MQP-ID-SFS-ABDV

Creation of an eGFP-PLCβ1 Fusion Protein Using CRISPR-Cas9 Mutagenesis

A Major Qualifying Project Report

Submitted to the faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

in Biochemistry

by

Kaitlyn Valla

CDR Deadline: April 26, 2018

APPROVED:

Suzanne F. Scarlata, PhD

Chemistry and Biochemistry

PROJECT ADVISOR

Abstract

Phospholipase Cβ1 (PLCβ1), most notably a member of the mammalian phosphoinositide pathway, has functions that extend beyond its phospholipid hydrolysis and calcium release induction in cells. Research has shown that PLCβ1 has been implicated in cancer cell proliferation, psychological disorders such as schizophrenia, and gene-silencing through RNA interference. The ability to accurately visualize PLCβ1 in cells without changing endogenous levels of the enzyme is paramount for biophysical experimentation. With the help of the CRISPR-Cas9 genome editing machinery, a GFP-PLCβ1 fusion protein was incorporated into the genome of cultured cells. The GFP-tagged PLCβ1 protein was visible in human neuronal cells and now allows us to characterize its function in the cytoplasm and nucleus of these cells.

Table of Contents

Acknowledgements

First and foremost I would like to thank Dr. Suzanne Scarlata for allowing me the opportunity to really explore my scientific interests in her lab. Dr. Scarlata showed both moral and financial support for my interest in CRISPR through her belief in me and my passion for science.

Next, I would like to thank Osama Garwain for his endless patience while teaching me most of the lab techniques I know today. Without Osama, I would not have been as successful with this project as I was. He was also paramount in making me feel at home in the lab and welcoming me into the Scarlata lab family.

Further, I would like to thank Dr. Dan McCollum and his post-doc Sebastian Mana-Capelli at the University of Massachusetts Medical School. Dr. McCollum's assistance with the CRISPR mutagenesis was imperative for my project's success. Also, without Sebastian, I would not have been able to construct my Cas9 plasmid and troubleshoot the problems I experienced as well as we did. The McCollum lab also provided me with the PX330 plasmid that was edited to make my specific Cas9 construct.

Finally, I would like to thank Dr. Destin Heilman for showing me that my true passions lie in biochemistry. Without Professor Heilman, I would likely not be where I am today: certain of the field of study that I want to pursue and exactly how to get there. It was with Professor Heilman that I first discovered my love of CRISPR.

Background

Phospholipase Cβ1 (PLCβ1) is a major component of the phosphoinositide pathway which, in neuronal cells, ultimately results in a release of intracellular calcium that facilitates cell-to-cell communication. PLCβ1 is an enzyme that is activated by Gαq when hormones or neurotransmitters (such as dopamine or acetylcholine) bind the G protein-coupled receptor. PLCβ1's most notable function is the hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP_2). This hydrolysis results in the formation of diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP3). It is the cellular re-localization of the IP3 to the endoplasmic reticulum that causes the intracellular calcium release (Kadamur, 2013). This hydrolysis of PIP_2 is said to happen in the TIM domain of the PLCβ1 enzyme which is where the catalytic site is located (see Figure 1).

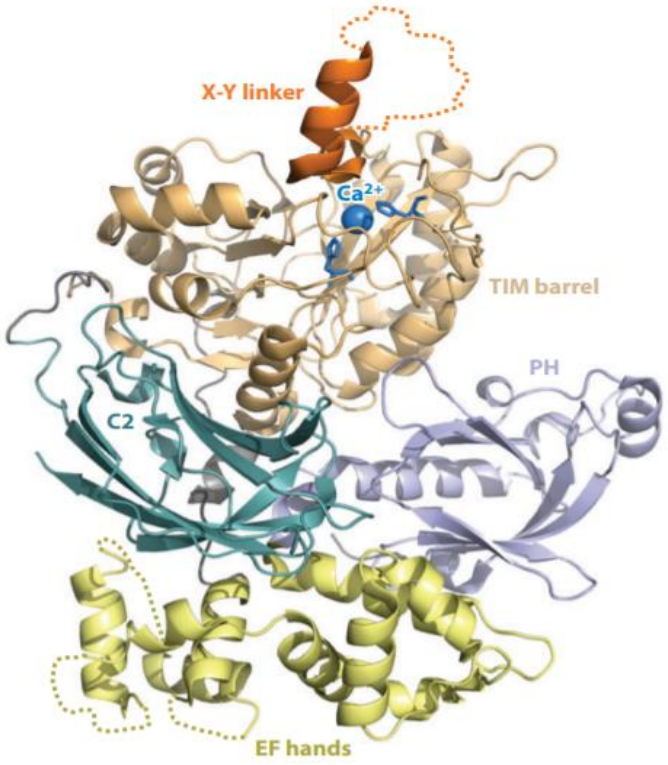


Figure 1: 3D Tertiary Structure of the PLCβ1 Enzyme – The PH domain rests at the N-terminus of the enzyme while the C2 domain rests at the C-terminus of the enzyme.

