Structure-Function Characterization of the VEGF-Ax mRNA Readthrough Element

G5 progress update submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy

Nico Wagner
Molecules, Cells, and Organisms
Department of Molecular and Cellular Biology, Harvard University

November 19, 2018
2:30 -4:30 PM
Northwest 195.05

Your Signature

Victoria D'Souza

Introduction

For my PhD, I have been working on vascular endothelial growth factor A (VEGF-A), a system recently shown to undergo a programmed ribosomal readthrough event, bypassing the stop codon and creating an extended protein with a frequency of 9-25% [1]. This was the first example of translation regulation via stop codon readthrough in a human mRNA. Furthermore, it was the first readthrough event discovered that requires a *trans*-acting factor, the heterogeneous nuclear ribonucleoprotein particle (hnRNP) A2/B1.

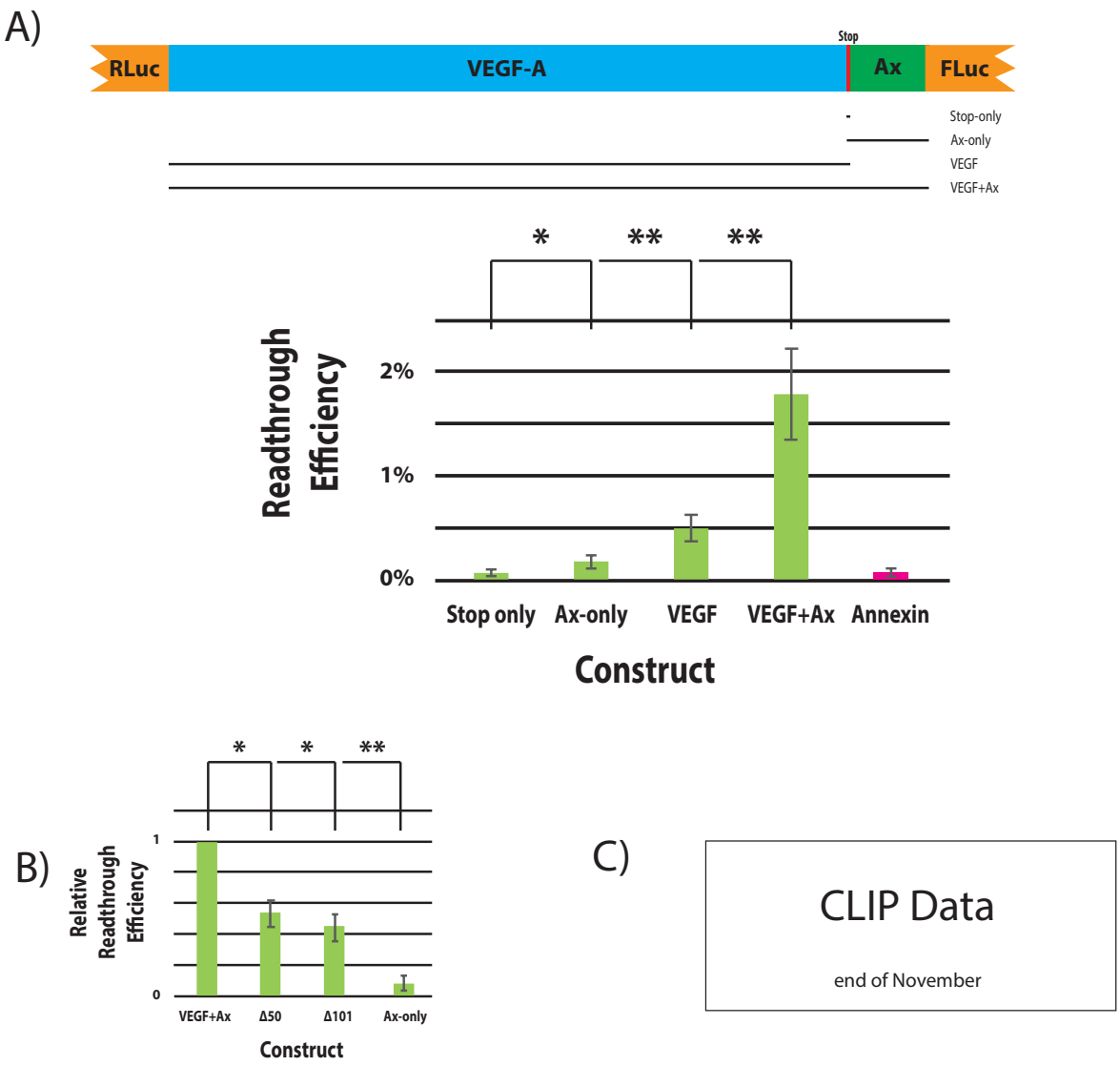
Here I show that VEGF-A contains a 5' readthrough signal and a second signal just downstream of the stop codon. Further, I show that these two signals, separated by 500 nucleotides, are able to coordinate with each other through hnRNP A2/B1.

