Chemically Inducible Dimerization and Chemically Disruptable Systems for Spatial and Temporal Control of Cellular Processes

Kyler J Radmall

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2020

Reading Committee:

Dustin J. Maly, Chair

Champak Chatterjee

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Chemically Inducible Dimerization and Chemically Disruptable Systems for Spatial and Temporal Control in Cellular Processes

Kyler J Radmall

Chair of the Supervisory Committee:
Professor Dustin J. Maly
Chemistry

Spatial proximity–defined as the physical distance between two biomolecules–plays an important role in controlling and regulating many cellular functions. Engineered cellular systems have been developed where the physical distance between two interacting pairs can be controlled with exogenous inputs. By constraining proteins of interest to a restricted space, the relative concentration of one protein as seen by another is increased, which can lead to increased reaction rates or binding events. Recently, our group developed a new CID system called Pleiotropic Response Outputs from a Chemically Inducible Single Receiver (PROCISiR), which allows the rapid co-localization of two proteins with clinically-approved drugs. While PROCISiR is capable of some level of reversibility, the rates observed are not useful for some applications. The first chapter of this thesis describes efforts to achieve a more rapidly reversible PROCISiR system by incorporating drug-resistant mutants of the NS3a protease. The second chapter of this thesis

describes efforts to increase the affinity of a previously-developed chemically-disrupted proximity system. Together, these efforts help expand the chemical genetic toolbox for rapidly controlling cellular proximity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

pM	**Picomolar**
nM	**Nanomolar**
µM	**Micromolar**
µL	**Microliter**
mL	**Milliliter**
µg	**microgram**
FKBP	**FK506 binding protein**
FRB	**FKBP rapamycin binding domain**
mRNA	**Messenger ribonucleic acid**
DNA	**Deoxyribonucleic acid**
EGFP	**Enhanced green fluorescent protein**
DMSO	**Dimethyl sulfoxide**
RT-qPCR	**Reverse transcriptase quantitative polymerase chain reaction**
BOC	**tert-butyloxycarbonyl**
TFA	**Trifluoroacetic acid**
TCO	**transcyclooctene**
HCV	**Hepatitis C virus**
VPR	**VP64 P65 Rita**
GTP	**Guanosine triphosphate**
GDP	**Guanosine diphosphate**
FACS	**Fluorescence activated cell sorting**

ACKNOWLEDGEMENTS

I first want to express my thanks and gratitude to my spouse Rachel. She has been a constant support in my decisions and encourages me to be the best version of myself. Without here I would have never discovered my love for chemistry, science and learning. She truly is my rock and is with me through thick and thin.

I also want to thank members of the Maly lab, they have been influential to my scientific journey. I want to thank Cindy and Emily for all the help and guidance they have given me with my projects and for answering my constant questions guiding me to think more critically about science. Also, thanks to Ethan and Jessica, who were part of my cohort, and have been a great support to me. Everyone else in the lab, Linglan, Sujata, Zack, and Gayani has influenced me for the better I say thank you.

I also want to thank my adviser Dusty for his patience and guidance. He has helped me grow in my scientific abilities and has been genuine in his advice and suggestions. I also want to thank Champak who is a great scientist who I look up to. I'm grateful for my time at the University of Washington and being here has helped me realized the direction I want to take with my career.

Chapter 1: Reversible chemically inducible dimerization

Introduction

Chemically inducible dimerization (CID) systems have played an important role in understanding and controlling many cellular processes such as signal transduction cascades[1,2,3]. One of the most widely used small molecules in CID systems is rapamycin, which mediates the dimerization between FKBP and FRB[4]. New analogs of rapamycin, called rapalogs, have been developed but these CID systems still rely on the use of endogenous proteins and have off target effects[5]. Our lab has developed a novel biorthogonal CID called Pleiotropic Response Outputs from a Chemically Inducible Single Receiver (PROCISiR) but it lacks the ability for rapid reversibility like some rapalog-based systems[6]. This led us to evaluate what we know about our CID system and attempt to make this system more rapidly reversible. To do this, we first considered how the PROCISiR system functions.

Dr. Foight in the Maly lab at the University of Washington has developed a novel CID system called PROCISiR[7]. This project started with the intent to design a CID system with no off-target effects that can also be used as a platform for multiple functional outputs. The central receiver for this CID system is a catalytically inactive (S139A) mutant of the Hepatitis C viral protease NS3/4a (NS3a). One of the main attractions of using NS3a is that there are many clinically approved protease inhibitors that can be selected for use as chemical inducers. Additionally, NS3a protease inhibitors have no mammalian targets due to the structural difference between viral and mammalian proteases. Where this system excels, compared with other CID systems, the ability for multiple and tunable outputs with a single central receiver leading to graded control; which allows for a proportional response depending on the ratio of small molecules added. Dr. Foight used Rosetta to develop two chemical readers (Danoprevir

NS3a Chemical Reader (DNCR2) and Grazoprevir NS3a Chemical Reader (GNCR1)) that each recognize a separate NS3a:drug interface, NS3a:Danoprevir and NS3a:Grazoprevir (**Figure 1**). DNRC2 and GNCR1 have high apparent affinities for their NS3a:drug bound complexes, 36 pM and 140 nM, respectively[7]. Further experimentation demonstrated that DNCR2 and GNCR1 do not bind to NS3a in the absence of drug. While PROCISiR works well and confers new control properties, like graded control, it lacks the ability for rapid dissociation. Adding a rapidly reversible behavior to this system opens up opportunities for additional control of cellular processes.

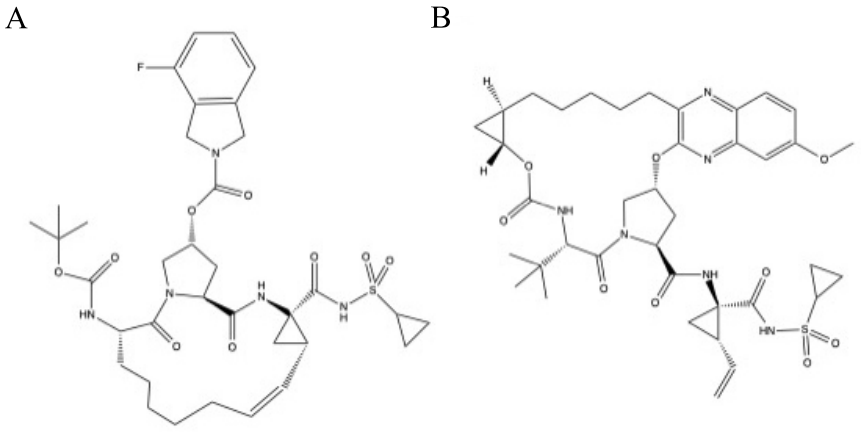

Figure 1**.** ***Chemical structure of NS3a protease inhibitors***. A) Structure of Danoprevir. B) Structure of Grazoprevir

Results

An experiment carried out by Dr. Foight that provides a foundation for this project is a transcriptional activation assay using PROCISiR. The goal of this experiment was to measure the dissociation of the NS3a:Danoprevir:DNCR2 ternary complex through competition with Grazoprevir in mammalian cells. In this setup, NS3a was fused to a catalytically dead endonuclease called dCas9. dCas9 is targeted to specific locations in the genome with a bound single stranded guide RNA (sgRNA). The NS3a-dCas9 construct was co-transfected with a

sgRNA targeting the promoter region of CXCR4, an endogenous mammalian gene, and a DNCR2-VPR protein fusion (**Figure 2A**).

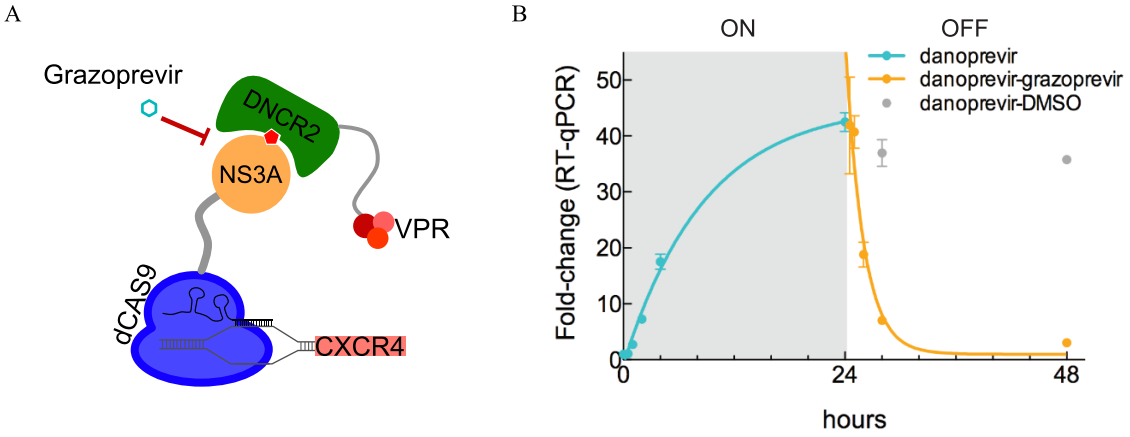

Figure 2. ***mRNA transcriptional deactivation with NS3a and DNCR2.*** A) Conceptual schematic of the mRNA reversibility experiment. The red pentagon represents danoprevir. B) RT-qPCR quantification of CXCR4 mRNA over 48 hrs. At the 24 hr point grazoprevir was added to the cells.