Extensive research within the Scarlata lab into the function of PLCβ1 and has revealed that PLCβ1 is responsible for far more than simply the hydrolysis of PIP_2. Research suggests that PLCβ1 is also a binding partner of Component 3 Promoter of the RNA induced silencing complex (C3PO). This finding introduces the potential for PLCβ1 to be involved in RNA interference machinery. Evidence suggests that a cytosolic population of PLCβ1 will bind to C3PO and hinder C3PO's nuclease activity therefore suppressing the silencing of specific genes (Scarlata, 2016). It was also found through immunoprecipitation experiments and mass spectrometry analysis that PLCβ1 binds to cyclin-dependent kinase 16 (CDK16). Through this interaction, once PLCβ1 is bound to CDK16, neuronal cell proliferation slows and cells begin to terminally differentiate. The converse is also true which means that, if PLCβ1 levels are decreased, CDK16 is free to promote neuronal cell proliferation. This interaction has been connected to G-protein signaling which supports the hypothesis that cellular differentiation and proliferation is controlled, at least in part, by external stimuli (Garwain, 2018). Moreover, there is ongoing research into PLCβ1's involvement in stress granule formation (Singla, Scarlata et al, unpublished). This research looks specifically at stress granules that form as a consequence of osmotic stress. As this research is ongoing in the Scarlata lab, the data offer yet another function for the PLCβ1 enzyme. All of these potential pathways means that PLCβ1 is doing much more than simply hydrolyzing PIP_2 in order to release intracellular calcium. With all of these new potential pathways (see Figure 2), it became paramount to be able to visualize PLCβ1 at its endogenous levels and in order to do that, the CRISPR-Cas9 gene editing technique was used.

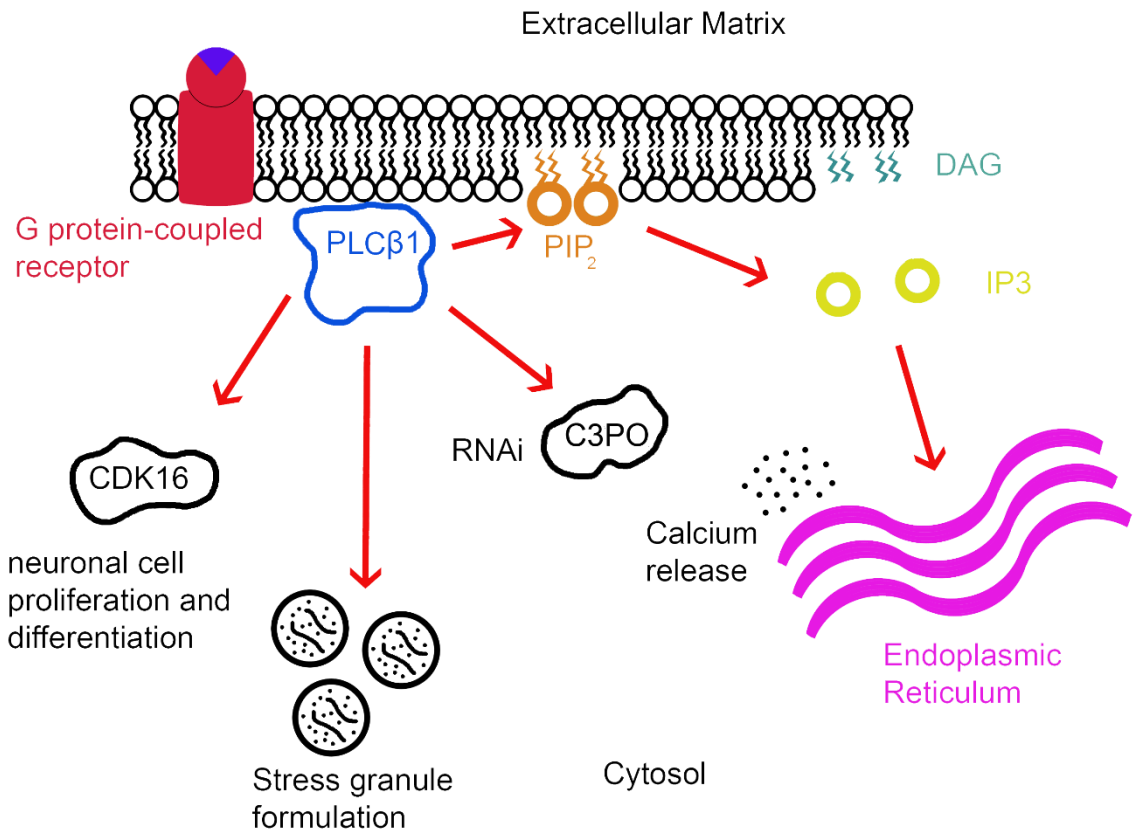


Figure 2: Suggested PLCβ1 Pathways – Implicated in more than just the phosphoinositide pathway, PLCβ1 has been shown to have a role in RNAi machinery, stress granule formation, and neuronal cell proliferation and differentiation.

CRISPR, which stands for Clustered Regularly Interspaced Short Palindromic Repeats, was first discovered as a prokaryotic immune defense against bacteriophages. When a bacteriophage injects its foreign DNA into the prokaryotic cell, *cas* enzymes and their requisite RNA scaffolds will insert part of the bacteriophage DNA into the genome of the prokaryotic cell. This insert of the bacteriophage DNA will allow for the prokaryotic cell to recognize that bacteriophage DNA again in the future and, in doing so, it confers a resistance to that particular bacteriophage (Barrangou, 2007). This ability of the prokaryotic cell to perform reactions that

allow for adaptive immunity was soon recognized for its eukaryotic genome editing potential. Edits were made to the cas9 endonuclease and its guide RNA scaffolds in order to optimize the system for eukaryotic use. These edits include the addition of two nuclear localization signals to ensure the correct localization of the endonuclease into the eukaryotic nucleus (Cong, 2013). Once the endonuclease and its guide RNAs could access the nucleus and make the specified double-stranded cut, the cell has two options with which to repair the break: non-homologous end joining (NHEJ) or homology-directed repair (HDR) (Doudna, 2014). With non-homologous end joining, the DNA is simply ligated back together which often results in either the insertion or deletion of a single base pair which leads to a frame shift. This frame shift likely renders the downstream protein non-functional. Homology-directed repair is when the cell uses nearby DNA as a template for the break repair. This is a good technique for when the insertion of a specific base pair or set of base pairs is desired. The nearby DNA can be introduced as a plasmid via transfection into the cell and, with the proper length of homology arms (the bigger the desired insert, the longer the homology arms), integrated into the eukaryotic genome (Shrivastav, 2016).