The presence of a distal 5' element conferring ribosomal readthrough is unusual. In retroviruses, where ribosomal recoding has been studied more extensively, readthrough signals tend to be localized in the immediately vicinity of the stop codon. While distant downstream elements conferring to frameshifting in *Pea enation mosaic virus* [2] and readthrough in *Potato leafroll virus* [3] have been studied, to the best of my knowledge, there are no examples of distant upstream signals.

In this DAC update I show a series of preliminary figures that will go into a publication. Along with each figure, I describe the main takeaway messages. *Italicized text refers to experiments not yet completed. A proposed timeline for these experiments can be found in the panels where the resulting data will be shown.*

Lastly, I present a series of proposed experiments for my 'final direction' toward publication.

Figure 1: VEGF-Ax mRNA contains multiple readthrough signals



The aim of figure 1 is to show that two signals are required for readthrough of the VEGF-A gene.

1A) Readthrough efficiency was determined using a dual luciferase reporter assay in which a sequence of interest was cloned between renilla and firefly luciferases on the p2Luc plasmid [4]. Upon transfection into bovine aortic endothelial cells (BAOEC), readthrough efficiencies were calculated. In doing so, I was, in agreement with another study [5], unable to reproduce the 9-25% readthrough previously attributed to the 63 nucleotides of the VEGF-Ax element [1]. Rather, I observed 0.2% readthrough, a ~4-fold increase over the natural leakiness of the UGA stop codon and Annexin A2 mRNA, a gene with no known readthrough activity. When cloning

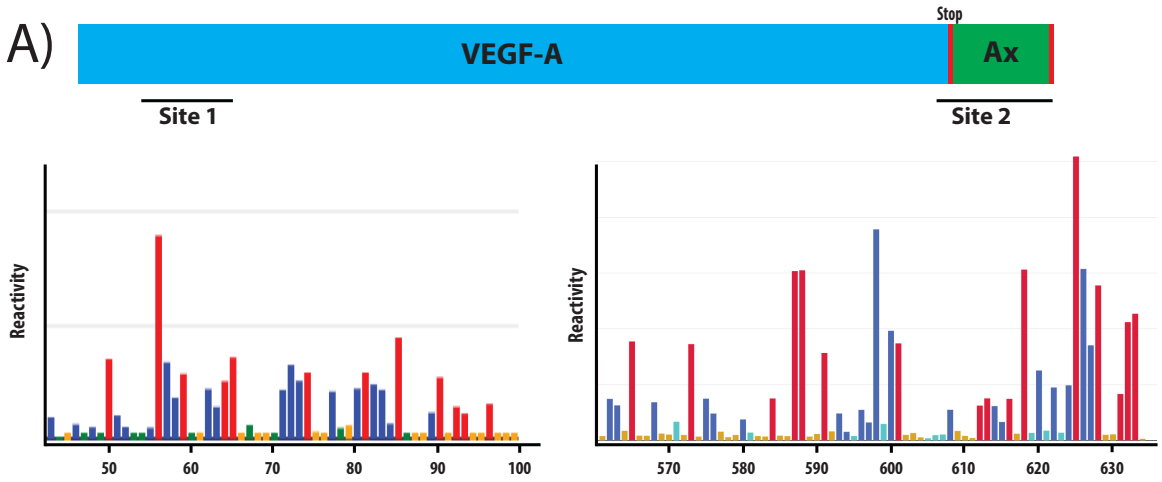
the full VEGF-Ax mRNA into p2Luc, I was able to increase readthrough to 1.8%. Interestingly, the VEGF mRNA lacking the Ax element was able to confer 0.5% readthrough, a 10-fold increase over natural leakiness. **Taken together, these results indicate that the full VEGF-Ax mRNA contains multiple readthrough signals: a 5' signal leading to a 10-fold readthrough increase over natural leakiness and a 3' signal within the Ax-element leading to an additional 4-fold increase in readthrough.**