In the presence of Danoprevir, DNCR2 binds to NS3a, localizing VPR, a transcriptional activator, to the CXCR4 promoter region. This led to an increase in CXCR4 mRNA levels, allowing for quantitative analysis of mRNA with RT-qPCR. After 24 hours with 10 µM of Danoprevir, the media was removed and replaced with DMSO or 10 µM Grazoprevir. The data (**Figure 2B**)[7] demonstrate that the dissociation of the DNCR2/NS3a/Danoprevir complex is very slow, with no diminution of CXCR mRNA levels 24 hours after washout (+DMSO). However, an increased dissociation rate ($t_{½}$ of about 2 hours) was observed when 10 µM Grazoprevir was added. It is unclear whether this rate corresponds to mRNA degradation or dissociation of the ternary complex. While these preliminary experiments confirm that drug swapping experiments can be used to make PROCISiR reversible, we speculate that the time scales observed are too slow for rapidly turning on or off signaling pathways, which occur on a scale of minutes[8]. Additionally, issues arise when using this system for multiple cellular outputs. For example, when one partner is not completely deactivated, readout could come from competing signals.

Therefore, I set out to answer how can we increase the rate of dissociation for the NS3a:dano:DNCR2 ternary complex, and measure these changes on a protein level using fluorescence techniques.

The rate of dissociation, k_{off}, can be related to K_d and the rate of association, k_{on} (**equation 1**).

$$K_d = \frac{k_{off}}{k_{on}} \quad (\mathbf{1})$$

For many protein-protein interactions k_{off} is the major determinant of K_d[17]. Therefore, we predicted that a more rapidly reversible system could be generated by increasing K_d. How can the K_d of the NS3a:Dano:DNCR2 ternary complex be increased? The first possible solution tested was using drug resistant mutants of NS3a. NS3a has well known mutations leading to drug resistance and many inhibitors are somewhat resistant to these mutations, meaning they still have a low K_d [9]. Analysis from crystal structures of NS3a drug bound complexes revealed potential residues that hinder drug binding and led us to test three mutations: R155K, A156T and D168A. The first approach to understanding what effect these mutations might have was a qualitative experiment. The set up for the experiment is shown in **Figure 3A**.

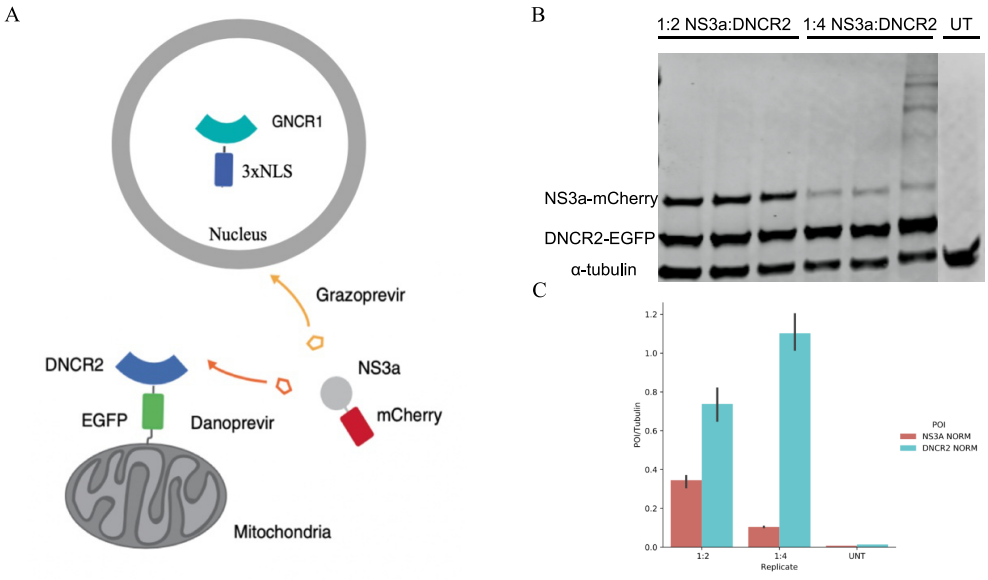

Figure 3. ***Design and analysis of protein level reversibility assay.*** A) Diagram indicating the addition of Danoprevir leads to mCherry and EGFP overlap with NS3a:dano:DNCR2 complex formation. B) Western Blot of NS3a-mCherry and TOM20-DNCR2-EGFP-P2A-3x-NLS-GNCR1 plasmids transfected in HEK293T cells with a 1:2 and 1:4 DNA ratio. UT = untransfected. C) Quantification of expression levels relative to α-tubulin**.**

Two constructs were co-transfected into Hek293T cells: mCherry-NS3a and TOM20-DNCR2-EGFP-P2A-3xNLS-GNCR in a 1:3 ratio. Assays were performed in a 12-well plate seeded at 200,000 cell/mL. A key factor for this assay was to ensure that DNCR2 was in excess with respect to NS3a. If NS3a was more abundant, then we could not assume that the majority of NS3a signal came from ternary complex dissociation (**Figure 3B**). After 24 hours, 500 nM of danoprevir was added to the cells, recruiting NS3a-mCherry to mitochondrial bound DNCR2-EGFP.

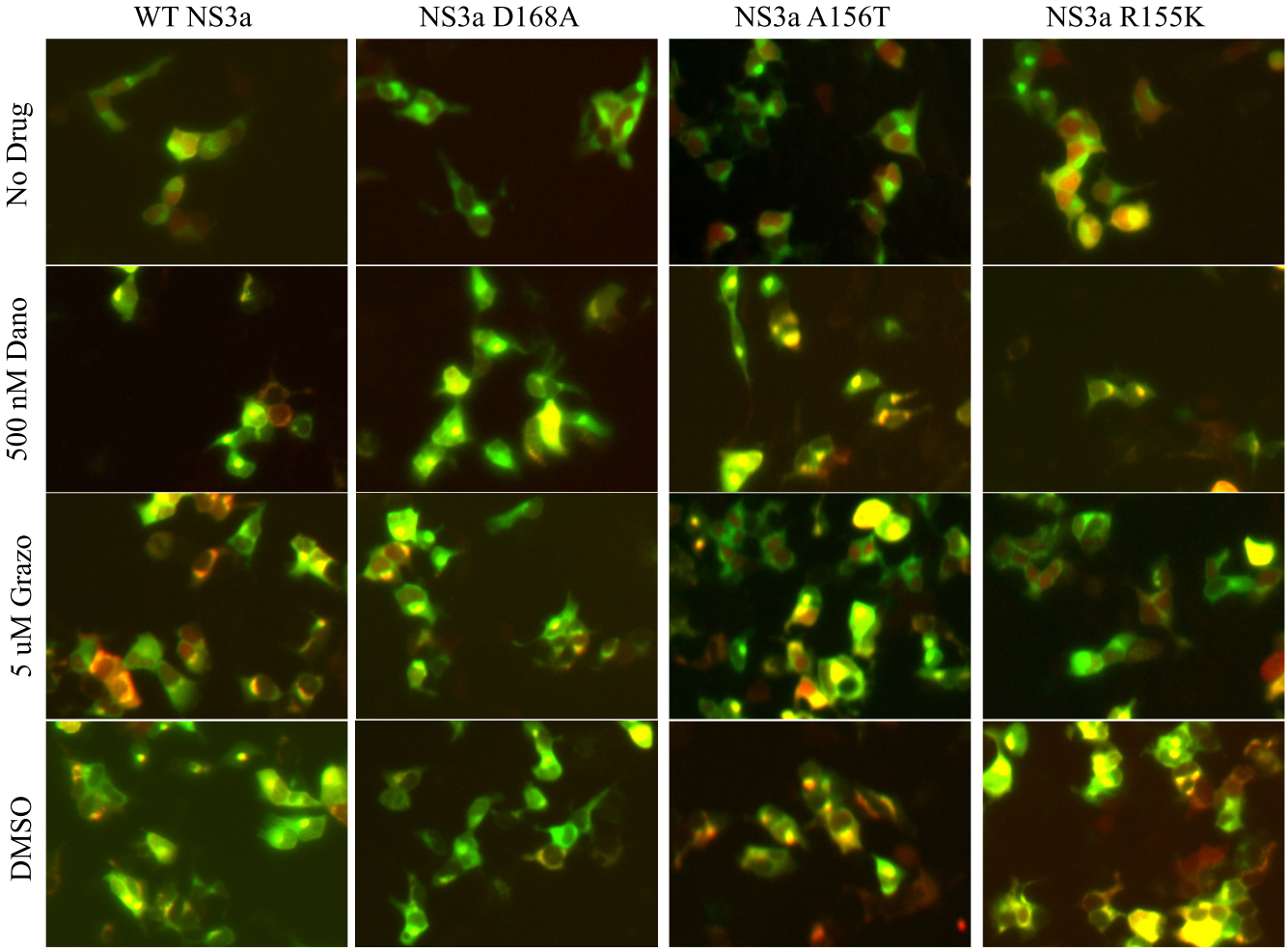

Figure 4. ***Zoe Images of Grazoprevir chase.*** Experimental results for the grazoprevir chase experiment with HEK293T. No drug represents cells before treatment. 500 nM Danoprevir are