The CRISPR/Cas9 genome editing technology was perfect for the creation of an enhanced green fluorescent protein (eGFP) and PLCβ1 fusion protein. The eGFP would be situated on the N-terminus of the PLCβ1 enzyme so as not to interfere with the enzyme's catalytic activity (see Dowal et al, 2006). The guide RNAs were designed to target the transcriptional start site of the PLCβ1 gene and, due to the sheer size of the eGFP protein sequence (717 bps), 900 bps of homology to either side of the transcriptional start site were used to anchor the eGFP sequence. A diagram illustrating the proposed sequence of molecular events can be seen in Figure 3.

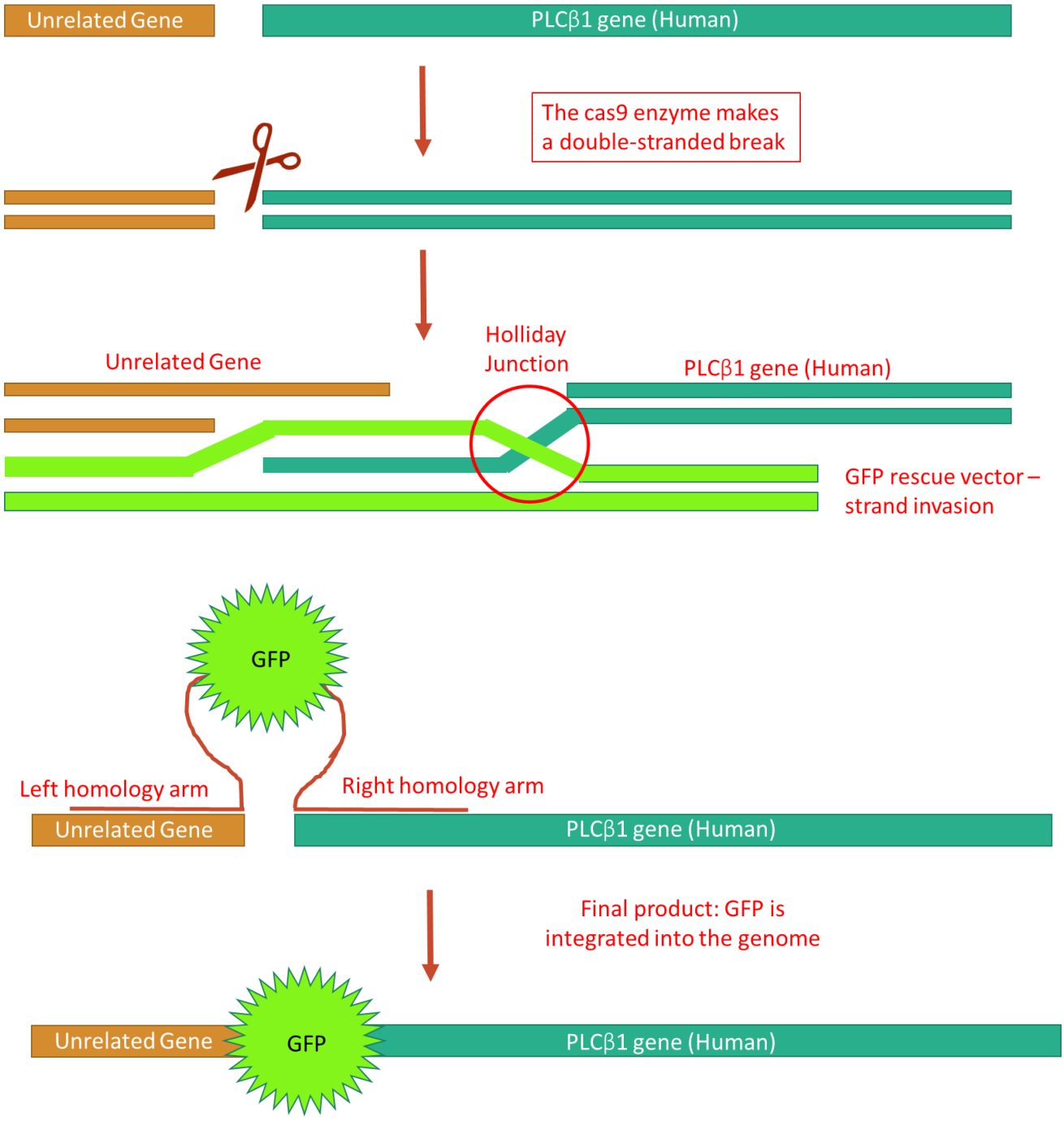


Figure 3: Illustration of the proposed sequence of events for the CRISPR-Cas9 insertion of eGFP before the PLCβ1 transcriptional start site.