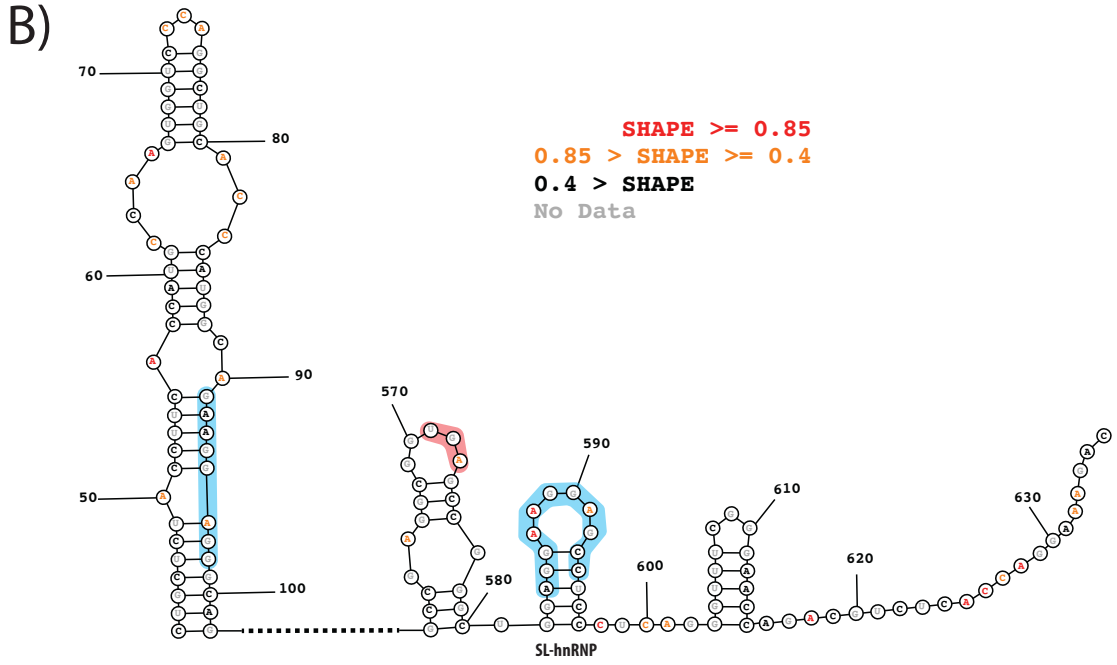
1B) Truncation experiments show that an element within the first 100nt is important for readthrough to occur.

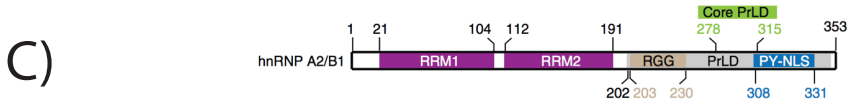
1C) Given that readthrough is dependent upon hnRNP A2/B1 binding in the Ax-element, I used cross-linking immunoprecipitation (CLIP) to test for additional hnRNP A2/B1 binding sites.

Once I receive my CLIP data, they will be inserted and discussed here. Results are expected any day as samples were submitted for sequencing on 10/29/2018.

Figure 2: Biochemical analysis of readthrough signals

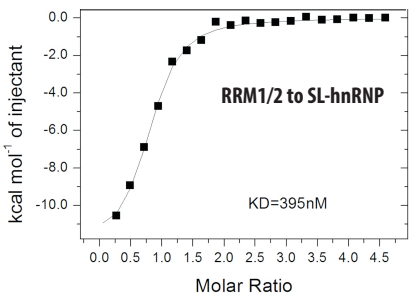






Site 1 ITC curve

end of November



The aim of figure 2 is to biochemically analyze the two hnRNP binding sites, site 1 and site 2.

