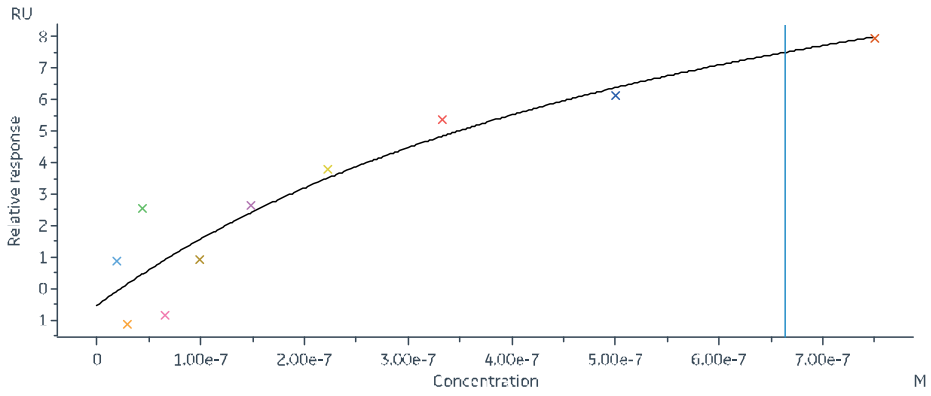
Fragment 22: SY8863



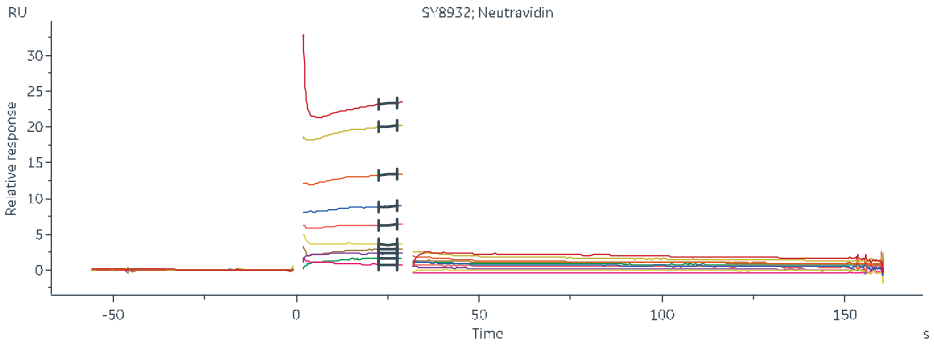


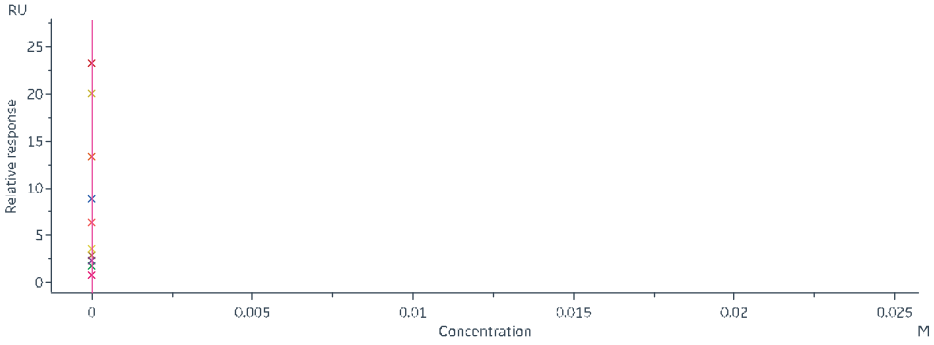
Fragment 23: SY8864



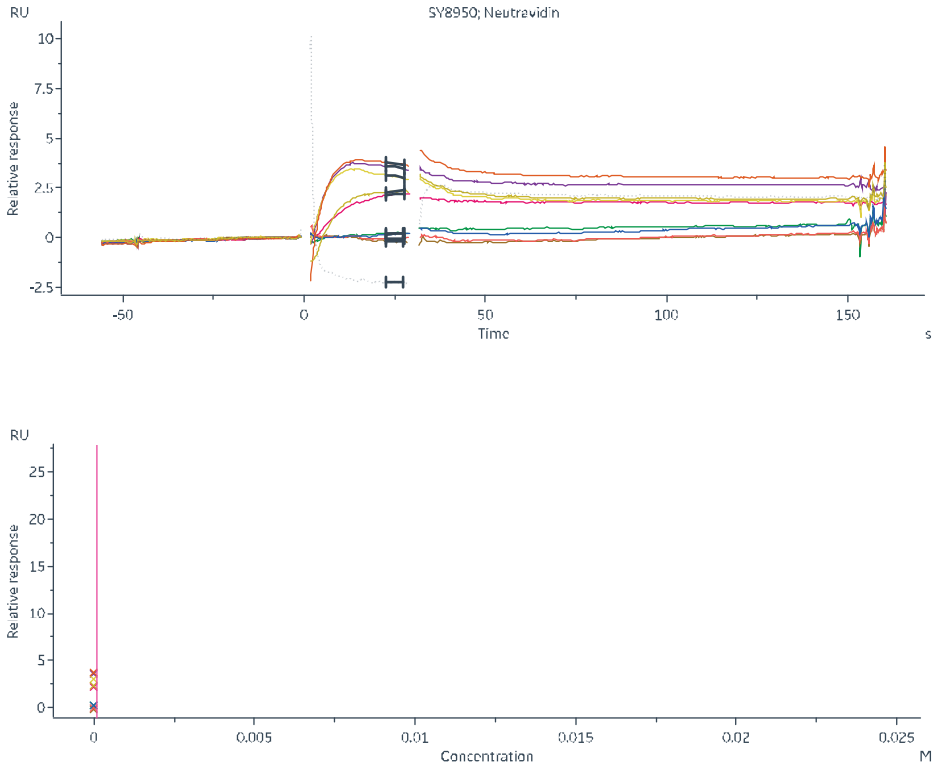


Fragment 24: SY8932

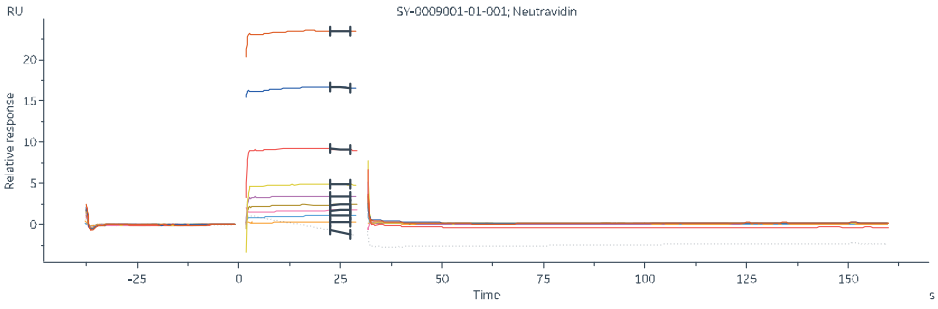


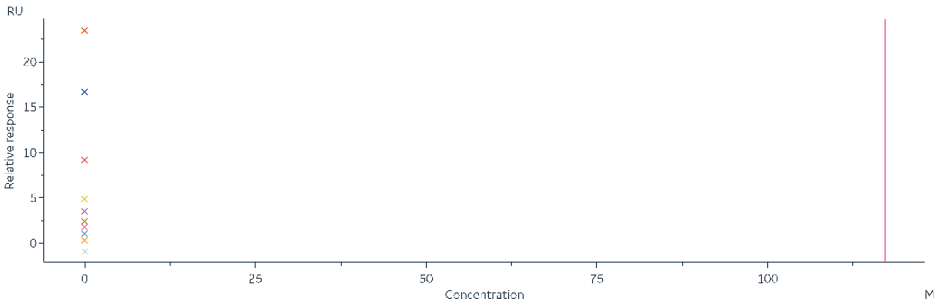


Fragment 25: SY8950

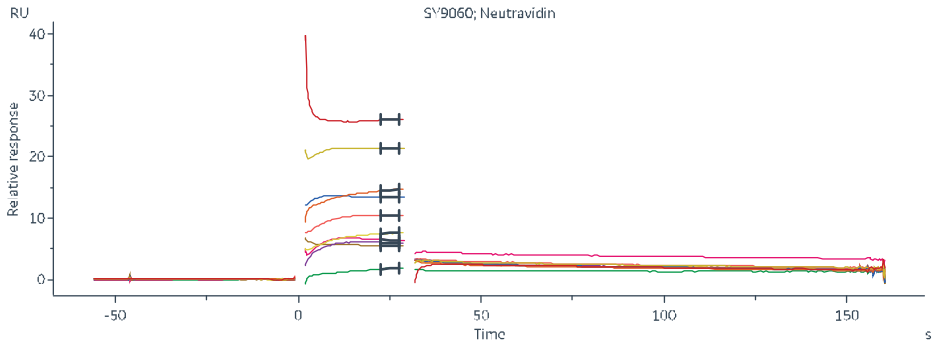


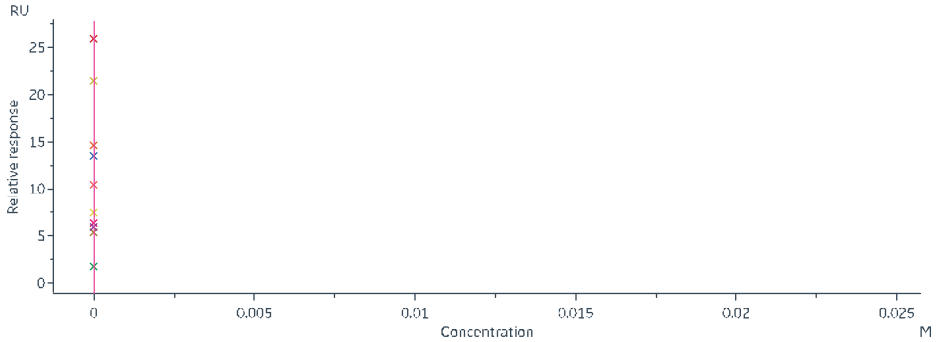
Fragment 26: SY9001



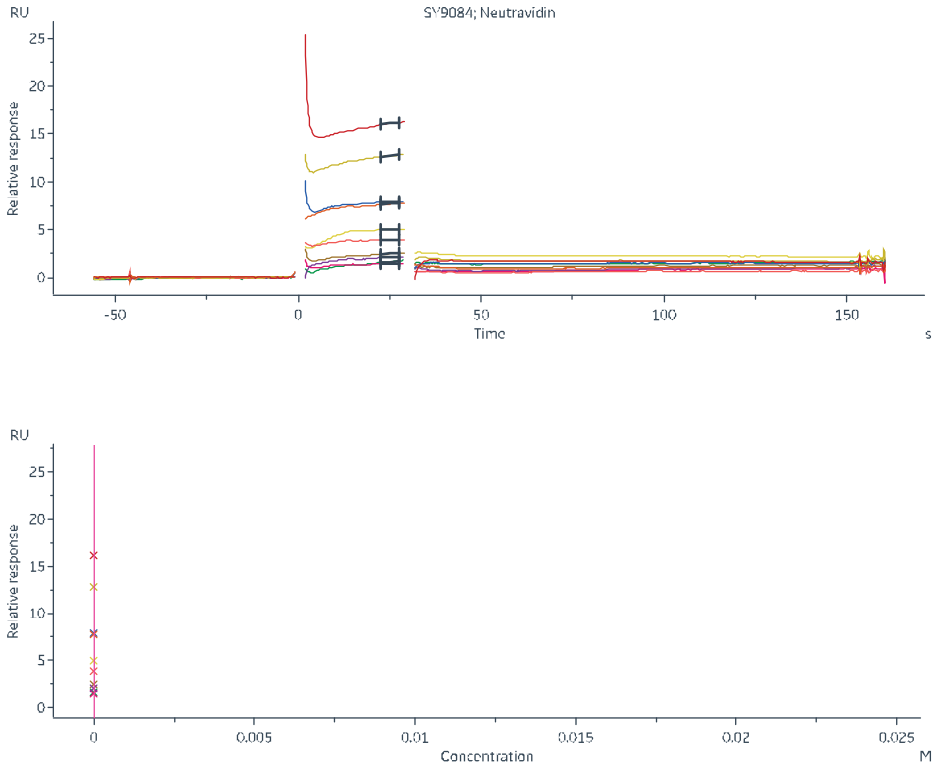


Fragment 27: SY9060

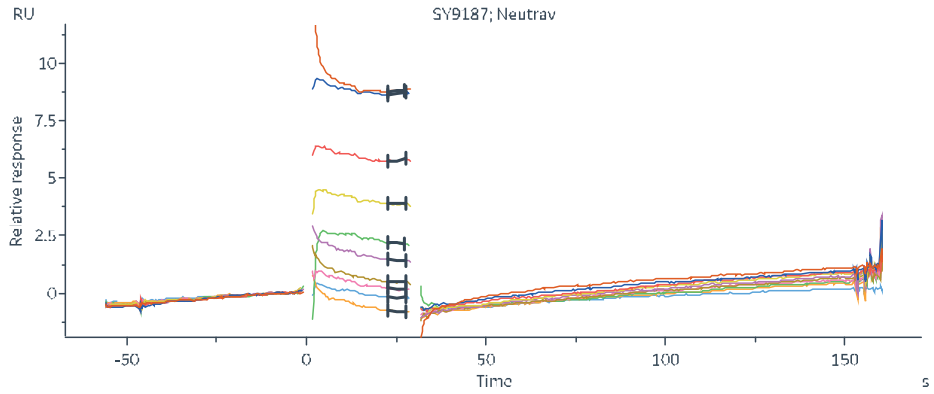


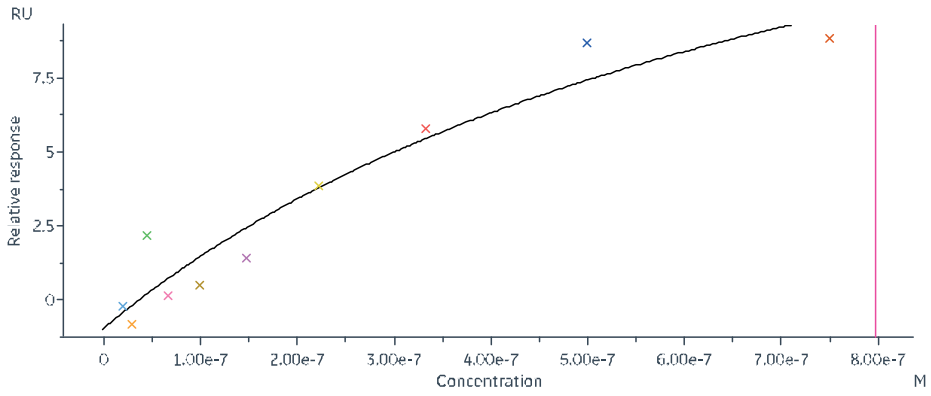


Fragment 28: SY9084

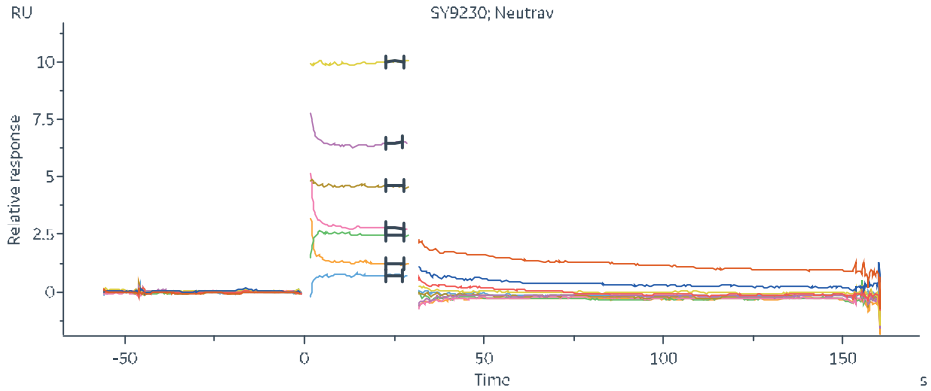


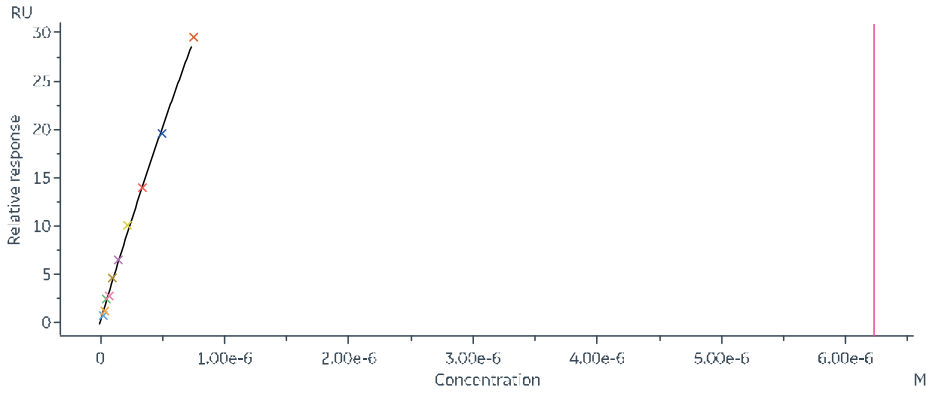
Fragment 29: SY9187





Fragment 30: SY9230





Fragment 31: SY9237

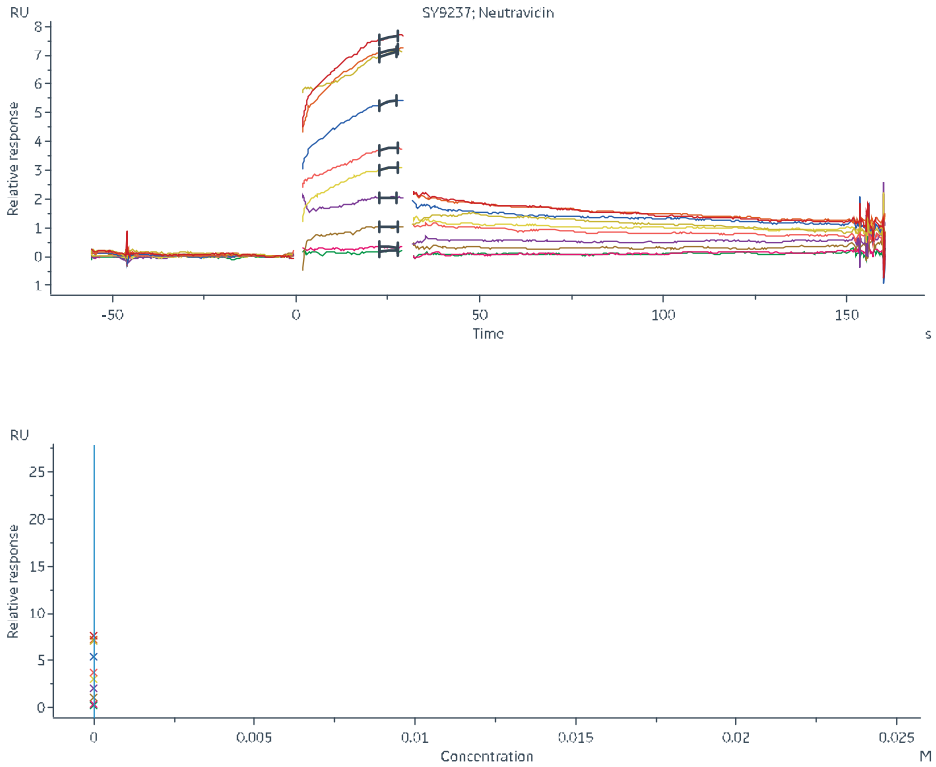


Figure 19. Dose Dependent Binding Sensorgrams and Affinity Plots for SPR Screen Fragment Hits