images of cells treated with Danoprevir for 1 hr. 5 μM Grazo are images of cells treated with Grazoprevir for 2 hrs. DMSO are images of cells treated with DMSO for 2 hrs.

After the addition of danoprevir, cells were incubated for 1 hour. The media was then removed and replaced with fresh media containing 5 μM of grazoprevir or DMSO for two hours. Adding grazoprevir to the cells competed with danoprevir for binding to NS3a, which when bound to grazoprevir is recruited to nuclear bound GNCR1. Images were taken using a ZOE fluorescent cell imager at the pre-drug phase, danoprevir incubation period, and at two hours after the grazoprevir media swap (**Figure 4**). Images show that D168A behaved similar to WT NS3a while R155K and A156T mutants had more mCherry-NS3a fluorescence located in the nucleus. It is difficult to determine how well these two mutants performed compared to WT NS3a without a more quantitative experiment. This led us to use fluorescence microscopy to test whether or not these NS3a mutations had a significant effect on reversibility. This experiment was identical to the first qualitative one except after two hours the cells were fixed and imaged with a SP8X Leica confocal microscope at the Keck center (**Figure 5A**). Using EzColocalization software[10] Pearson's r coefficients were determined to quantify fluorescent overlap between the different NS3a variants (**Figure 5B**). This experiment quantifies the amount of protein that moves from the mitochondria to the nucleus in a given time allotment. The amount of protein moved is associated with the fluorescent signal and its overlap. R155K and A156T mutants had less correlation with the mitochondrial DNCR2-EGFP signal than WT NS3a-mCherry indicating more NS3a dissociated and migrated to the nucleus in two hours. It is clear that mutants A156T and R155K show promise for reversibility but futher analysis will be needed to determine higher-resolution kinetics. In a follow up experiment, a double mutant, A156T R155K, was cloned and tested but showed no dimerization of NS3a and the associated readers GNCR1 or DNCR2.

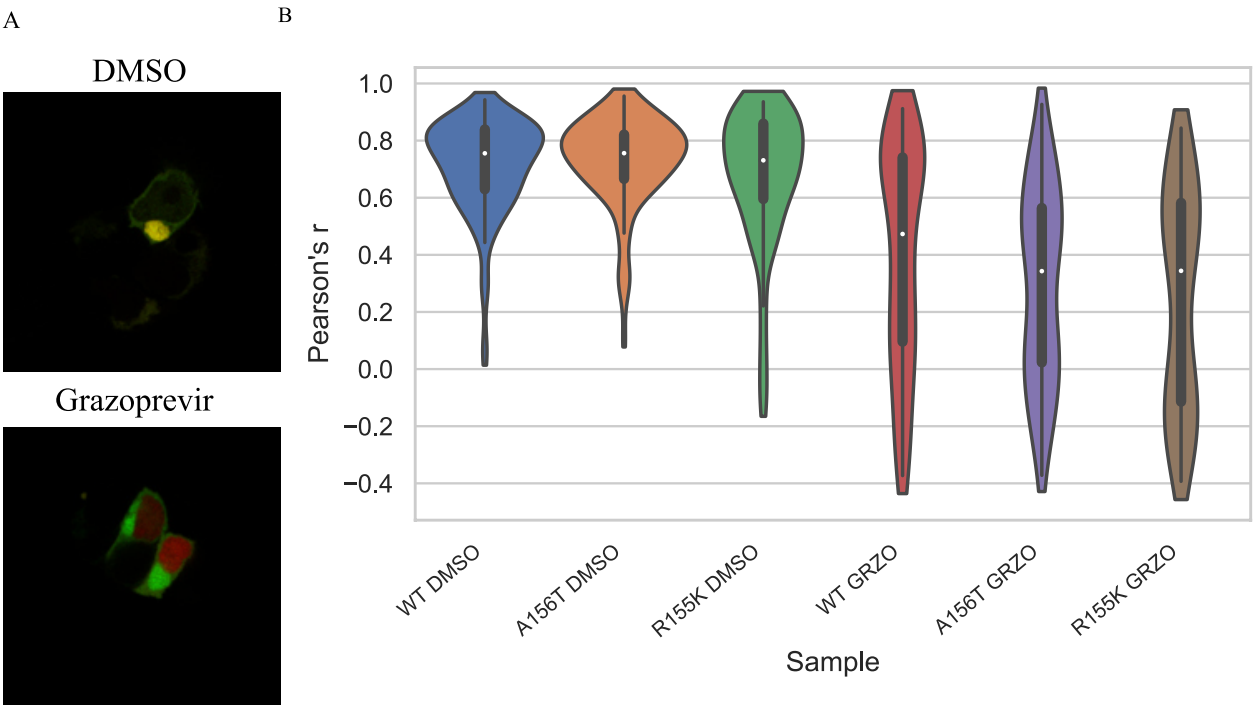

Figure 5. ***Quantification of Grazoprevir chase experiment.*** A) Example images taken from Leica SP8X confocal microscope from DMSO and Grazoprevir conditions from transfected HEK293T cells. B) Violin plot of Pearson's r coefficients for each experimental condition calculated using EzColocalization software.

A second solution that aims to increase the dissociation of the ternary complex is modifying the chemical structure of danoprevir. An ideal danoprevir variant would allow rapid chemically inducible dimerization, but also give temporal control over dissociation through measures such as chemical reactions or light sensitive conformational changes. When looking at the crystal structure of NS3a:dano:DNCR2 ternary complex, there is a solvent exposed amino group that could readily be modified after removal of the BOC group. We set out to use click chemistry between trans-cyclooctenne and tetrazine as a chemical reaction method to induce dissociation. For this to work, danoprevir variants still need to induce dimerization. Preliminary data shows that removing the BOC group with TFA, leading to free amino-danoprevir, does not induce dimerization as efficiently with NS3a and DNCR2. Linglan Fang, in the Maly lab, synthesized two other Danoprevir variants: 4-trans-cyclooctene-danoprevir (4TCO-dano) and 2-

trans-cyclooctene-danoprevir(2TCO-dano). It was shown that 4TCO-dano does indeed induce dimerization but 2TCO-dano does not (**Figure 6**).