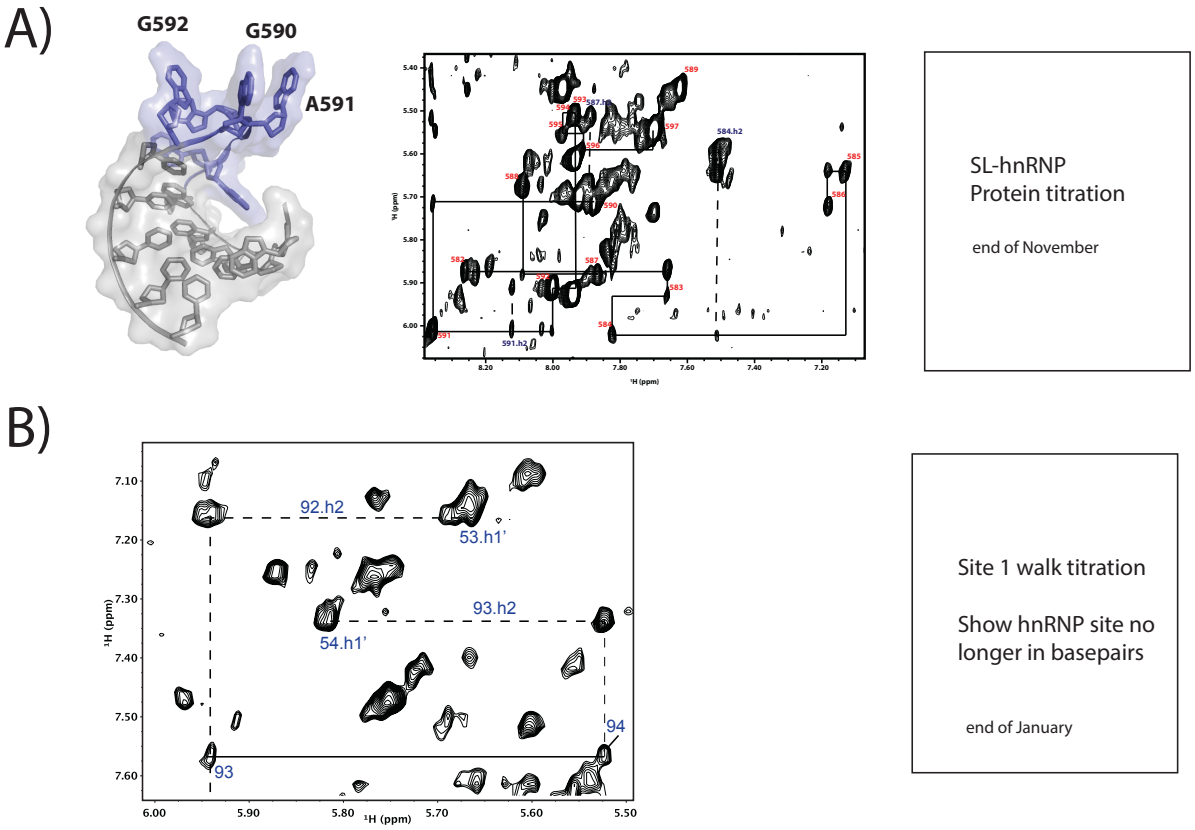
2A) Given the presence of two hnRNP A2/B1 binding sites, a collaborator and I performed DMS-MaPseq experiments [6] on the full VEGF-A mRNA and focused more closely on sites 1 and 2. This figure shows the raw population averaged DMS reactivities of these two sites.

2B) Shown are the predicted secondary structures of sites 1 and 2. Highlighted in blue are the hnRNP binding sites and shown in red is the stop codon involved in readthrough. Interestingly, the site 1 binding element is occluded in base pairs while chemical probing of site 2 reveals the presence of a preformed stem-loop (SL-hnRNP) containing the hnRNP binding consensus sequence. This predicted structure is in agreement with previous studies showing that hnRNPs of the A and B family tend to bind looped regions of RNA [7].

2C) I synthesized both the 60nt site 1 RNA and 16nt SL-hnRNP of site 2 and performed isothermal titration calorimetry (ITC) experiments to determine binding affinity between these RNAs and hnRNP A2/B1. In accordance with previous studies [7, 8], the hnRNP construct used for binding assay consisted of 2 RNA recognition motifs (RRMs), but lacked the aggregation-prone C-terminal domain.
For SL-hnRNP, we obtained a binding curve with K_d = 395nM. This K_d is in agreement with previously published binding affinity of similar sequences [1, 8].

Once completed, the ITC binding curve of site 1 will be added to this figure.