Fragments 12 and 14 were able to be fit with a 1:1 binding model and resulted in acceptable affinity K_d measurements. Other fragment sensorgrams were unable to be fit with kinetic or affinity K_d parameters for a 1:1 binding model. This suggested that there was potentially complex binding occurring wherein the RNA was bound by the fragment and then a second fragment then bound to the formed complex.

Taking a deeper look at the structures of the 31 fragments, some recurring features were observed. The most common shared functional groups were, thiazole, pyrazole, pyridine among the hits.

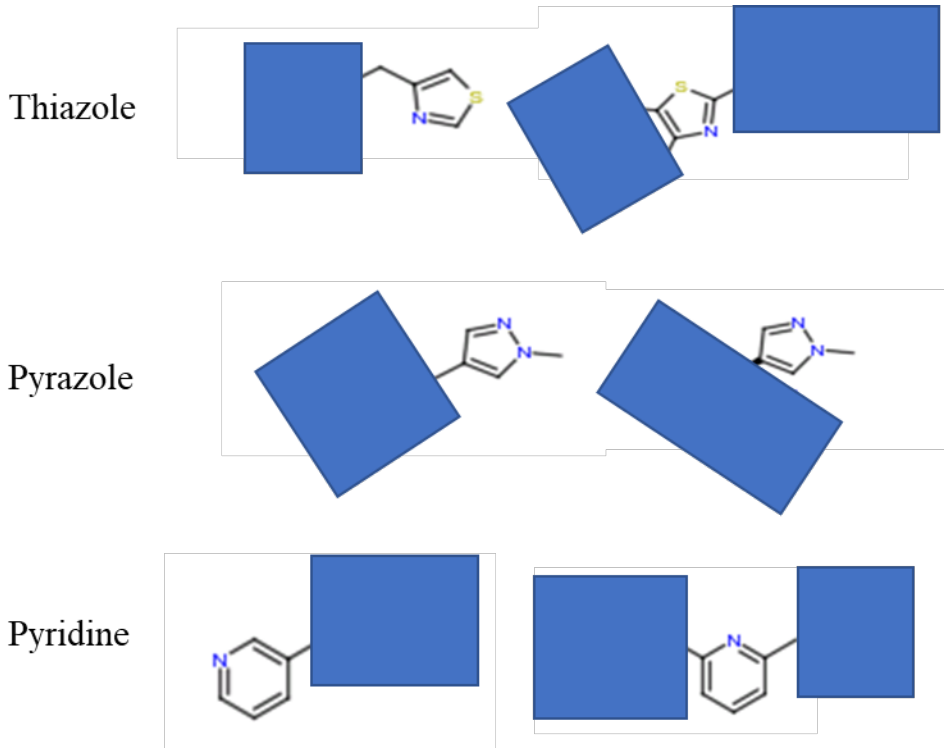


Figure 20. Blinded fragments show structural diversity.

The SPR biophysical screen began with 1794 unique fragments and resulted in the experimental determination that 31 fragments show dose-dependent direct binding to the 30mer RNA hairpin oligomer.

Virtual Fragment Screen Data and Discussion

Virtual screening was completed on the 1ZEV structure to allow for modeling of a well-defined RNA hairpin structure with expected base pairing.

Figure 21. Crystal Structure of a Pathogenic RNA: CUG Repeats; PDB: 1ZEV, Mooers et al, 2005

Hit calling for the virtual screen had several evaluation criteria as outlined by Bender et al in their 2021 Nature protocols publication. This includes evaluation of the docking score, broken molecules, internal strain, interaction partners, unsatisfied hydrogen bond donors and acceptors, novelty filters, visual inspection, and scaffold clustering. The purpose of hit calling filters is to categorize and evaluate fragments based on both positive and negative features. Positive features may include fragments that

indicate binding to key residues or fragments able to bind in a favorable orientation. Negative features include indication of a strained docking orientation or buried hydrogen bonds. Fragment categorization according to standardized filters is necessary to determine a feasible number of hits to evaluate in confirmatory experiments.