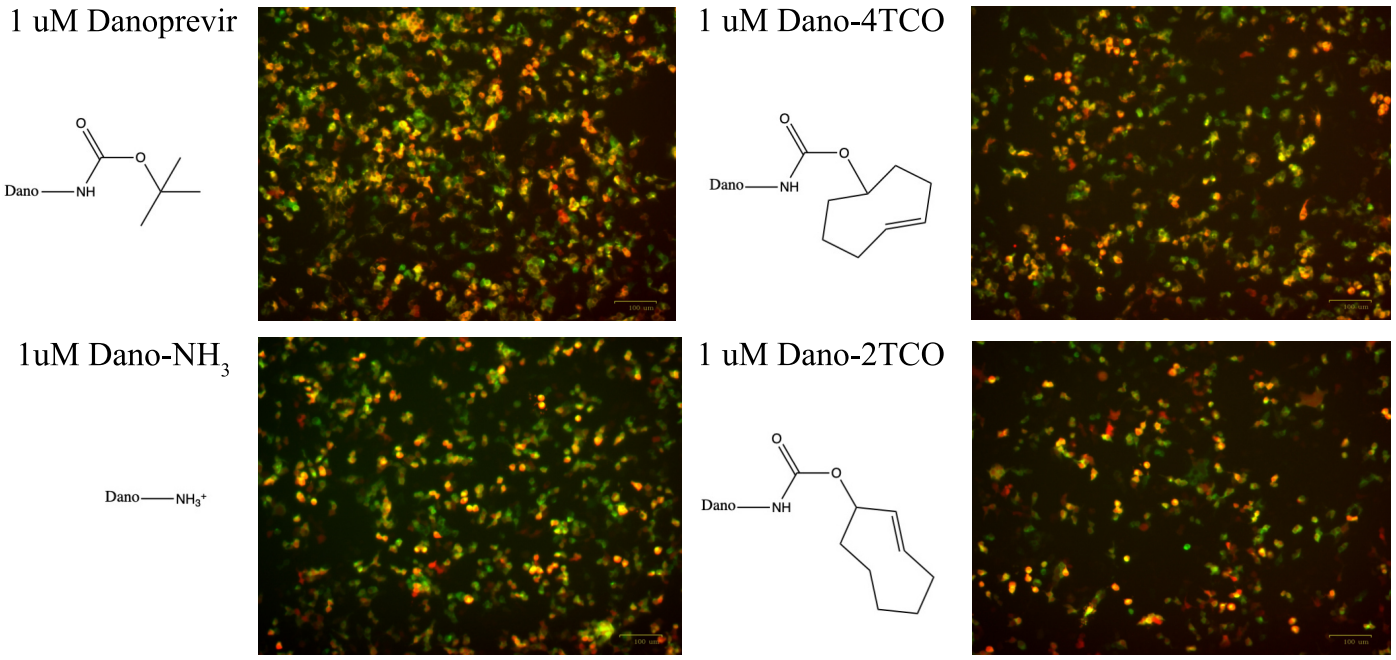

Figure 6. ***Zoe results from testing danoprevir derivatives***. Images of HEK293T cells transfected with NS3a-mCherry, Tom20-DNCR2-EFGP, and 3xNLS-GNCR1 after 1 hr of incubation with Danoprevir derivatives.

By using trans-cyclooctene, a bulky substituent can be added upon addition of a tetrazine derivative. Dr. Foight performed one experiment with 4TCO-dano and tetrazine but did not observe any dissociation. It is possible that the substituent did not have enough steric hinderance for disruption and needs additional experimentation. Further exploration of tetrazine derivatives could be beneficial to the project.

Discussion

Although much insight has been gained from experimental data, further experiments and optimization are needed in order to achieve reversibility on a smaller time scale. NS3a mutations have proven helpful to a certain extent but kinetic experiments are needed to quantify the k_{off} for each mutant. After mutant experiments, it is probable that chemical modification may be the route to pursue. It seems likely that adding a larger substituent to tetrazine could result in

dissociation when it reacts with 4TCO-dano. Another viable avenue is conjugating a photoswitch that changes conformation when irradiated with light such as azobenzenes. If these experiments prove unsuccessful, it would be useful to explore photocleavable groups that lead to the formation of free amine danoprevir promoting dissociation. Furthermore, Grazoprevir modification, though more difficult, might increase the diversity of options with PROCISiR. Developing a rapidly reversible system can have a wide variety of applications and will be beneficial when using PROCISiR as a tool set to control proximity-based cellular functions. A hopeful future application for the lab would be demonstrating reversibility with AcrIIA4 a potent Cas9 inhibitor. Using an AcrIIA4-NS3a fusion, AcrIIA4 can be localized and excluded from the nucleus by adding danoprevir. This could form a ternary complex at the plasma membrane or mitochondria depending on the DNCR2 construct, and thus activate gene editing. A researcher could then temporally control the dissociation of the ternary complex, localize AcrIIA4 back into the nucleus and inhibit gene editing.

Materials and Methods

Cloning

Primers were designed using ApE software and purchased from IDT. For all PCR fragments, segments were amplified on ProFlex PCR system. PCR cycles were as follows: 98C for 10 s, annealing temp 30 s, 72C for 30s per kb, repeated for 25 cycles. Fragments were combined with DNA HiFi assembly mix (NEB) and incubated for 1 hr at 50C for Gibson assembly. Gibson products were then transformed into DH5α E. Coli competent cells (NEB) and grown on Carbenicillin LB agar plates. Colonies were picked under sterile conditions, grown in LB Broth at 37C overnight, and DNA was harvested with Qiagen miniprep kit. **Table 1** shows the constructs and amino acid sequences.

Table 1. Constructs used for transfection in Grazoprevir chase experiment

Construct	Sequence (Amino acid for constructs)
Tom20-DNCR2-EGFP-P2A-3xNLS-GNCR1	MVGRNSAIAAGVCGALFIGYCIYFDRKRRSDPNFSSDEEEARELIERAKEAAERAQEAAERT GDPRVRELARELKRLAQEAAEEVKRDPSSSDVNEALKLIVEAIEAAVDALEAAERTGDPEVR ELARELVRLAVEAAEEVQRNPSSSDVNEALHSIVYAIEAAIFALEAAERTGDPEVRELARELVR LAVEAAEEVQRNPSSRNVEHALMRIVLAIYLAEENLREAEESGDPEKREKARERVREAVERA EEVQRDPSGWLNHEQKLISEEDLGSGTGSGTMVSKGEELFTGVVPILVELDGDVNGHKFSV SGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPE GYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYI MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPN EKRDHMVLLEFVTAAGITLGMDELYKGSGATNFSLLKQAGDVEENPGPMDPKKKRKVDPK KKRKVDPKKKRKVGSGDIEKLCKKAEEEAKEAQEKADELRQRHPDSQAAEDAEDLANLAVA AVLTACLLAQEHPNADIAKLCIKAASEAAEAASKAAELAQRHPDSQAARDAIKLASQAARA VILAIMLAAENPNADIAKLCIKAASEAAEAASKAAELAQRHPDSQAARDAIKLASQAAEAVE RAIWLAAENPNADIAKKCIKAASEAAEEASKAAEEAQRHPDSQKARDEIKEASQKAEEVKER CKSGTGYPYDVPDYA*
mCherry-NS3aD168A	MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLP FAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQD GEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYD AEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKGSGT GDYKDDDDKKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVS TATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC TCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSAGGPLLCPAGHAVGIFRAAVST RGVAKAV**A**FIPVESLETTMRSP*
mCherry-NS3aR155K	MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLP FAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQD GEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYD AEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKGSGT GDYKDDDDKKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVS TATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC TCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSAGGPLLCPAGHAVGIF**K**AAVST RGVAKAVDFIPVESLETTMRSP*
mCherry-NS3aA156T	MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLP FAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQD GEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYD AEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKGSGT GDYKDDDDKKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVS TATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC TCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSAGGPLLCPAGHAVGIFR**T**AVST RGVAKAVDFIPVESLETTMRSP*

Tissue Culture

Cells were cultured on a T75 flask in Dulbecco's Modified Eagle Medium (DMEM) (Gibco) with 10% Fetal Bovine Serum (FBS) (Gibco) and passaged at 90% confluency. For microscopy experiments, Hek293T cells were plated in a 12-well plate (ThermoFisher) on top of an 18 mm cover slip (ThermoFisher) with a cell density of 20,000 cell/mL.

Transfection

A 1:3 plasmid ratio was used for NS3a-mCherry and TOM20-DNCR2-EGFP-P2A-3xNLS-GNCR1 constructs maintaining a total DNA concentration of 1 µg. Turbofectin8.0 reagent was used according to supplier's recommendation of 3 µL of turbofectin8.0 (Origene) to 1 µg of DNA. DNA and Turbofecin8.0 were added to 100 µL of serum free DMEM (Gibco), allowed to incubate for 15 minutes, and then added dropwise to each well.

Microscopy

Cells were fixed with 4% paraformaldehyde (Electron Microscopy Sciences) in DPBS (Gibco) for 8 minutes. The coverslips were removed, washed in DI water, placed on a glass slide with Fluoromount-G (Southern Biotechnology), and then sealed with clear nail polish. The slides were then analyzed at the UW Keck Center using the SP8X Leica confocal microscope and images were processed with EzColocalization software.

Western Blot

Protein samples were loaded on Any kD™ Mini-PROTEAN® TGX™ precast polyacrylamide gel (BioRad) with Precision Plus Protein Ladder (BioRad). The gel was run with 1X MOPS buffer, then the gel was placed in the center of the Transblot® Turbo™ Midi transfer packs, and then placed in the Trans-Blot® Turbo™ instrument which transferred the proteins to the nitrocellulose film. The film was placed in Odyssey® blocking buffer for 1 hour

on a rotating stage, and then primary antibodies, Myc-Tag (71D10) Rabbit mAb (CST #2278) and α-Tubulin (DM1A) Mouse mAb (CST #3873) at 1:1,000 and 1:500 dilution respectively, were added and placed on a rotating stage overnight at 4C. The nitrocellulose film was washed three times with 1X TBST buffer, and secondary antibodies, IRDye® 800CW Goat anti-Rabbit and IgG (Li-Cor) IRDye® 680LT Goat anti-Mouse IgG (Li-Cor), were added in with Odyssey® blocking buffer for 1 hour at room temperature. Then the film was washed twice with 1X TBST and imaged with Li-Cor Odyssey® infrared imager. The images were processed using ImageStudio™ Lite software.