Project Purpose

In an effort to visualize endogenous phospholipase Cβ1 (PLCβ1) and facilitate biophysical studies into its functions within and in addition to the phosphoinositide pathway, CRISPR-Cas9 mutagenesis was used to mutate the endogenous PLCβ1 locus. The mutation was made via co-transfection of a plasmid expressing the Cas9 endonuclease and a separate plasmid that could be utilized by the cell to undergo homology-directed repair (HDR) of the double-stranded break made by the Cas9 enzyme. The rescue vector that was used for the HDR was designed in such a way that the eGFP sequence would be inserted directly before the transcriptional start site (at the N-terminus of the protein) and would therefore be expressed along with the PLCβ1 gene under the control of the PLCβ1 promoter. Successful creation of the PLCβ1-eGFP fusion protein was visualized in cells using confocal microscopy and mutated cells were separated from non-mutated cells via chemical selection with puromycin and fluorescence-activated cell sorting in order to create a novel cell line.

Methods

Cell type. The SK-N-SH cell line is a human neuroblastoma cell line derived from a 4-year old female. Its morphology is such that there is a mixture of cells in different states of differentiation. (SK-N-SH ATCC HTB-11™) This cell line was used so that future studies in the lab could be conducted in neuronal cells that are of human origin.

Cell culture. SK-N-SH cells were seeded in Falcon cell culture dishes and maintained in GIBCO's Dulbecco Modified Eagle Medium (DMEM) that was supplemented with 10% Fobium Fetal Bovine Serum (FBS) and 5% antibiotic (penicillin/streptomycin). When the cells needed to be split into new dishes, 2mL of trypsin was added to each 100mm dish. The dish was then allowed to sit in the incubator for approximately 15 minutes while the cells detached. Sufficient cell detachment was determined via visualization under a light microscope at 10X magnification.

Cas9 plasmid construction:

Plasmid. pX330-U6-Chimeric_BB-CBh-hSpCas9 was a gift from Feng Zhang (Addgene plasmid #42230). The physical aliquot of the plasmid received by the Scarlata lab came from a stock in the McCollum lab at the University of Massachusetts Medical School. See Appendix 1-2 for plasmid maps.

Plasmid Transformation. pX330 was transformed into One Shot Stbl3 chemically competent E. coli (Invitrogen) following the protocol provided by Invitrogen. Following overnight growth of the transformed colonies, one colony was inoculated and grown up in a flask of 300mL of LB broth that was agitated at 37°C overnight.

Plasmid Maxi-Prep. Preparation of pX330 plasmid stocks was done following the Qiagen Endo-free Plasmid Maxi-Prep kit and protocols.

Design of guide RNAs. Guide RNAs were designed using an online tool developed by the Feng Zhang lab out of MIT. The tool can be accessed at the following URL: crispr.mit.edu. The target sequence selected was human and ~200 base pairs of sequence surrounding the N-terminus of the PLCβ1 gene was entered into the sequence box. A series of potential guide RNAs were generated, scored, and listed along with their accompanying PAM sequences (See Appendix 3). Three were chosen based on generated score and proximity to the start site of the PLCβ1 gene. Guide RNAs were further optimized for easy ligation into the pX330 plasmid:

5' – C A C C G (gRNA sequence) – 3'

3' C (gRNA sequence) C A A A – 5'

The following DNA oligos were ordered from Integrated DNA Technologies:

1. 5′ CACCGCGGGTTGAGCCCCGGCCAT 3′
   5′ AAACGATGGCCGGGGCTCAACCCGC 3′

2. 5′ CACCGCAGATGGCCGGGGCTCAAC 3′
   5′ AAACGGTTGAGCCCCGGCCATCTGC 3′

3. 5′ CACCGTGCACTCCGGGTTGAGCCC 3′
   5′ AAACGGGCTCAACCCGGAGTGCAC 3′

DNA Oligo Annealing & Phosphorylation. The oligos were annealed and phosphorylated by combining 1uL of each single-stranded oligo in a pair with 1uL of 10X T4 Ligation Buffer (NEB), 6.5uL ddH_2O, and 0.5uL of T4 PNK (NEB) for a total volume of 10uL. The mixture was

then annealed in a thermocycler under the following conditions: 37°C for 30 minutes followed by 95°C for 5 minutes and then decreased temperature in 5° increments until 25°C was attained.

Restriction of pX330. 1-3 ug of the pX330 plasmid were digested at 37°C for 60 minutes. The reaction mixture was as follows: 1.1-3.3uL pX330, 1uL FastDigest Bbs1 (NEB), 1uL calf intestinal phosphatase (NEB), 2uL CutSmart buffer (NEB), and ddH_2O up to a volume of 20-50uL.

Clean-up of pX330 restriction. Following the restriction reaction of pX330, the total reaction mixture underwent PCR clean-up following Qiagen's QIAquick Gel Extraction Kit protocol.

Ligation of gRNAs & pX330. The annealed guide RNAs and the pX330 plasmid were ligated together in a benchtop ligation reaction that contained the following reaction mixture: 2uL (50ng) of Bbs1 digested plasmid, 1uL of annealed oligo duplexes (1:200 dilution), 5uL of 2X Quickligation buffer (NEB), and ddH_2O to bring the volume up to 10uL. After the ddH_2O was added, 1uL of Quick Ligase (NEB) was added and the reaction was set to incubate at room temperature for 10 minutes.