Glide and docking scores are both standard parameters utilized to determine the ligand binding mode and estimated binding energy of favorable binding modes respectively. Glide scores are calculated by summing the dock score with Epik state penalties of tautomeric states. Based on literature precedent and the scope of this study, dock score and visual inspection were used as the primary hit calling criteria.

The virtual screen library contained 5871 unique conformers of the 1794 fragments in the Enamine Golden Fragment Library. 2933 conformers had measurable docking scores.

Analysis of all 2933 conformers resulted in the following docking statistics:

Table 10. Analysis of Docking Statistics of the Virtual Fragment Screen

Statistic	**Docking Score**
Mean	-4.058
Standard Deviation	1.38
Range (3X Standard Deviation)	-8.13 - -0.022

Primary hit determination was done by analyzing the conformers which had docking scores above three standard deviations above the mean as well as docking scores above the mean response. 3 conformers had docking scores above three standard

deviations over the mean. 1524 conformers had docking scores above the mean. This study intended to be inclusive of a wide range of fragments hence the 1524 conformers were determined to be the primary virtual screen hits. Based on literature precedence, the primary hit list was refined to only include hits with docking scores of <-7. This list of 45 conformers contained 37 unique hits and 8 hits which each had two separate conformers with docking scores <-7. This hit list was then visually inspected to determine if the fragments were binding within the expected intercalation site between C10 and G9. Visual inspection became a critical tool in final hit list determination as it allowed for the removal of fragments with acceptable docking scores but with contacts made outside of the intercalation site.

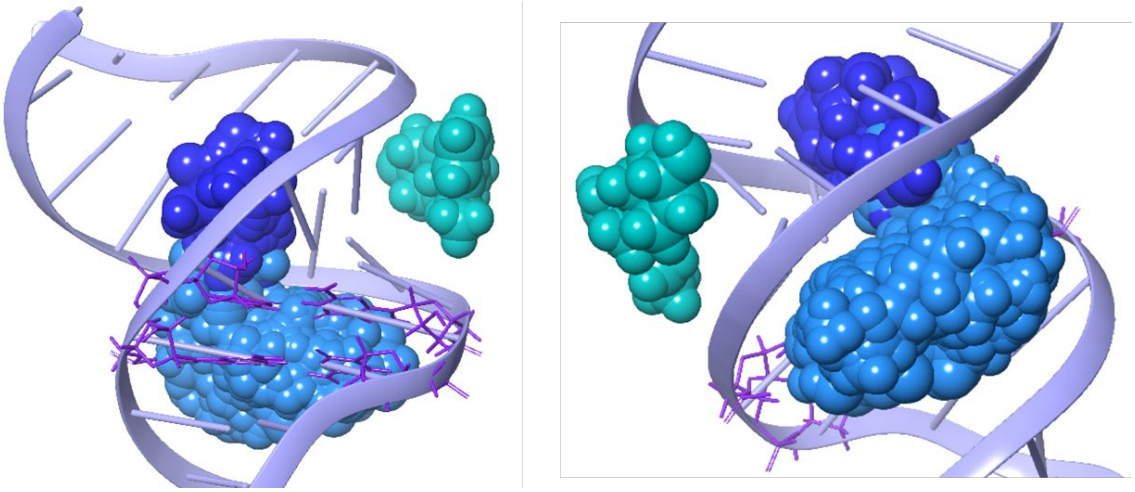


Figure 22. Docking of 45 primary hits to the 1ZEV RNA hairpin structure.

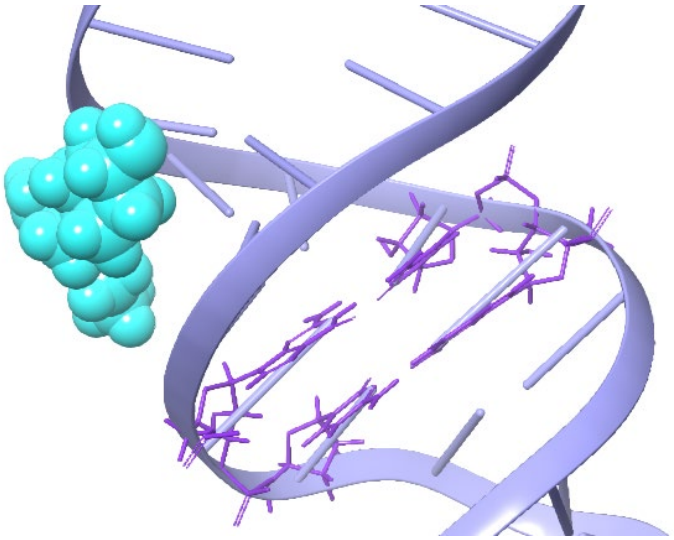


Figure 23. Fragment shows a docking score of -7.69 and makes contacts with the RNA structure outside of the intercalation site between C10 and G9.

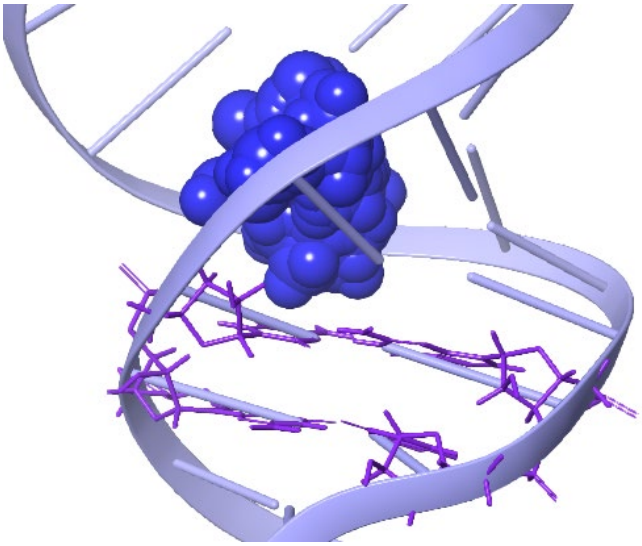


Figure 24. Fragment shows a docking score of < -7 and makes minimal contacts with the intercalation site of C10 and G9.

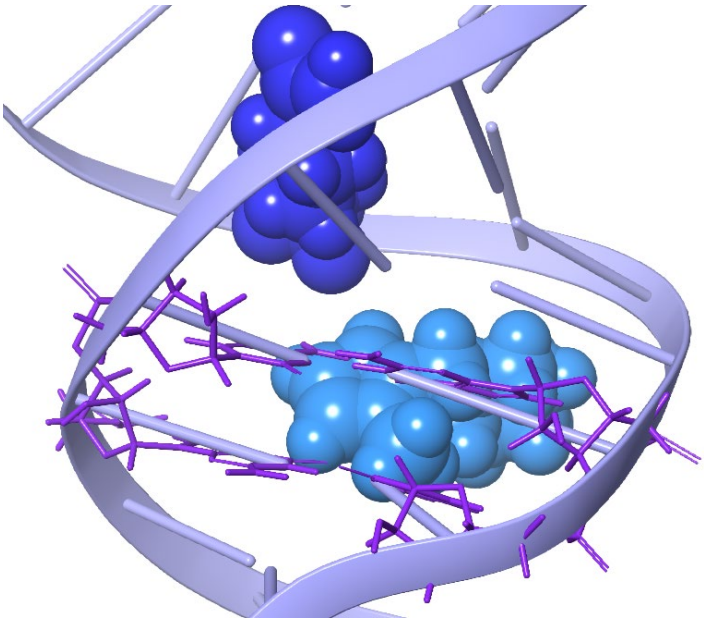


Figure 25. Two displayed fragments show the difference in intercalation site contacts through both have dock scores of < -7. The light blue fragment shows hypothesized contacts between C10 and G9 while the dark blue fragment shows major contacts between C13 and G12.

After utilizing dock scores and a primary filter and visual inspection as a secondary filter, this results in the determination of a refined hit list of 29 fragments with a range of dock scores from -7.051 to -8.234.