Chapter 2: Increasing the affinity of apo NS3a reader

Introduction

A first generation engineered protein switch was developed in order to understand RAS-ERK signaling dynamics. This switch, called chemically-inducible activator of RAS (CIAR), is composed of an activated guanine exchange factor, Son of Sevenless (SOScat), and connected to each termini is an anti-apoptotic protein Bcl-xL and binding peptide BH3. With addition of a chemical disruptor, BH3 no longer binds to Bcl-xL and SOScat is capable of catalyzing the exchange of RAS-GDP to RAS-GTP[11, 12]. Because Bcl-xL and BH3 are endogenous to mammalian cells, using small molecules to disrupt the BH/ BcL-xL interaction can have unintended off-target effects. This led to the development of a second generation switch composed by replacing BH3 and Bcl-xL with the HCV serine protease NS3a and a peptide inhibitor of NS3a termed ANR (*apo* NS3a reader)[13]. Initial experiments demonstrated that this switch had slightly higher activity in the absence of drug, which needs further optimization to overcome this “leakiness” problem.

To correct unwanted basal activity of the protein switch, it is necessary to increase the affinity between ANR and NS3a. An assay was developed by Dr. Deiter from the Maly lab to measure the interaction between ANR and NS3a (**Figure 7**).

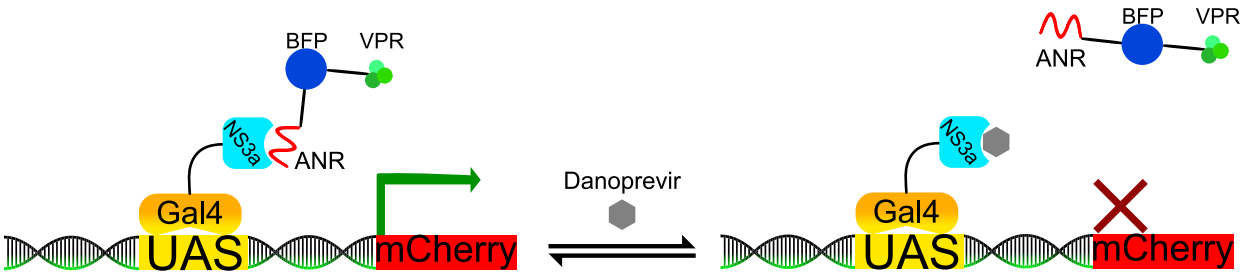

Figure 7. ***Schematic of mCherry fluorescence reporter assay.*** Diagram depicting reporter assay. When ANR is bound to NS3a VPR will induce transcription of mCherry fluorescent protein, which can be measured using flow cytometry. BFP is included as a gate for determining plasmid incorporation. The addition of danoprevir prevents mCherry transcription.

In this system Gal4UAS-mCMV-mCherry CMV-Gal4DBD-NS3a was stably integrated into HEK293T landing pad cells using lentiviral infection and sorted with FACS using iRFP670 fluorescent protein for detecting gene integration. The HEK293T landing pad cells were developed by the Fowler lab and enable the incorporation of ANR-VPR through DNA recombination[14]. This recombination site allowed each GFP-ANR-VPR construct to be integrated in the same location of the genome with the same expression level across the ANR variants that would be included in a library-based screen.

ANR is a 21 amino acid peptide, or 63 nucleotides, with a K_i of about 10 nM for NS3a[15]. The ANR library used a doped oligo strategy, meaning that for every nucleotide in the sequence 3% was an equal mixture of three bases not in the wild type sequence, while 97% was the same nucleotide as wild type ANR[16]. Using this strategy created a single ANR variant for each amino acid position giving ~400 possible mutants to increase binding affinity. The tighter the binder the higher the mCherry fluorescence should be which can be sorted with FACS and sequenced using next-generation sequencing.

Results

Deep sequencing data from the first generated library provided several residues within ANR that warranted further analysis of binding affinity (**Figure 8A**). Our approach for testing the relative binding affinity for ANR variants was to use the established reporter system (**Figure 7**). ANR variants were cloned into the original vector and transfected into the reporter cell line and then analyzed using flow cytometry. Original experiments were conducted with single ANR variants which showed tighter binding in G1E, D4E, Y8F, and D16N mutations, which ultimately led to the combination of single variants into a triple ANR variant (**Figure 8B**).

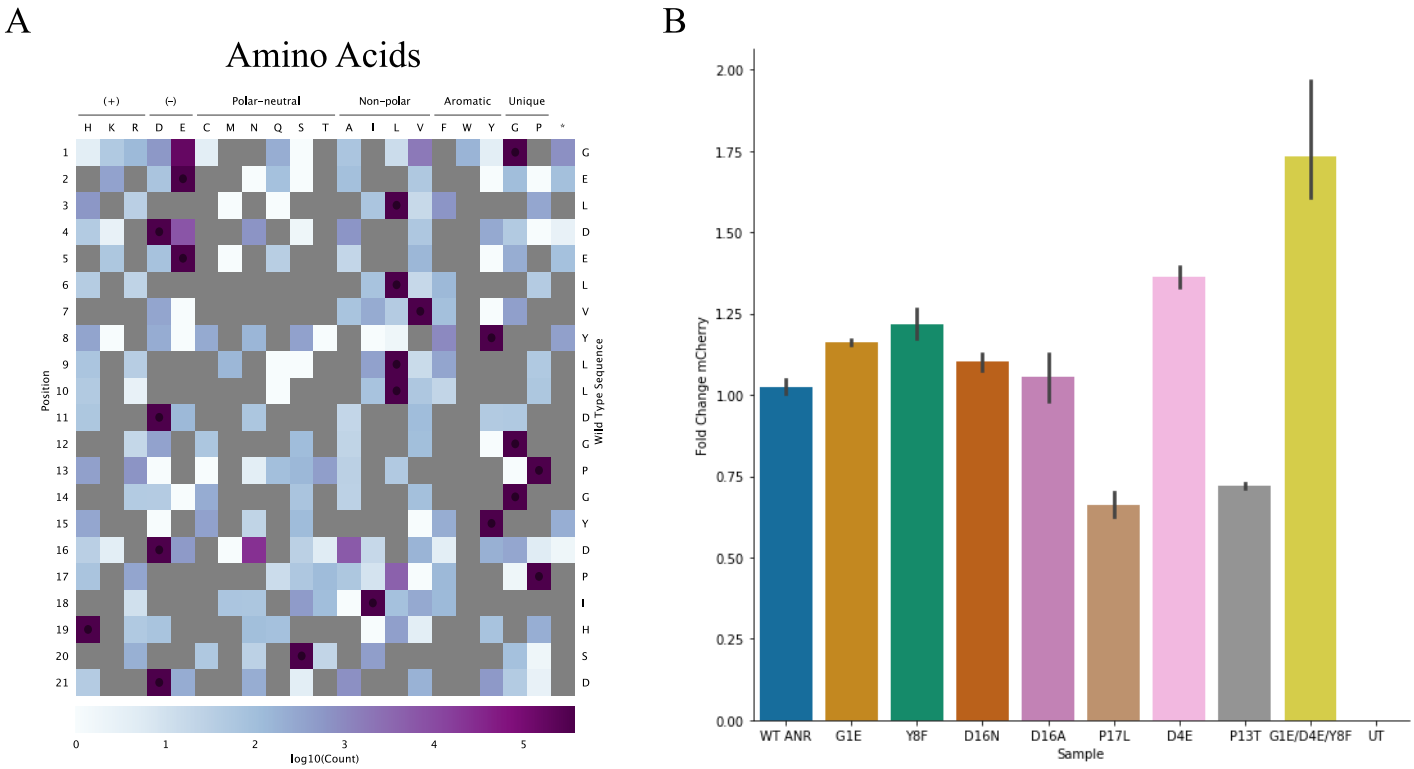

Figure 8. ***Data from first round ANR library.*** A) Deep sequencing map with frequency of mutations. Grey = nothing observed; black dot = wild type residue. B) Fold change in median mCherry fluorescence for each ANR variant compared to wild type.