Figure 3: Structural analysis of readthrough signals

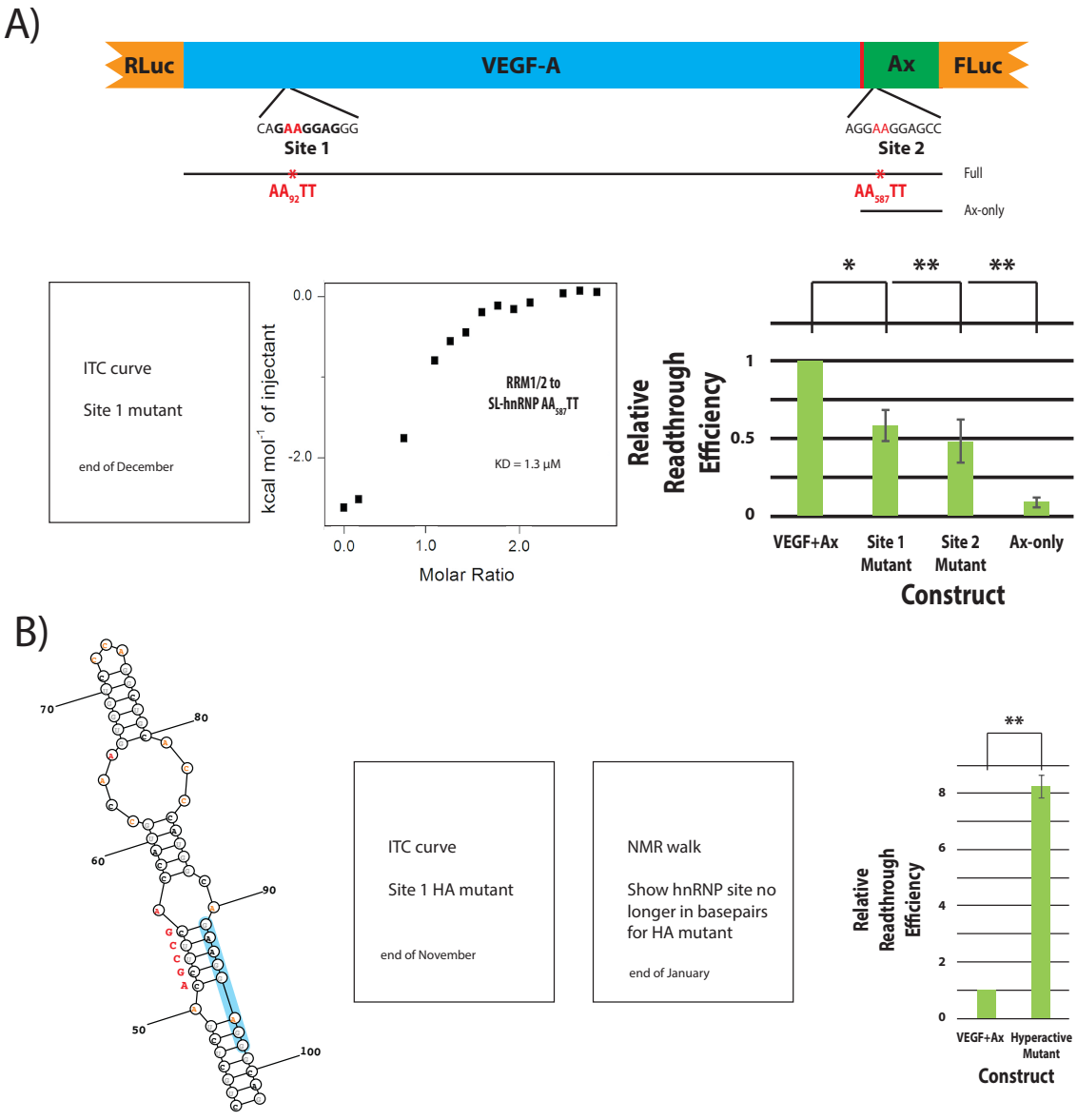


The aim of this figure is to explore the structure of the pre-formed SL-hnRNP and the differences of protein-bound & unbound RNA using NMR.

3A) Left: In accordance with DMS-MaPseq secondary structure predictions, SL-hnRNP folds into a stem-loop conformation. Interestingly, two bases within the loop required for hnRNP binding point out, ready to bind to hnRNP. This is similar to hnRNP A1 binding to an HIV-1 exon splicing silencer [9].
Middle: Shows the walks used to assign the peaks of this structure.
Right: *Will show which peaks shift upon titrating hnRNP A2/B1 RRMs1/2 into SL-hnRNP.*

3B) Shows a preliminary walk of site 1, confirming that the hnRNP binding is in a stem-loop conformation.
The goal for these panels is to track the peaks corresponding to the site 1 hnRNP element during protein titration. Our hypothesis is that the bases involved in hnRNP binding are base-paired in the absence of protein (as shown in DMS-MaPseq data in figure 2) and will undergo a chemical shift in the presence of protein.