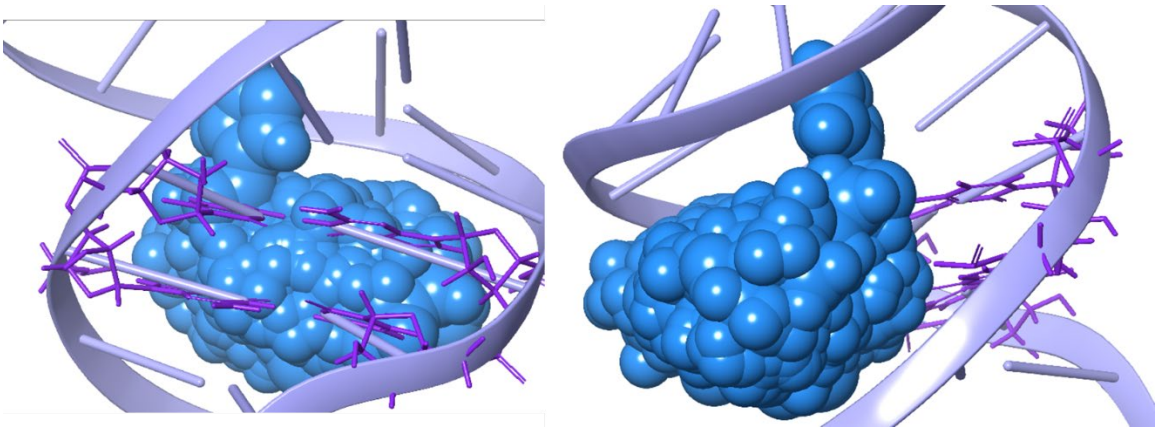


Figure 26. Refined hit list of 28 fragments docked to the C10 and G9 grid region.

Table 11. Classification of Binders

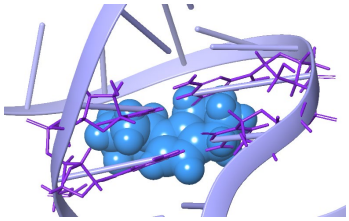
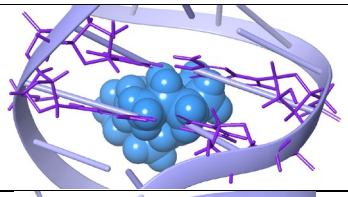
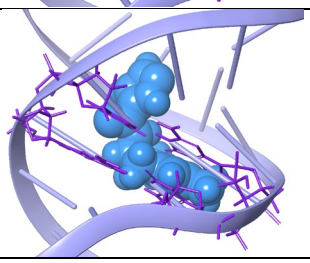
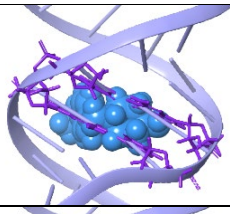
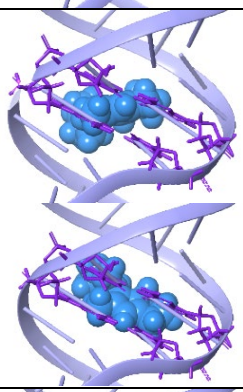
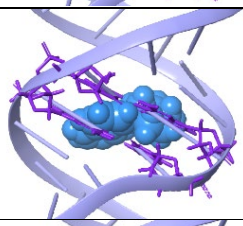
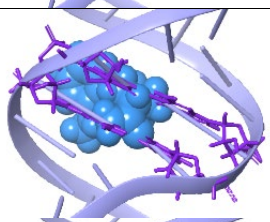
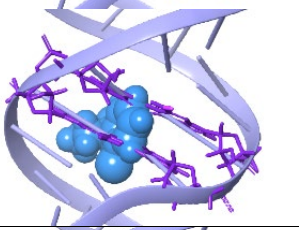
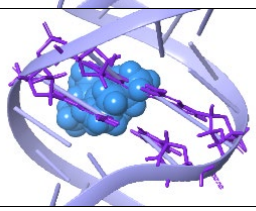
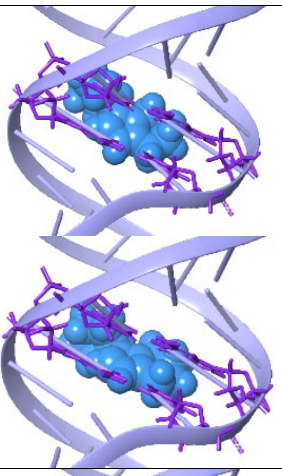
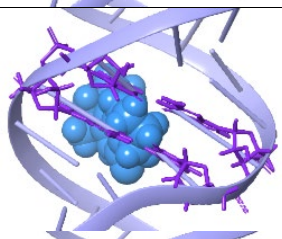
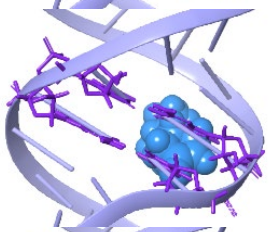
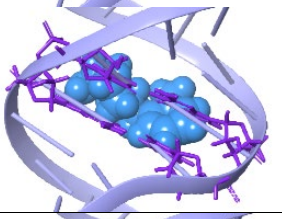
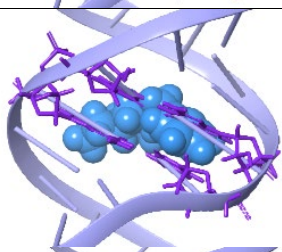
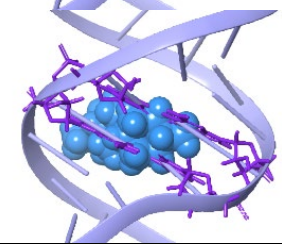
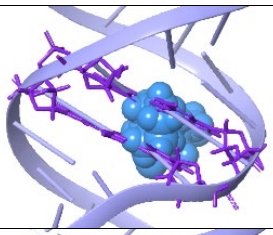
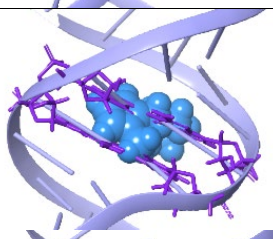
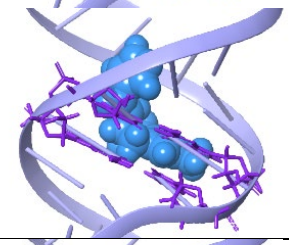
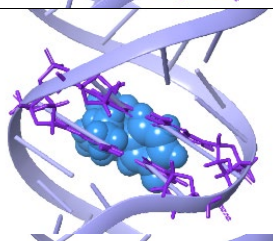
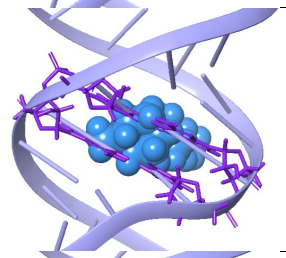
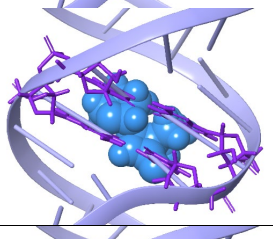
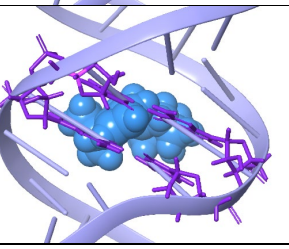
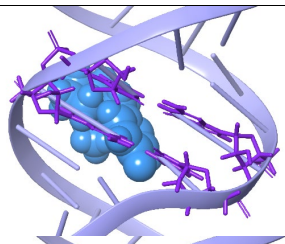
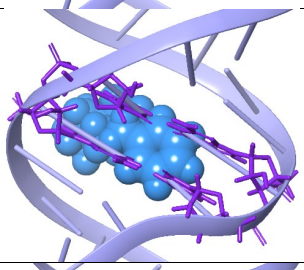
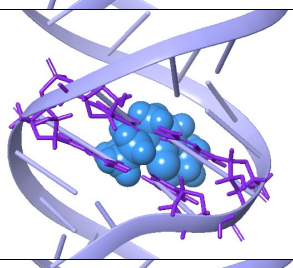
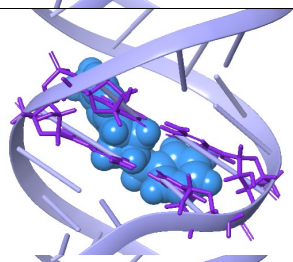
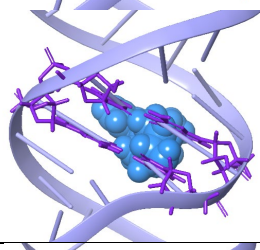
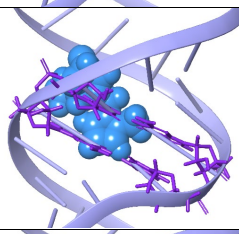
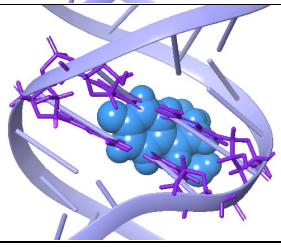
Type of Binder	Example Image
Figure shows clear fragment interaction in the intercalation site.	
Figure shows portions of the fragment making contacts with the intercalation site	
Figure shows portions of the fragment making contacts at the intercalation site and above the site	