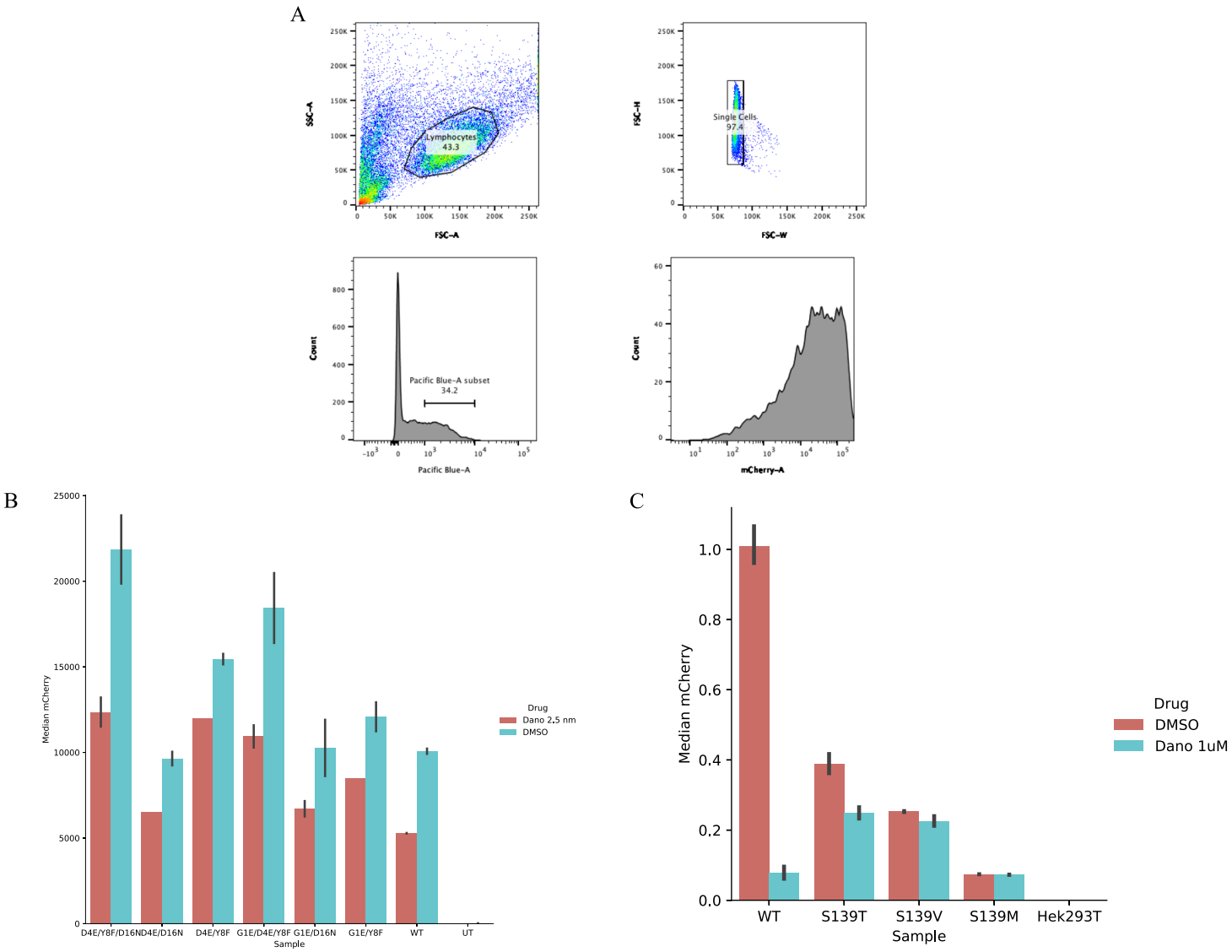

Figure 9. ***Testing of NS3a and ANR variants with flow cytometry.*** A) Gating strategy for FACS analysis. B) Median mCherry fluorescence for double and triple ANR variants. C) Normalized median mCherry fluorescence for NS3a active site variants, S139M represents the other 6 variants tested but not shown (other 6 mutations are S139 to F, H, L, N, W, and Y).

The G1E/D4E/Y8F variant provided the highest fold difference in mCherry fluorescence and led us to hypothesize that these variants might have additive effects on binding. To understand which residues were providing the most energetically favorable interactions, combinations of double and triple ANR variants were cloned and tested. At this point, all components of the reporter assay were moved to a single plasmid to increase transfection efficiency and cloning

capabilities, and the fluorescent sorting gate was changed to BFP (**Figure 9B**). The D4E/Y8F residues seem to contribute significantly for ANR binding affinity, while more terminal residues such as G1E and D16N contribute minimally to binding, which is congruent with single ANR variant data. While these results were promising, it was apparent that these variants might not provide sufficient binding affinity for future applications. This led to the screening of NS3a mutants as an additional method of increasing binding interactions.

Conventionally, some applications use a catalytically dead NS3a serine protease where the catalytic serine residue is mutated to alanine (S139A). However, since ANR is a peptide inhibitor of active NS3a it was thought that something other than alanine might provide added contact for higher affinity. We generated nine NS3a active site mutants and screened them using the mCherry fluorescence assay (**Figure 9C**). Unfortunately, most of these mutations had detrimental effects to ANR binding and also were not disrupted with the addition of danoprevir, with the exception of serine to threonine, rendering the chemically disruptable system non-functioning. With these results, we set out to design another library to further optimize binding affinity of ANR.

Using the highest affinity ANR variant became the basis for generating a new library. While considering which variants to use, it was noted that it would also be useful to have a shorter ANR peptide for future experiments. The crystal structure of ANR bound to NS3a revealed that the last four N-terminal residues of the sequence were missing, suggesting that they may not be necessary for binding affinity. Subsequently, the last four residues of ANR (I, H, S, and D) were removed giving a truncated ANR sequence that was tested (**Figure 10B**). It was shown that truncated ANR and wild type (WT) ANR had little difference in fluorescence output.

Following this result, each of the previous double and triple variants were truncated and analyzed with flow cytometry (**Figure 10C**).

A

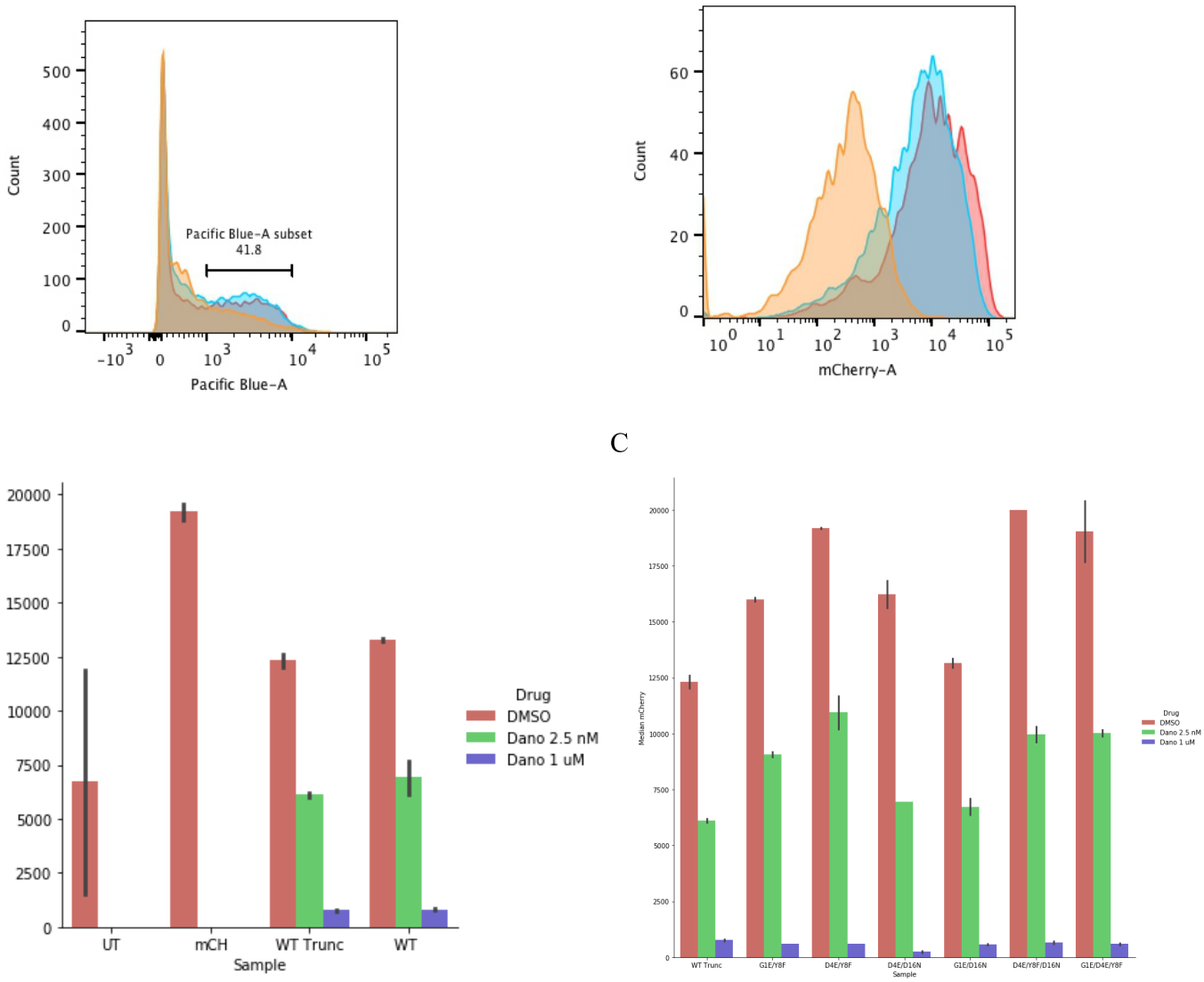

Figure 10. ***Flow cytometry analysis of truncated ANR.*** A) Example of BFP positive cell gate and median mCherry fluorescence curve with decreasing [dano] towards the right. B) Comparison of truncated ANR to wt ANR. mCH is a control mCherry plasmid transfected into HEK293T cells. C) Median mCherry signal with truncatedANR variants.