Figure 4: Contribution of VEGF-A mRNA toward readthrough



The aim of this figure is to explore the contribution of the VEGF-A mRNA toward readthrough.

4A) *Shown here will be ITC binding curves of the RRM1/2 construct described in figure 2 to hnRNP binding site mutants of sites 1 and 2.*

An initial ITC run of the site 2 mutant ($AA_{587}TT$) indicates that binding is reduced, but not abolished, in this construcy. This is in accordance with previous studies of binding site mutants

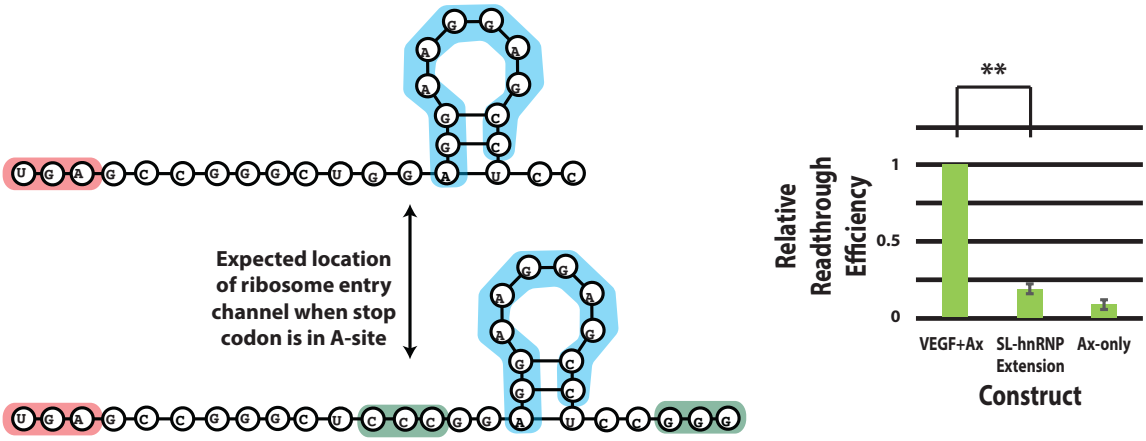
[8] which show that single and double nucleotide mutations do not abolish binding. This reduced binding also agrees with my *in-cellulo* data, which show a ~2-fold readthrough decrease for the same mutants, both at site 1 and site 2.

We predict site 1 ITC curves for $AA_{92}TT$ to exhibit a similar behavior.

4B) Using the secondary structure predicted through chemical probing, I designed a site 1 mutant which would prevent the hnRNP consensus sequence from forming a stem, making it more accessible to hnRNP and leading to hyperactive readthrough levels. When testing for readthrough efficiency of this site 1 hyperactive mutant, I found it to be an ~8-fold increase over wild-type VEGF levels. **This experiment shows that there is cross-talk between the 5' signal and the stop codon.**

Using NMR, we would like to show that the site 1 consensus sequence is no longer base paired in this hyperactive mutant. Additionally, we would like to show via ITC that the hnRNP is able to more easily access the binding site.

Figure 5: Contribution of hnRNP A2/B1 toward readthrough



The aim of this figure will be to show how the hnRNP contributes toward readthrough.

One data point for this I already have is that the spacing between the stop codon and the site 2 hnRNP binding sequence matters. By increasing the spacing (green highlighted nucleotides), readthrough drops to levels as if the Ax-element were absent. This is intriguing since the spacing between the stop codon the hnRNP binding site matches the length of the ribosome entry channel. One model for this is that the hnRNP, when bound at site 2, leads to ribosome stalling. I have submitted a sample focusing on this for CryoEM and should hear back before the end of November.

Here I propose a series of experiments to test the hypothesis that hnRNP A2/B1 dimerizes via its C-terminal domain upon binding to both site 1 and site 2. As of submitting this report, none of these experiments have been conducted, but I hope to present some preliminary data during my DAC. As discussed at my previous DAC, this series of experiments will constitute the final direction toward wrapping up this project. As such, this is the section where I would appreciate feedback the most.

- ITC of full-length VEGF-A mRNA with RRMs1/2
  I am hoping this experiment will provide us with an n value of binding ratios.
  Complete by mid December.