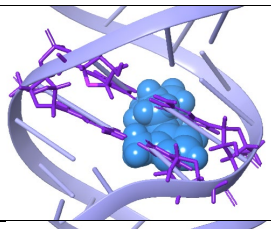
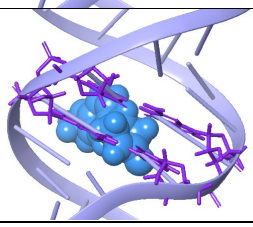
Table 12. 28 Fragment hits with docking scores

Fragment Number	Docked Image	Docking Score
SY8441		-8.234
SY7777		-8.043 -8.321
SY9005		-7.989
SY9082		-7.940
SY8338		-7.931
SY7685		-7.930

SY8512		-7.874
SY8665		-7.744
SY9277		-7.693
SY8719		-7.691
SY7651		-7.612
SY7675		-7.538

SY8350		-7.421
SY7719		-7.379
SY8087		-7.369
SY8293		-7.277
SY8112		-7.266
SY8889		-7.265
SY7639		-7.247

SY8492		-7.239
SY9267		-7.200
SY8264		-7.119
SY7609		-7.111
SY7959		-7.102
SY9173		-7.09
SY8546		-7.079

SY8748		-7.062
SY8688		-7.051

Taking a deeper look at the structures of the 28 fragments, some recurring features were observed. The most common shared functional groups were di-pyridine, pyrazole, and imidazole.

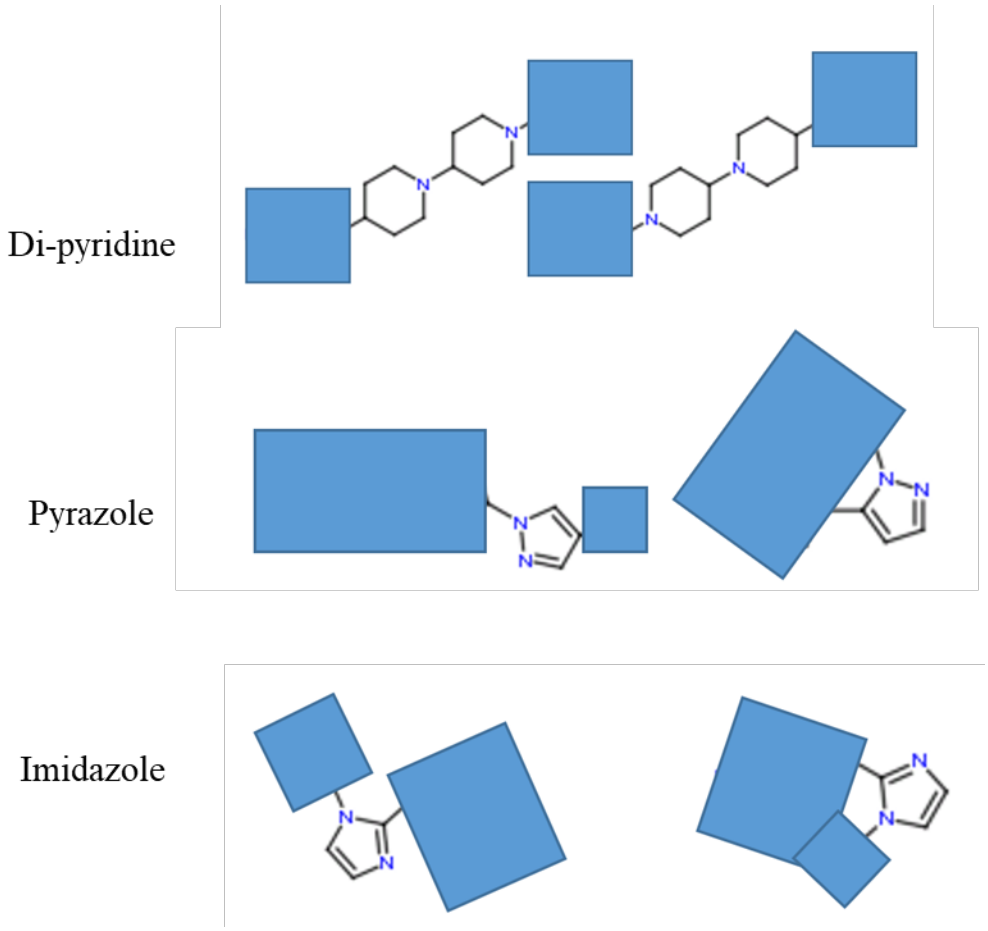


Figure 27. Blinded fragments show structural diversity.

Chapter IV.

Discussion

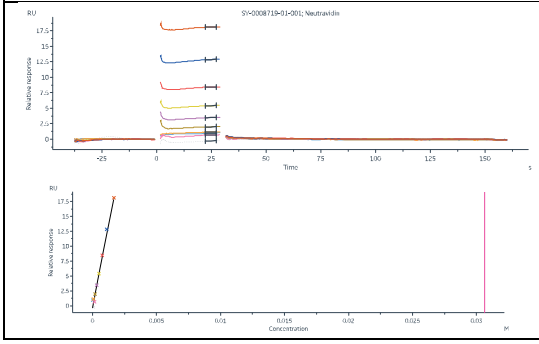
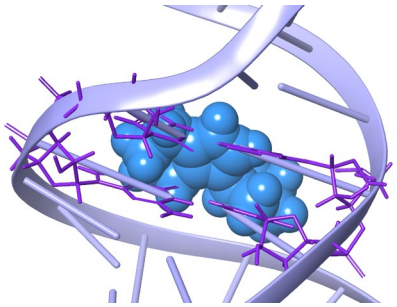
This section details how the screening data across both methods was analyzed. This data then was used to determine if the hypothesis could be supported using this data package.

Data Discussion

This study sought to understand the hypothesis of if in-silico virtual screening can identify novel RNA binding chemical matter with similar results as an in-vitro biophysical screen.

The SPR biophysical screen resulted in the determination that 31 unique fragments directly bound the 30mer RNA hairpin structure. The in-silico virtual screen resulted in the determination that 28 unique fragments made energetically favorable contacts to the CUG30 RNA hairpin structure. Comparing the fragment hit lists between both methods, there were two fragment hits from both methods. This means that the hit percentage overlap rate is 3.2%.