The D4E/Y8F, G1E/D4E/Y8F, and D4E/Y8F/D16N truncated variants show the most promise for the next starting point in the new library. Since G1E/D4E/Y8F added more charge to the peptide which could lead to unintended consequences, it was decided to use the variants that kept

the same overall charge as WT ANR. A titration of WT and D4E/Y8F, D4E/Y8F/D16N truncated ANR variants was performed to quantify the EC_{50} values (**Figure 11**). It is shown that the EC_{50} values for the double and triple ANR variants differ by less than 1 nM, confirming that variants closer to the termini are less crucial to binding than internal residues.

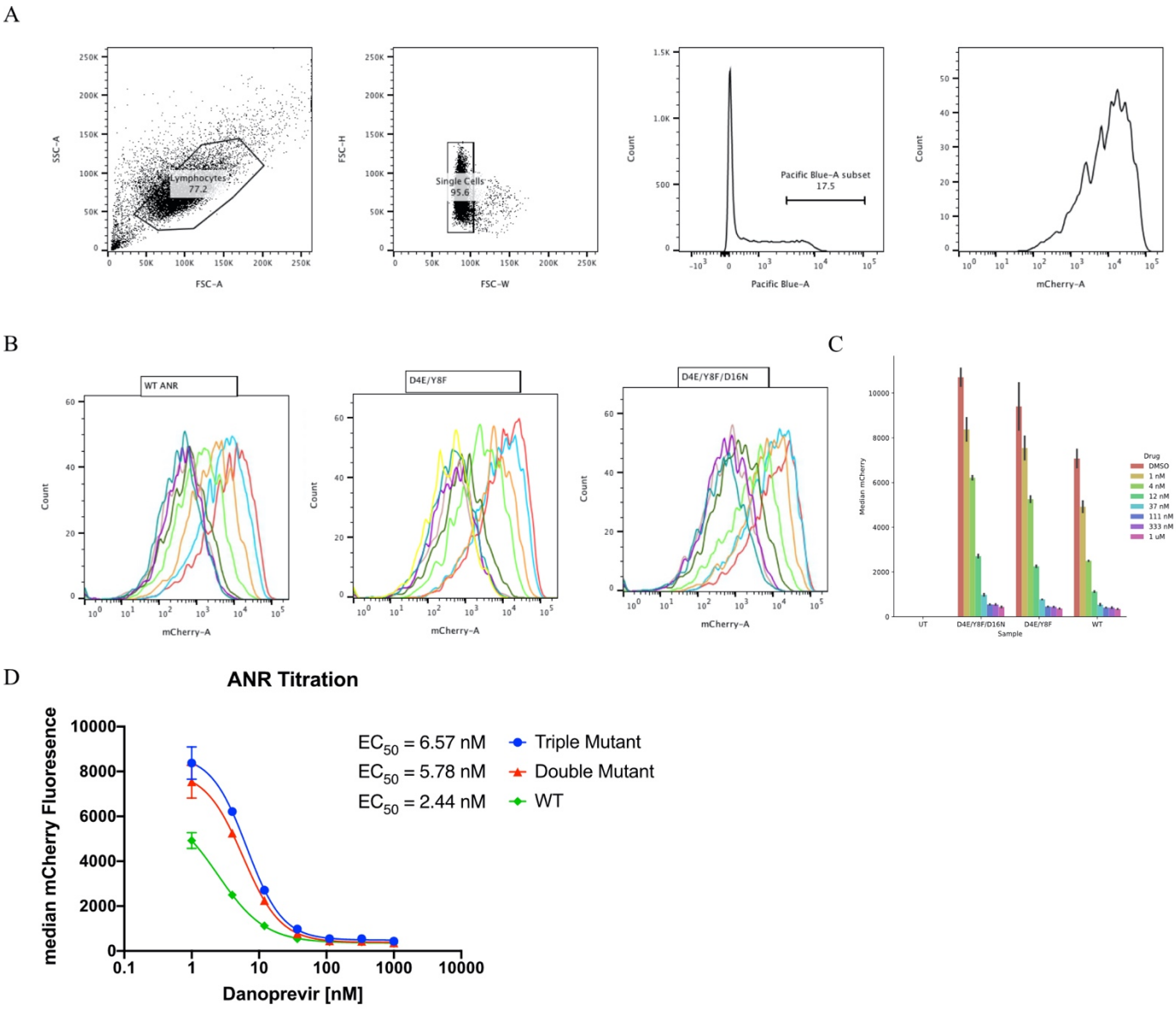

Figure 11. ***Titration of ANR variants with danoprevir.*** A) FACS analysis gating strategy using LSRII flow cytometer. B) Fluorescence curves for WTand variant ANR peptides with decreasing danoprevir concentrations to the right. C) Median mCherry fluorescence plots. D) Titration curve using median fluorescent values and reported EC_{50} values, ANR mutants have the last 4 residues truncated.

Discussion

The first ANR library led to the discovery of a triple mutant variant with an increased IC_{50} of about 2.7 times greater than wild type ANR. This translates to higher concentrations of drug needed to disrupt the NS3a:ANR interaction, but is it enough for creating a functioning biorthogonal switch comparable to CIAR? We have our doubts and further optimization is required for increased binding affinity. One issue with the first library was the convergence to the WT ANR sequence, giving no new variants after successive sorting. Plans were made to make a second library using more informed decisions with hopes of increasing ANR affinity. We hypothesized that there could be conformation restraints of internal neighboring residues and mutating only a single residue can lead to a change in conformation that is unfavorable to binding. To counter this, we devised a co-variant NNK library. Looking at the crystal structure, it was decided that five co-variant libraries would be generated, mutating two neighboring residues per library. NNK is a mixture of nucleotides where N is 1:1:1:1 ratio of A, C, T, and G nucleotides and K is a 1:1 ratio of G and T nucleotides. This will generate 5,120 DNA variants of ANR and approximately 2,000 protein variants. Plans to use the same strategy of generating a stable cell line and using the landing pad cells for variant integration were unable to be carried out due to the current pandemic.

Materials and Methods

Cloning

Primers were purchased from IDT and designed with SnapGene® software. DNA fragments were amplified with PCR. Cycles were as follows: 98C for 10 s, annealing temp 30 s, 72C for 30s per kb, repeated for 25 cycles. DNA fragments were separated with Gel electrophoresis and extracted with a DNA gel extraction kit (Zymo Research). DNA fragments

were combined with DNA HiFi assembly mix (NEB) and incubated at 50C for 1 hour. DNA was then transformed in E. Coli DH5α competent cells (NEB), harvested with DNA miniprep (Qiagen) and analyzed with Sanger sequencing (GENEWIZ).

Quickchange

Primers were designed with mutant nucleotides and the entire plasmid was amplified with PCR. The PCR cycles were adjusted to 15 cycles instead of 25. The parent plasmid was digested using DpnI digestion enzyme (NEB) with Fast Digest buffer (NEB). The plasmids were transformed, harvested and analyzed using DH5α E. Coli competent cells (NEB) as mentioned in the cloning method. Cloning and Quickchange generated sequences can be seen in **Table 2**, other constructs differ only by mutations specified in experiments.