Plasmid Transformation. pX330 + gRNAs was transformed into One Shot Stbl3 chemically competent E. coli (Invitrogen) following the protocol provided by Invitrogen. Following overnight growth of the transformed colonies, one colony was inoculated and grown up in a flask of 300mL of LB broth that was agitated at 37°C overnight.

Plasmid Maxi-Prep. Preparation of pX330 + gRNAs plasmid stocks was done following the Qiagen Endo-free Plasmid Maxi-Prep kit and protocols.

Rescue vector construction.

Plasmid. The eGFP rescue vector was built and designed in VectorBuilder software. The vector was designed to have 900bp of homology on either side of the transcriptional start site for PLCβ1 with the 717bp sequence for eGFP in between the two homology arms (See Appendix 4). The homology arms were mutated slightly at the PAM sequences of the gRNAs to ensure that the Cas9 endonuclease would not be able to cut the same chromosome a second time (See Appendix 5).

Plasmid Transformation. The eGFP rescue vector was transformed into One Shot Stbl3 chemically competent E. coli (Invitrogen) following the protocol provided by Invitrogen. . Following overnight growth of the transformed colonies, one colony was inoculated and grown up in a flask of 300mL of LB broth that was agitated at 37°C overnight.

Plasmid Maxi-Prep. Preparation of the eGFP rescue vector plasmid stocks was done following the Qiagen Endo-free Plasmid Maxi-Prep kit and protocols.

Mammalian Cell Mutagenesis.

Cas9/Rescue vector co-transfection. SK-N-SH and HEK-293 cells were co-transfected with the Cas9 endonuclease plasmid and the eGFP rescue vector using Lipofectamine 3000 (Thermofisher) using the suggested transfection conditions provided by Thermofisher. For a standard 35mm glass bottom dish: 250uL Opti-Mem (Gibco), 7.5uL of Lipofectamine 3000, 5uL of eGFP rescue vector, and 4uL of Cas9 endonuclease plasmid.

Chemical selection. SK-N-SH cells were chemically selected with 2ug/mL of puromycin in order to select for the cells that were actively expressing the Cas9 endonuclease as this plasmid also conferred puromycin resistance.

Visualization of GFP+ cells. Cells were visualized on the ISS Alba FCS/FLIM microscope using the GFP filter to detect GFP fluorescence in the cells following co-transfection.

FACS analysis. Fluorescence-activated cell sorting was performed at the University of Massachusetts Medical School's Flow Cytometry Core. Cells were detached and transported to UMass Medical in PBS supplemented with 5% FBS. Cells were first sorted by cell size and granularity to ensure that only cells of healthy size were sorted for fluorescence. Cells were then sorted by intensity of FITC fluorescence (as a measure of GFP intensity) against APC fluorescence (as a measure of auto-fluorescence). Cells expressing GFP were sorted into a 96-well plate in media that was supplemented with 20% FBS.

Results

The ultimate purpose of this project was the creation of a novel human cell line that would express endogenous fluorescent PLCβ1 protein. Following the restriction and ligation of the pX330 plasmid with the designed gRNAs, successful insertion of the gRNAs was determined by success of the colony growth following inoculation and through sequencing of the plasmid that the colonies took up. Following transformation, the transformation reaction was plated onto ampicillin plates and allowed to culture overnight. The following morning we visualized a lawn of bacteria indicating successful ligation (see Figure 4).