Table 13. Screening Profile of Shared Hit #1 (SY8719)

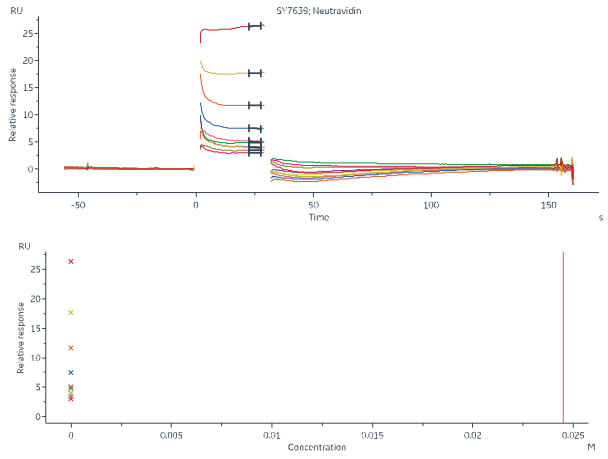
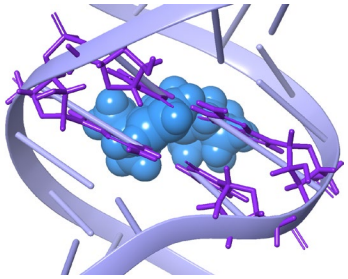
SPR Dose Response Curve	Docking Image	Glide Score
		-7.691

The SPR data for Shared Hit #1 (SH1) shows direct target binding through concentration-dependent binding response. SH1 also exhibits fast kinetics as evidenced by the association and dissociation phases on the sensorgram. However, the data does not indicate surface saturation being reached, therefore the K_d cannot be fit with this data. The fragment exhibited poor solubility, so the experiment was limited in increasing the concentration of compound to reach surface saturation.

The virtual screen data shows the fragment making contacts with the intended intercalation site and the glide score of -7.691 indicates that the docking pose is energetically favorable.

The structure of SH1 is unique among the fragment set and no hits from the virtual screen or SPR screen are of a similar scaffold.

Table 14. Screening Profile of Shared Hit #1 (SY7639)

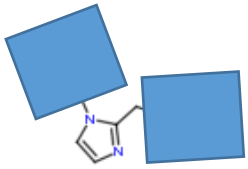
SPR Dose Response Curve	Docking Image	Glide Score
		-7.247

The SPR data for Shared Hit #2 (SH2) shows direct target binding through concentration-dependent binding response. SH2 also exhibits fast kinetics as evidenced by the association and dissociation phases on the sensorgram. However, the data does not indicate surface saturation being reached, therefore the K_d cannot be fit with this data. The fragment exhibited poor solubility, so the experiment was limited in increasing the concentration of compound to reach surface saturation.

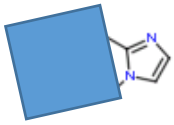
The virtual screen data shows the fragment making contacts with the intended intercalation site and the glide score of -7.247 indicates that the docking pose is energetically favorable. SH2 also has a portion of the molecule which extends below the intended intercalation site.

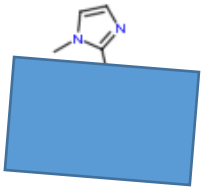
The structure of SH2 is contains both an imidazole and a seven membered aromatic ring. The imidazole is a structural feature that is well represented in the virtual screening hit list.

Hit compound:

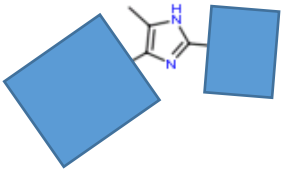


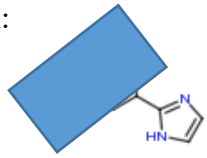
Imidazole Fragments from virtual screen:





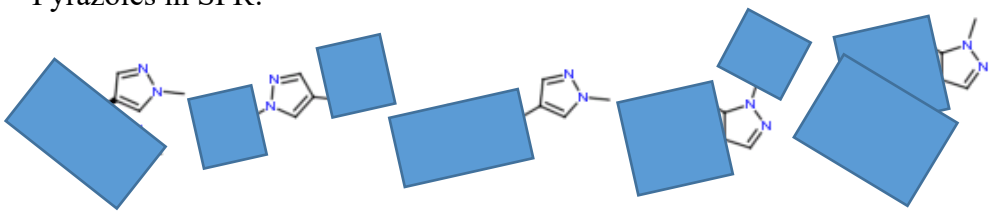
Imidazole Fragments from SPR screen:



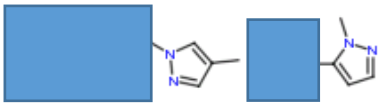


Between the SPR and virtual screen data, both methods hit lists included a good representation of pyrazole containing compounds. However, there were no pyrazole containing compounds which overlapped as hits in both methods.

Pyrazoles in SPR:



Pyrazoles in virtual screen:



Method Comparison

This study proposed to answer the overarching hypothesis of if an in-silico virtual screen can identify novel RNA binding chemical matter with similar results as an in-vitro biophysical screen. The resulting study successfully executed an SPR screen and matching computational screen to test this hypothesis. The study data shows a 3.2% hit overlap between both methods. Hence, the hypothesis is supported as the data shows that an in-silico virtual screen can identify novel CUG30 RNA hairpin binding matter with statistically significant similar results as a biophysical screen.

While the hypothesis is supported by the data, some key questions remain as to why there is not greater overlap between methods. The first possible reason for the < 5% overlap is that this SPR experiment required the surface immobilization of the CUG30 hairpin to the chip surface. During the method development there was no reason to question the surface conformation of the oligonucleotide. However, it does represent a potential difference from the hairpin model used in the virtual screening. The second reason for the < 5% overlap is that SPR can include several different types of fragment binding including complex binding. Furthermore, each fragment in the SPR screen was limited in solubility. This presented a challenge in being able to fit a K_d by kinetics or affinity for the chemical matter.

Finally, SPR sensorgrams capture binding of a fragment to any portion of the immobilized CUG30 oligomer as opposed to grid-restricted docking as was done in the virtual screen. Limitations in virtual screening include target flexibility, a difference in solvent models, sequence specific docking bias based on docking grid determination, and lack of ability to account for unique binding features such as complex binding.

Though the data presented in this study shows a modest hit overlap, the data package still supports that both methods can be used complementary to one another in early drug discovery campaigns. Both biophysical assays and computational screening give unique insight into what early chemical matter on a screening campaign will entail. Computational screening tools allow for significant numbers of fragments and compounds to be screened with relative ease and relative low cost. SPR is a high throughput method that gives direct target binding data, which is critical for assessing target engagement. Furthermore, there are increasing numbers of publications on the value of using both computational methods alongside biophysical characterization for early hit determination to reduce false positives, artifactual binding, and to reduce the time and effort required to develop validated chemical matter across novel targets and different disease types. This study adds to both the fields of fragment-based drug discovery and Myotonic Dystrophy. The data presented supports the existing data that when both computational and biophysical tools are used in parallel or sequentially, they have the capability to shape robust drug discovery efforts. The data also presents novel chemical matter than can be utilized in a robust drug discovery program to develop drugs for the treatment of Myotonic Dystrophy.

Future Work

Finding fast, effective, and optimal methods in early drug discovery continues to be a challenge. The data presented in this study illustrate that both biophysical and computational tools can be used concurrently to aid in narrowing down a large library of chemical matter to selected chemical matter that are validated in both methods. The utilization of both methods concurrently should continue in both academic and industry groups alike as drug discovery projects in Myotonic Dystrophy and other disease areas are undertaken. The work presented here should move forward into initial medicinal chemical efforts to elaborate upon the fragment structures identified. From there stoichiometry and other mechanism-understanding studies should be enabled. Finally, the assessment of compounds in relevant biological and pre-clinical studies should be enabled.

While this study was being designed and undertaken, other academic groups published work on Myotonic Dystrophy drug discovery as well as review articles detailing the value of utilizing both computational and biophysical screening in fragment-based drug discovery. The Disney group published a study of an RNA-repeat focused DNA-encoded library to inform the design of a degrader of CUG repeat expansions. In March 2023, Pascual-Gilbert et al published a review of the Myotonic Dystrophy type 1 drug development pipeline. In this review they share that there are now 20 drug candidates across different drug classes including small molecules, nucleic acids, fatty acid conjugation, monoclonal antibody conjugation, peptide backbone conjugation, adeno-associated virus, and genome/transcriptome engineering. The small molecules in clinical trials represent a collection of repurposed drugs like Tideglusib originally

targeted for Alzheimer's, Mexiletine originally developed for anti-arrhythmia, Pitolisant originally developed as a stimulant drug for narcolepsy, Metformin for anti-diabetes, and the antibiotic Erythromycin. While progress has been made in the field, it still leaves a gap in targeted RNA therapies to treat Myotonic Dystrophy.