- Pull down assay
  I propose to pull down VEGF-A mRNA from bovine endothelial cells after a round of crosslinking. I have designed a construct that would allow me to pull down VEGF mRNA

using a biotinylated primer and beads. After washing, I would digest the RNA and separate protein on an SDS-PAGE gel followed by Western blotting against hnRNP A2/B1.
Complete by end of December.

- Delete C-terminal domain in tissue culture
  I have created constructs which will allow me to overexpress full length hnRNP and RRMs1/2 in tissue culture. By comparing the effects of overexpressing these two constructs I hope to circumstantially show the relevance of the C-terminal domain. I hypothesize that overexpressing RRMs1/2 will show decreased readthrough relative to overexpressing the full length hnRNP.
  Complete by mid December.

- DNA-paint
  The idea with this experiment is to show that the 5' and 3' ends of the VEGF mRNA are in close proximity, which would be expected if hnRNP forms a dimer.
  Complete by end of February.

References

1. Eswarappa SM, Potdar AA, Koch WJ, Fan Y, Vasu K, Lindner D, et al. Programmed translational readthrough generates antiangiogenic VEGF-Ax. Cell. 2014;157(7):1605-18. Epub 2014/06/21. doi: 10.1016/j.cell.2014.04.033. PubMed PMID: 24949972; PubMed Central PMCID: PMCPMC4113015.
2. Gao F, Simon AE. Multiple Cis-acting elements modulate programmed -1 ribosomal frameshifting in Pea enation mosaic virus. Nucleic Acids Res. 2016;44(2):878-95. Epub 2015/11/19. doi: 10.1093/nar/gkv1241. PubMed PMID: 26578603; PubMed Central PMCID: PMCPMC4737148.
3. Xu Y, Ju HJ, DeBlasio S, Carino EJ, Johnson R, MacCoss MJ, et al. A Stem-Loop Structure in Potato Leafroll Virus Open Reading Frame 5 (ORF5) Is Essential for Readthrough Translation of the Coat Protein ORF Stop Codon 700 Bases Upstream. J Virol. 2018;92(11). Epub 2018/03/09. doi: 10.1128/JVI.01544-17. PubMed PMID: 29514911; PubMed Central PMCID: PMCPMC5952135.
4. Grentzmann G, Ingram JA, Kelly PJ, Gesteland RF, Atkins JF. A dual-luciferase reporter system for studying recoding signals. RNA. 1998;4(4):479-86. Epub 1998/06/18. PubMed PMID: 9630253; PubMed Central PMCID: PMCPMC1369633.
5. Loughran G, Howard MT, Firth AE, Atkins JF. Avoidance of reporter assay distortions from fused dual reporters. RNA. 2017;23(8):1285-9. Epub 2017/04/27. doi: 10.1261/rna.061051.117. PubMed PMID: 28442579; PubMed Central PMCID: PMCPMC5513072.
6. Zubradt M, Gupta P, Persad S, Lambowitz AM, Weissman JS, Rouskin S. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. Nat Methods. 2017;14(1):75-82. Epub 2016/11/08. doi: 10.1038/nmeth.4057. PubMed PMID: 27819661; PubMed Central PMCID: PMCPMC5508988.
7. Jain N, Morgan CE, Rife BD, Salemi M, Tolbert BS. Solution Structure of the HIV-1 Intron Splicing Silencer and Its Interactions with the UP1 Domain of Heterogeneous Nuclear Ribonucleoprotein (hnRNP) A1. J Biol Chem. 2016;291(5):2331-44. Epub 2015/11/27. doi: 10.1074/jbc.M115.674564. PubMed PMID: 26607354; PubMed Central PMCID: PMCPMC4732216.
8. Wu B, Su S, Patil DP, Liu H, Gan J, Jaffrey SR, et al. Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1. Nat Commun. 2018;9(1):420. Epub 2018/01/31. doi: 10.1038/s41467-017-02770-z. PubMed PMID: 29379020; PubMed Central PMCID: PMCPMC5789076.
9. Levengood JD, Rollins C, Mishler CH, Johnson CA, Miner G, Rajan P, et al. Solution structure of the HIV-1 exon splicing silencer 3. J Mol Biol. 2012;415(4):680-98. Epub 2011/12/14. doi: 10.1016/j.jmb.2011.11.034. PubMed PMID: 22154809.