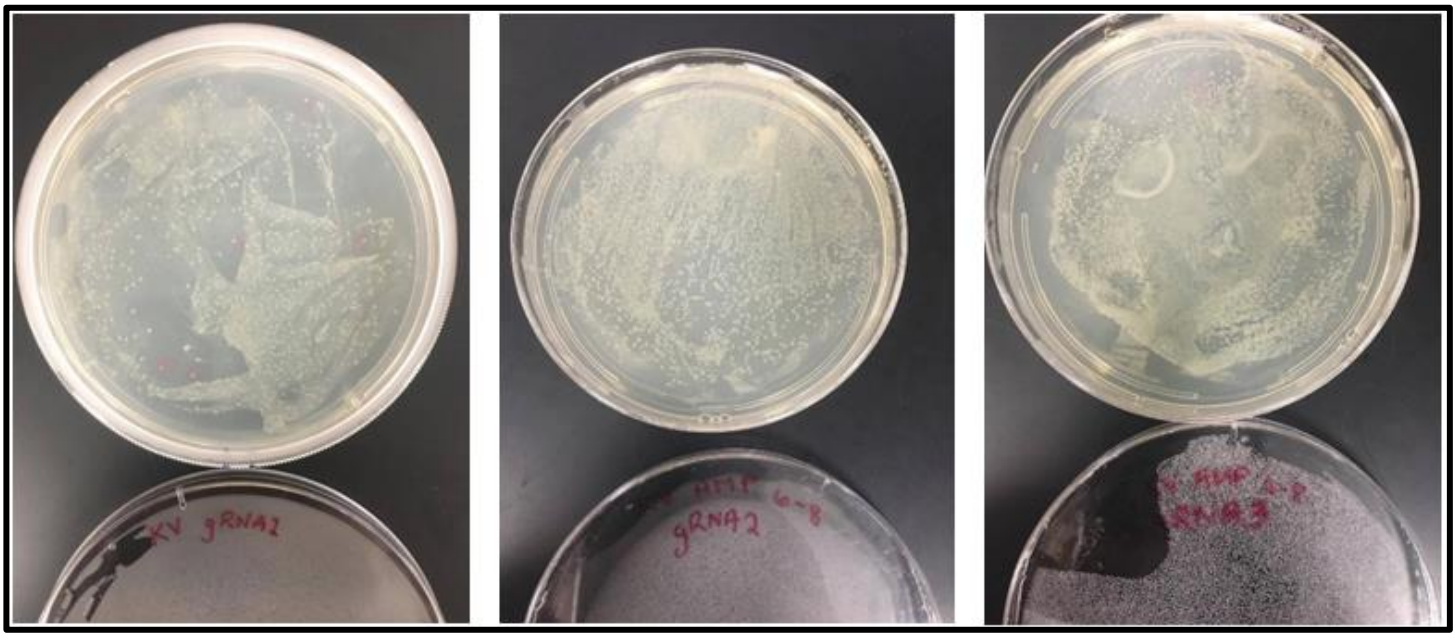


Figure 4: Growth of a Bacterial Lawn Indicating Successful Ligation – the gRNAs were successfully inserted into the pX330 plasmid and taken up by the One Shot Stbl 3 E. coli.

Once the cultures were grown, a single colony was inoculated, grown up in LB broth, maxi-prepped, and a fraction of the plasmid stock was sent to Eton Biosciences for sequencing. Plasmid sequencing revealed successful integration of two of the three guide RNAs into the

pX330 vector (see Figure 5). gRNA 3 was not available in a high enough concentration following maxi-prep.

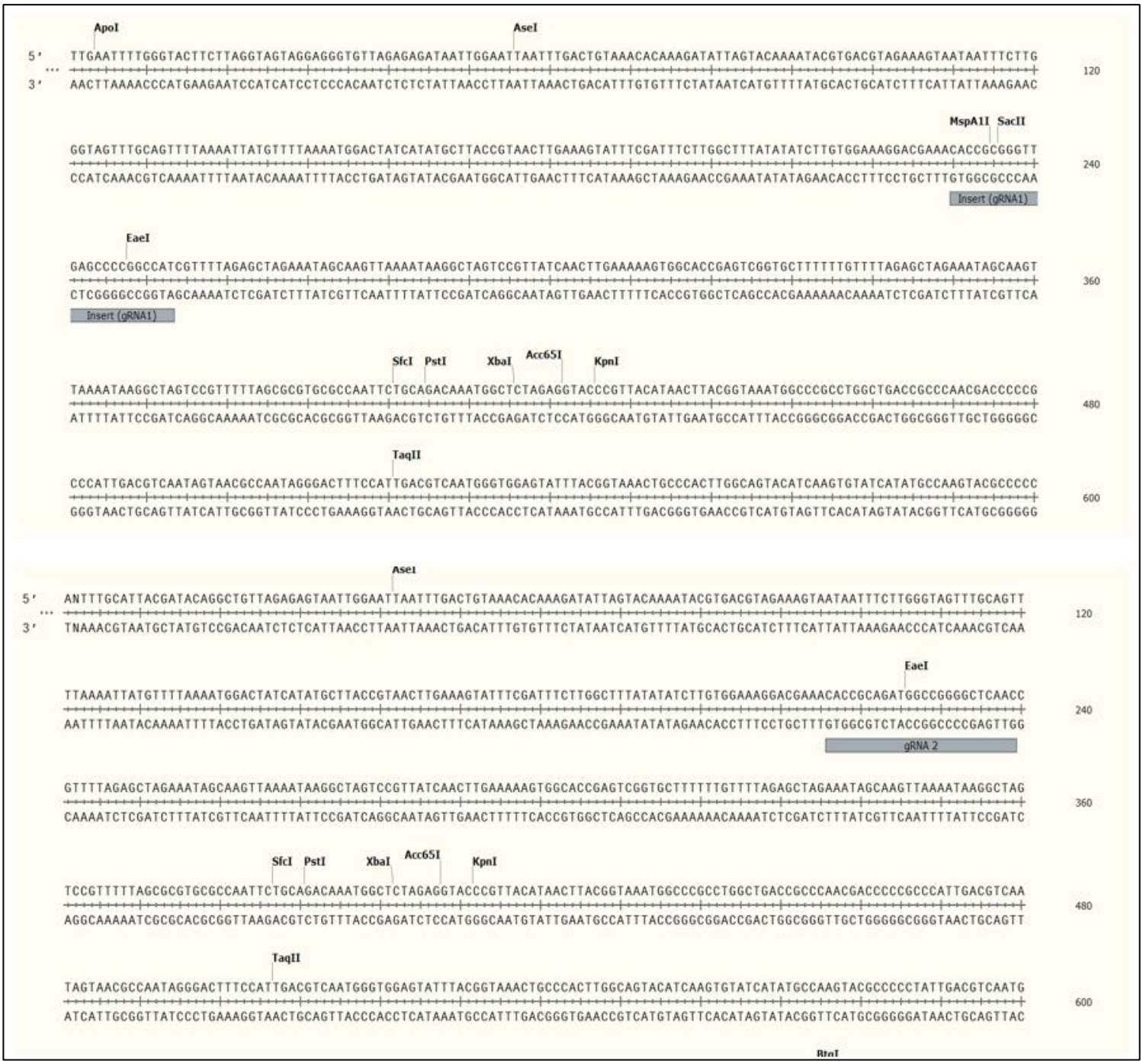


Figure 5: Sequence confirmation of the successful ligation of gRNA 1 & gRNA 2

Once the ligation of the gRNAs into the pX330 plasmid proved successful, both the pX330 plasmid with the newly introduced guides and the rescue vector were co-transfected into HEK 293 cells, HeLa cells, and SK-N-SH cells. The HeLa cells did not survive post-transfection. The HEK 293 cells grew rapidly and were brought to the University of Massachusetts Medical School to be sorted. The HEK 293 cells were brought the UMass Flow Cytometry Core to undergo fluorescence activated cell sorting. The HEK 293 cells were first sorted for cell size and granularity (SSC vs. FSC) in the non-transfected population (see Figure 6) and the co-transfected population (see Figure 7).