Table 2. Constructs used in transfection for ANR experiment

Construct	Sequence (Amino acids)
mCMV-Gal4UAS-mCHerry-CMV-Gal4DBD-NS3a-P2A-ANR(WT)-BFP-VPR	MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVESRLERLEQLFLLIFPREDLDMILKMDSLQDIKALLGTPAAASTAGSGGMAKGSVVIVGRINLSGDTAYSQQTRGLEGCQETSQTGRDKNQVEGEVQVVSTATQSFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSAGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSPGSGATNFSLLKQAGDVEENPGPGSASGM**GELDELVYLLDGPGYDPIHSD**GVLSGSGTGSGTGSGTGTTSGTGTGGSTGEQKLISEEDLGSGSSELIKENMHMKLYMEGTVDNHHFKCTSEGEGKPYEGTQTMRIKVVEGGPLPFAFDILATSFLYGSKTFINHTQGIPDFFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGCLIYNVKIRGVNFTSNGPVMQKKTLGWEAFTETLYPADGGLEGRNDMALKLVGGSHLIANIKTTYRSKKPAKNLKMPGVYYVDYRLERIKEANNETYVEQHEVAVARYCDLPSKLGHKLNGSGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSPKKKRKVGSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDTSLF*

mCMV-Gal4UAS-mCHerry-CMV-Gal4DBD-NS3a-P2A-ANR(WT Truncation)-BFP-VPR	MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTE VESRLERLEQLFLLIFPREDLDMILKMDSLQDIKALLGTPAAASTAGSGGMAKGS VVIVGRINLSGDTAYSQQTRGLEGCQETSQTGRDKNQVEGEVQVVSTATQSFL ATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTP CTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSAGGPLLCPAGHAV GIFRAAVSTRGVAKAVDFIPVESLETTMRSPGSGATNFSLLKQAGDVEENPGPG SASGM**GELDELVYLLDGPGYDP**GVLSGSGTGSGTGSGTGTTSGTGTGGSTGE QKLISEEDLGSGSSELIKENMHMKLYMEGTVDNHHFKCTSEGEGKPYEGTQTM RIKVVEGGPLPFAFDILATSFLYGSKTFINHTQGIPDFFKQSFPEGFTWERVTTYE DGGVLTATQDTSLQDGCLIYNVKIRGVNFTSNGPVMQKKTLGWEAFTETLYPA DGGLEGRNDMALKLVGGSHLIANIKTTYRSKKPAKNLKMPGVYYVDYRLERIKE ANNETYVEQHEVAVARYCDLPSKLGHKLNGSGSDALDDFDLDMLGSDALDDF DLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSPKKKRKVGSQYLPDTDDR HRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPAPQPYPF TSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQ APAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNST DPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPP DPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMF LPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHE PVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKE EAAICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDTFLNDEC LLHAMHISTGLSIFDTSLF*

Tissue Culture

Cells were cultured in a T75 flask in Dulbecco's Modified Eagle Medium (DMEM) (Gibco) with 10% Fetal Bovine Serum (FBS) (MiliporeSigma) and passaged at 90% confluency. For FACS experiments, Hek293T cells were plated in a 12-well plate (ThermoFisher) at a density of 200,000 cells/mL.

Transfection

1 µg of DNA of mCMV-Gal4UAS-mCHerry-CMV-Gal4DBD-NS3a-P2A-ANR-BFP-VPR plasmid was added to each well with transfection reagent. Turbofectin8.0 reagent was used according to supplier's recommendation of 3 µL of turbofectin8.0 (Origene) to 1 µg of DNA. DNA and Turbofecin8.0 were added to 100 µL of serum free DMEM (ThermoFischer), allowed to incubate for 15 minutes, and then added dropwise to each well.

Flow Cytometry

Cells were detached from plates with 125 µL of 0.25% Trypsin-EDTA (ThermoFisher) and incubated for about 3 minutes or until all cells detached. Trypsin was quenched with 875 µL of DMEM (Gibco) and cells were transferred to 1.5 mL Eppendorf vials. Cells were spun down at 4000 rpm for 10 minutes, and supernatant was aspirated. Cells were then resuspended in 275 µL of DPBS (Gibco) with 10% FBS (MiliporeSigma), transferred to sterile 5 mL Roundbottom tubes (ThermoFisher), taken to Flow Core center at UW and analyzed on LSRII BD™. Flow cytometry data was analyzed using FlowJo software. Titration curves and IC50 values were calculated with GraphPad Prism8.

REFERENCES

1. Stanton, B. Z.; Chory, E. J.; Crabtree, G. R. Chemically Induced Proximity in Biology and Medicine. *Science* **2018**, *359* (6380), eaao5902.

2. Banaszynski, L. A.; Liu, C. W.; Wandless, T. J. Characterization of the FKBP·Rapamycin·FRB Ternary Complex. *J. Am. Chem. Soc.* **2005**, *127* (13), 4715–4721.

3. Douglass, E. F.; Miller, C. J.; Sparer, G.; Shapiro, H.; Spiegel, D. A. A Comprehensive Mathematical Model for Three-Body Binding Equilibria. *J. Am. Chem. Soc.* **2013**, *135* (16), 6092–6099.

4. Tanaka, H.; Kuroda, A.; Marusawa, H.; Hatanaka, H.; Kino, T.; Goto, T.; Hashimoto, M.; Taga, T. Structure of FK506, a Novel Immunosuppressant Isolated from Streptomyces. *J. Am. Chem. Soc.* **1987**, *109* (16), 5031–5033.

5. Liberles, S. D.; Diver, S. T.; Austin, D. J.; Schreiber, S. L. Inducible Gene Expression and Protein Translocation Using Nontoxic Ligands Identified by a Mammalian Three-Hybrid Screen. *Proc Natl Acad Sci USA* **1997**, *94* (15), 7825.

6. Voß, S.; Klewer, L.; Wu, Y.-W. Chemically Induced Dimerization: Reversible and Spatiotemporal Control of Protein Function in Cells. *Current Opinion in Chemical Biology* **2015**, *28*, 194–201.

7. Foight, G. W.; Wang, Z.; Wei, C. T.; Jr Greisen, P.; Warner, K. M.; Cunningham-Bryant, D.; Park, K.; Brunette, T. J.; Sheffler, W.; Baker, D.; et al. Multi-Input Chemical Control of Protein Dimerization for Programming Graded Cellular Responses. *Nature Biotechnology* **2019**, *37* (10), 1209–1216.

8. Kholodenko, B. N. Cell-Signalling Dynamics in Time and Space. *Nature Reviews Molecular Cell Biology* **2006**, *7* (3), 165–176.

9. Romano, K. P.; Ali, A.; Aydin, C.; Soumana, D.; Özen, A.; Deveau, L. M.; Silver, C.; Cao, H.; Newton, A.; Petropoulos, C. J.; Huang, W.; Schiffer, C. A. The Molecular Basis of Drug Resistance against Hepatitis C Virus NS3/4A Protease Inhibitors. *PLoS Pathog* **2012**, *8* (7), e1002832.

10. Stauffer, W.; Sheng, H.; Lim, H. N. EzColocalization: An ImageJ Plugin for Visualizing and Measuring Colocalization in Cells and Organisms. *Scientific Reports* **2018**, *8* (1), 15764.

11. Rose, J. C.; Huang, P.-S.; Camp, N. D.; Ye, J.; Leidal, A. M.; Goreshnik, I.; Trevillian, B. M.; Dickinson, M. S.; Cunningham-Bryant, D.; Debnath, J.; Baker, D.; Wolf-Yadlin, A.; Maly, D. J. A Computationally Engineered RAS Rheostat Reveals RAS–ERK Signaling Dynamics. *Nature Chemical Biology* **2017**, *13* (1), 119–126.

12. Dieter, E. M.; Maly, D. J. Chapter Six - A Chemically-Controlled System for Activating RAS GTPases. In *Chemical and Synthetic Biology Approaches To Understand Cellular Functions - Part C*; Shukla, A. K., Ed.; Methods in Enzymology; Academic Press, 2020; Vol. 633, pp 103–117.

13. Cunningham-Bryant, D.; Dieter, E. M.; Foight, G. W.; Rose, J. C.; Loutey, D. E.; Maly, D. J. A Chemically Disrupted Proximity System for Controlling Dynamic Cellular Processes. *Journal of the American Chemical Society* **2019**, *141* (8), 3352–3355.

14. Matreyek, K. A.; Stephany, J. J.; Fowler, D. M. A Platform for Functional Assessment of Large Variant Libraries in Mammalian Cells. *Nucleic Acids Research* **2017**, *45* (11), e102–e102.

15. Kügler, J.; Schmelz, S.; Gentzsch, J.; Haid, S.; Pollmann, E.; van den Heuvel, J.; Franke, R.; Pietschmann, T.; Heinz, D. W.; Collins, J. High Affinity Peptide Inhibitors of the Hepatitis C Virus NS3-4A Protease Refractory to Common Resistant Mutants. *Journal of Biological Chemistry* **2012**, *287* (46), 39224–39232.

16. Neylon, C. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Research* **2004**, 32(4), 1448–1459.

17. Fersht, A. *Structure and Mechanism in Protein Science*; Series in Structural Biology; WORLD SCIENTIFIC, 2017; Vol. Volume 9.