References

Chial, H. (2008) Rare Genetic Disorders: Learning About Genetic Disease Through Gene Mapping, SNPs, and Microarray Data. *Nature Education* 1(1):192

Coonrod, L. A., Nakamori, M., Wang, W., Carrell, S., Hilton, C. L., Bodner, M. J., Siboni, R. B., Docter, A. G., Haley, M. M., Thornton, C. A., & Berglund, J. A. (2013). Reducing levels of toxic RNA with small molecules. *ACS Chemical Biology*, *8*(11), 2528–2537. https://doi.org/10.1021/cb400431f

Costales, M. G., Childs-Disney, J. L., Haniff, H. S., & Disney, M. D. (2020). How We Think about Targeting RNA with Small Molecules. *Journal of Medicinal Chemistry*, *63*(17), 8880–8900. https://doi.org/10.1021/acs.jmedchem.9b01927

Coyne, A. G., Scott, D. E., & Abell, C. (2010). Drugging challenging targets using fragment-based approaches. *Current Opinion in Chemical Biology*, *14*(3), 299–307. https://doi.org/10.1016/j.cbpa.2010.02.010

Duan, D., Goemans, N., Takeda, S. *et al.* Duchenne muscular dystrophy. *Nat Rev Dis Primers* 7, 13 (2021). https://doi.org/10.1038/s41572-021-00248-3

Duchenne and Becker muscular dystrophy (DMD/BMD). Kennedy Krieger Institute. (n.d.). Retrieved March 28, 2022, from https://www.kennedykrieger.org/patient-care/conditions/muscular-dystrophy/duchenne-and-becker

Erlanson, D. A. (2011). Introduction to fragment-based drug discovery. *Topics in Current Chemistry*, 1–32. https://doi.org/10.1007/128_2011_180

Find a neuromuscular disease. Muscular Dystrophy Association. (2020, October 5). Retrieved March 28, 2022, from https://www.mda.org/disease/list

Genetic disorders: What are they, types, symptoms & causes. Cleveland Clinic. (2021, August 20). Retrieved March 28, 2022, from https://my.clevelandclinic.org/health/diseases/21751-genetic-disorders

Giannetti, A. M. (2011). From experimental design to validated hits. *Methods in Enzymology*, 169–218. https://doi.org/10.1016/b978-0-12-381274-2.00008-x

Kim, E. Y., Barefield, D. Y., Vo, A. H., Gacita, A. M., Schuster, E. J., Wyatt, E. J., Davis, J. L., Dong, B., Sun, C., Page, P., Dellefave-Castillo, L., Demonbreun, A., Zhang, H. F., & McNally, E. M. (2019). Distinct pathological signatures in human cellular models of myotonic dystrophy subtypes. *JCI Insight*. https://doi.org/10.1172/jci.insight.122686

Kirsch, P., Hartman, A. M., Hirsch, A. K. H., & Empting, M. (2019, November 26). *Concepts and core principles of fragment-based drug design*. Molecules (Basel, Switzerland). Retrieved October 24, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6930586/

Kumar, A., Park, H. J., Fang, P., Parkesh, R., Guo, M., Nettles, K. W., & Disney, M. D. (2011). Myotonic dystrophy type 1 rna crystal structures Reveal heterogeneous 1 × 1 Nucleotide uu internal Loop conformations. *Biochemistry*, *50*(45), 9928–9935. https://doi.org/10.1021/bi2013068

Li, Q. (2020). Application of fragment-based drug discovery to versatile targets. *Frontiers in Molecular Biosciences*, *7*. https://doi.org/10.3389/fmolb.2020.00180

Luo, J., Wei, W., Waldispühl, J., & Moitessier, N. (2019). Challenges and current status of computational methods for docking small molecules to nucleic acids. *European Journal of Medicinal Chemistry*, *168*, 414–425. https://doi.org/10.1016/j.ejmech.2019.02.046

Matsuno, K., Masuda, Y., Uehara, Y., Sato, H., Muroya, A., Takahashi, O., Yokotagawa, T., Furuya, T., Okawara, T., Otsuka, M., Ogo, N., Ashizawa, T., Oshita, C., Tai, S., Ishii, H., Akiyama, Y., & Asai, A. (2010). Identification of a New Series of STAT3 Inhibitors by Virtual Screening. *ACS Medicinal Chemistry Letters*, *1*(8), 371–375. https://doi.org/10.1021/ml1000273

Mazzoli, M., Ariatti, A., Garuti, G. C., Agnoletto, V., Genovese, M., Gozzi, M., Kaleci, S., Marchioni, A., Malagoli, M., & Galassi, G. (2020). Predictors of prognosis in type 1 myotonic dystrophy (DM1): longitudinal 18-years experience from a single center. *Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology*, *39*(3), 109–120. https://doi.org/10.36185/2532-1900-015

Mooers, B. H., Logue, J. S., & Berglund, J. A. (2005). The structural basis of myotonic dystrophy from the crystal structure of Cug Repeats. *Proceedings of the National Academy of Sciences*, *102*(46), 16626–16631. https://doi.org/10.1073/pnas.0505873102

Murata, A., Nakamori, M., & Nakatani, K. (2019). Modulating RNA secondary and tertiary structures by mismatch binding ligands. *Methods*, *167*, 78–91. https://doi.org/10.1016/j.ymeth.2019.05.006

Nguyen, L., Luu, L. M., Peng, S., Serrano, J. F., Chan, H. Y., & Zimmerman, S. C. (2015). Rationally designed small molecules that target both the dna and rna causing myotonic dystrophy type 1. *Journal of the American Chemical Society*, *137*(44), 14180–14189. https://doi.org/10.1021/jacs.5b09266

Ozimski, L. L., Sabater-Arcis, M., Bargiela, A., & Artero, R. (2020). The hallmarks of myotonic dystrophy type 1 muscle dysfunction. *Biological Reviews*, *96*(2), 716–730. https://doi.org/10.1111/brv.12674

Pascual-Gilabert, M., López-Castel, A., & Artero, R. (2021). Myotonic dystrophy type 1 drug development: A pipeline toward the market. *Drug Discovery Today*, *26*(7), 1765–1772. https://doi.org/10.1016/j.drudis.2021.03.024

Pascual-Gilabert, M., Artero, R., & López-Castel, A. (2023). The myotonic dystrophy type 1 drug development pipeline: 2022 edition. *Drug Discovery Today*, *28*(3), 103489. https://doi.org/10.1016/j.drudis.2023.103489

Ranum, L. P. W., & Day, J. W. (2004). Myotonic dystrophy: RNA Pathogenesis comes into focus. *The American Journal of Human Genetics*, *74*(5), 793–804. https://doi.org/10.1086/383590

Shepherd, C. A., Hopkins, A. L., & Navratilova, I. (2014). Fragment screening by SPR and advanced application to GPCRs. *Progress in Biophysics and Molecular Biology*, *116*(2-3), 113–123. https://doi.org/10.1016/j.pbiomolbio.2014.09.008

Silvestre, H. L., Blundell, T. L., Abell, C., & Ciulli, A. (2013). Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery. *Proceedings of the National Academy of Sciences*, *110*(32), 12984–12989. https://doi.org/10.1073/pnas.1304045110

Types of muscular dystrophy. Patient Care at NYU Langone Health. (n.d.). Retrieved March 28, 2022, from https://nyulangone.org/conditions/muscular-dystrophy/types

Winter, A., Higueruelo, A. P., Marsh, M., Sigurdardottir, A., Pitt, W. R., & Blundell, T. L. (2012). Biophysical and computational fragment-based approaches to targeting protein–protein interactions: applications in structure-guided drug discovery. *Quarterly Reviews of Biophysics*, *45*(4), 383–426. https://doi.org/10.1017/s0033583512000108

Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., & Swanson, M. S. (2007). Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Research*, *35*(16), 5474–5486. https://doi.org/10.1093/nar/gkm601