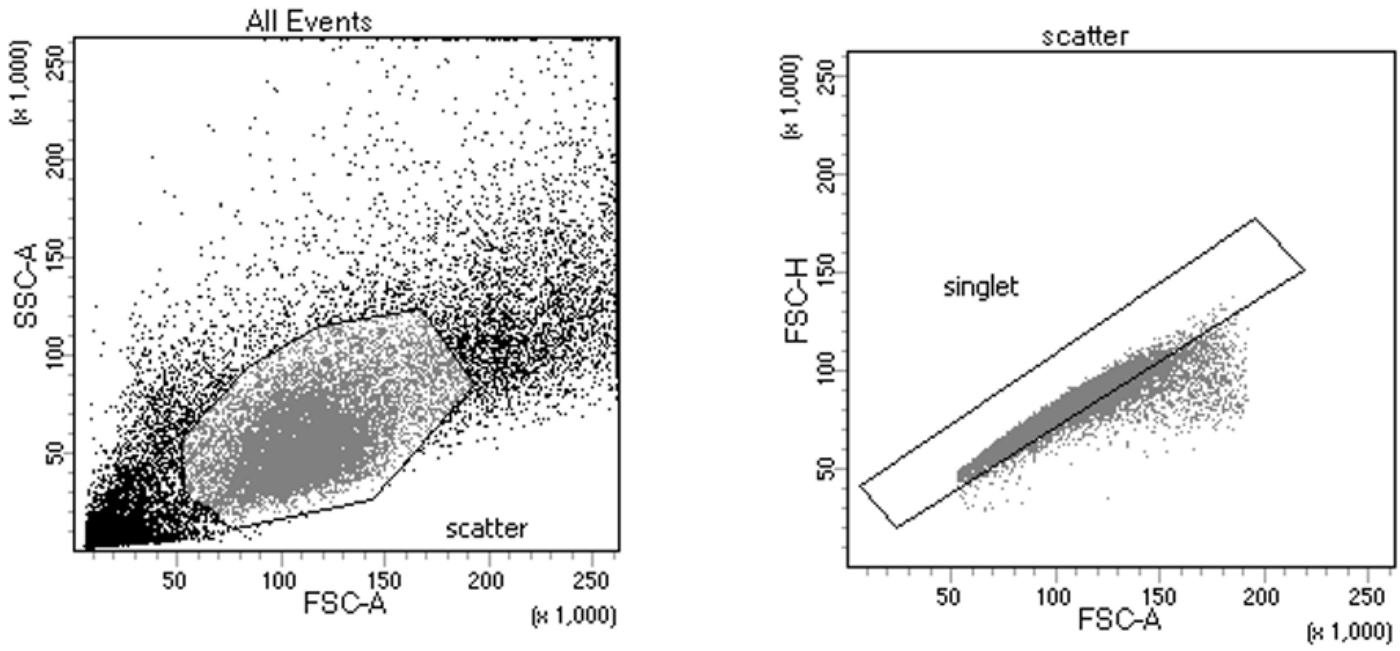


Figure 6: Non-transfected HEK cell sort for cell size and granularity

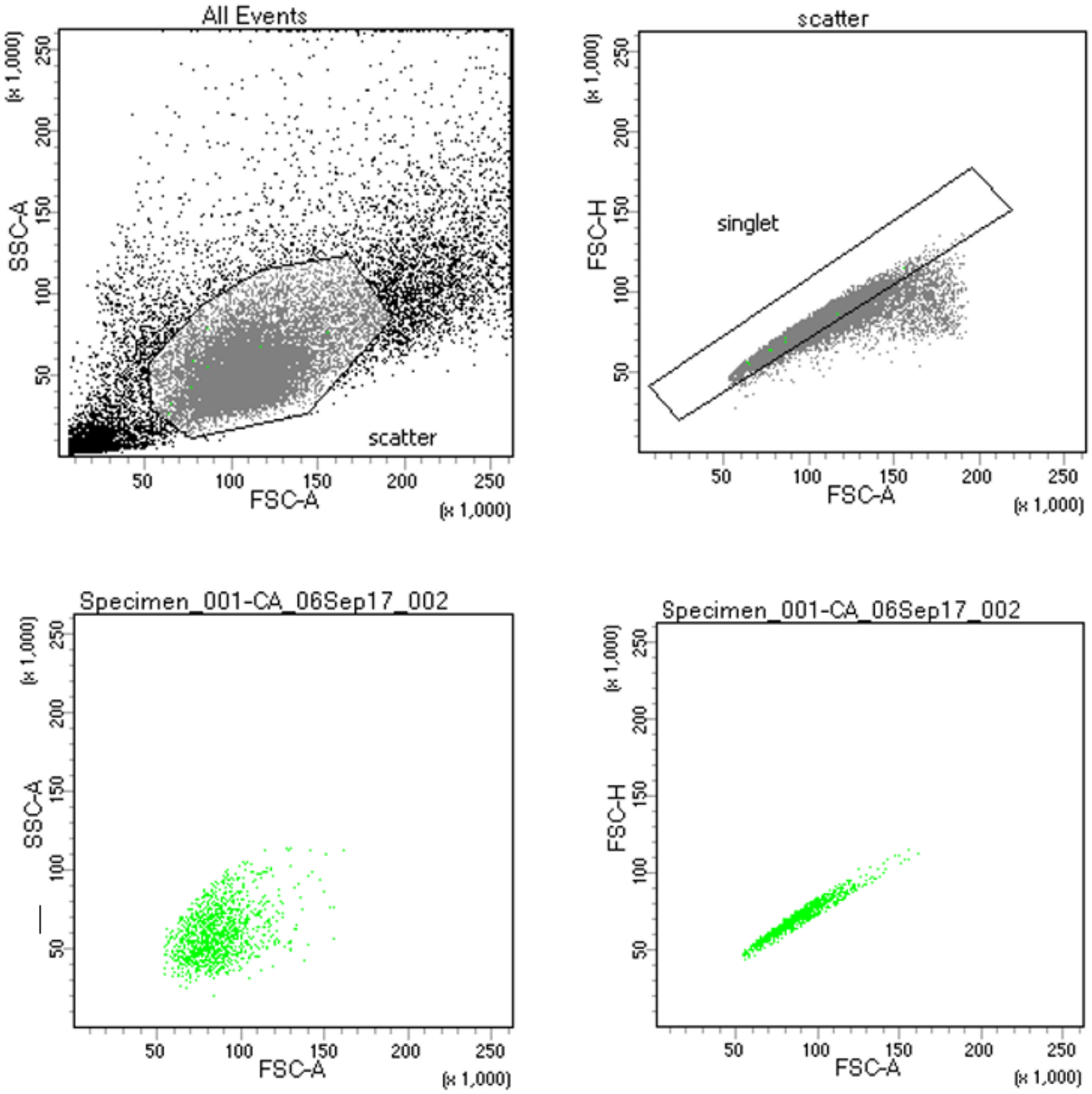
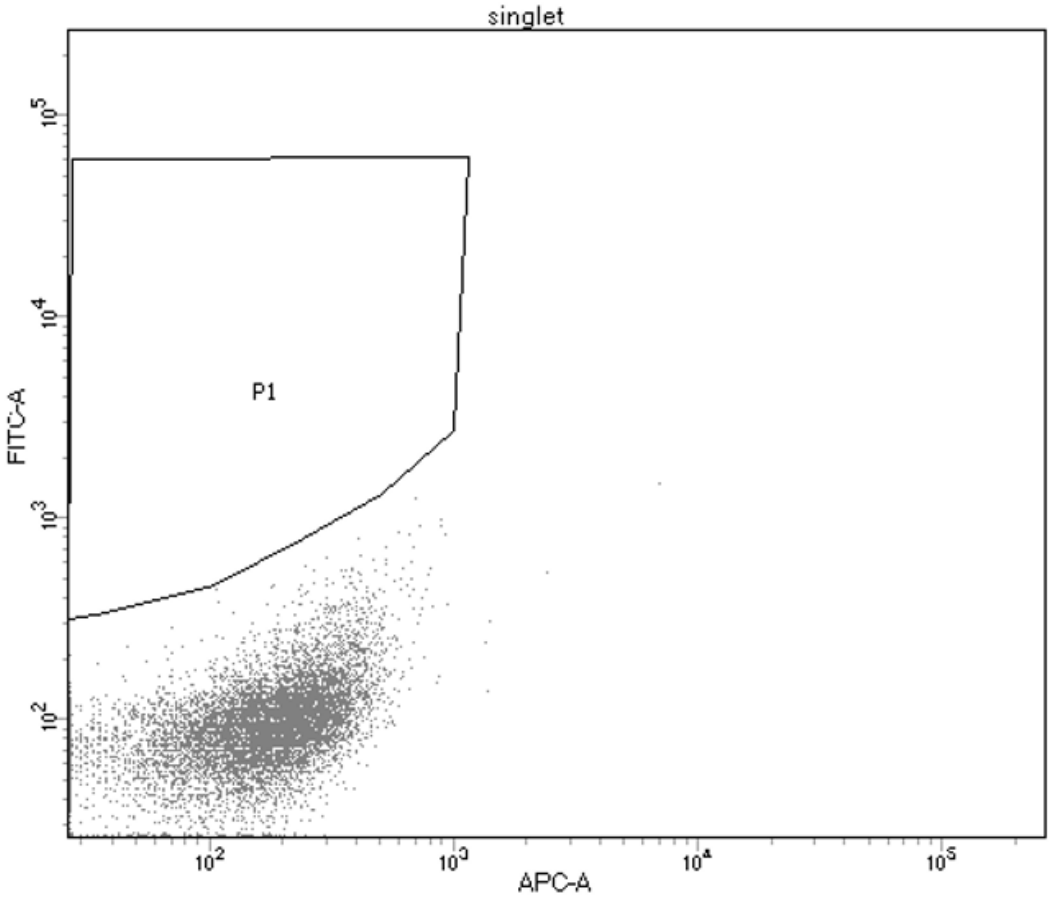


Figure 7: Co-transfected HEK cell sort for cell size and granularity

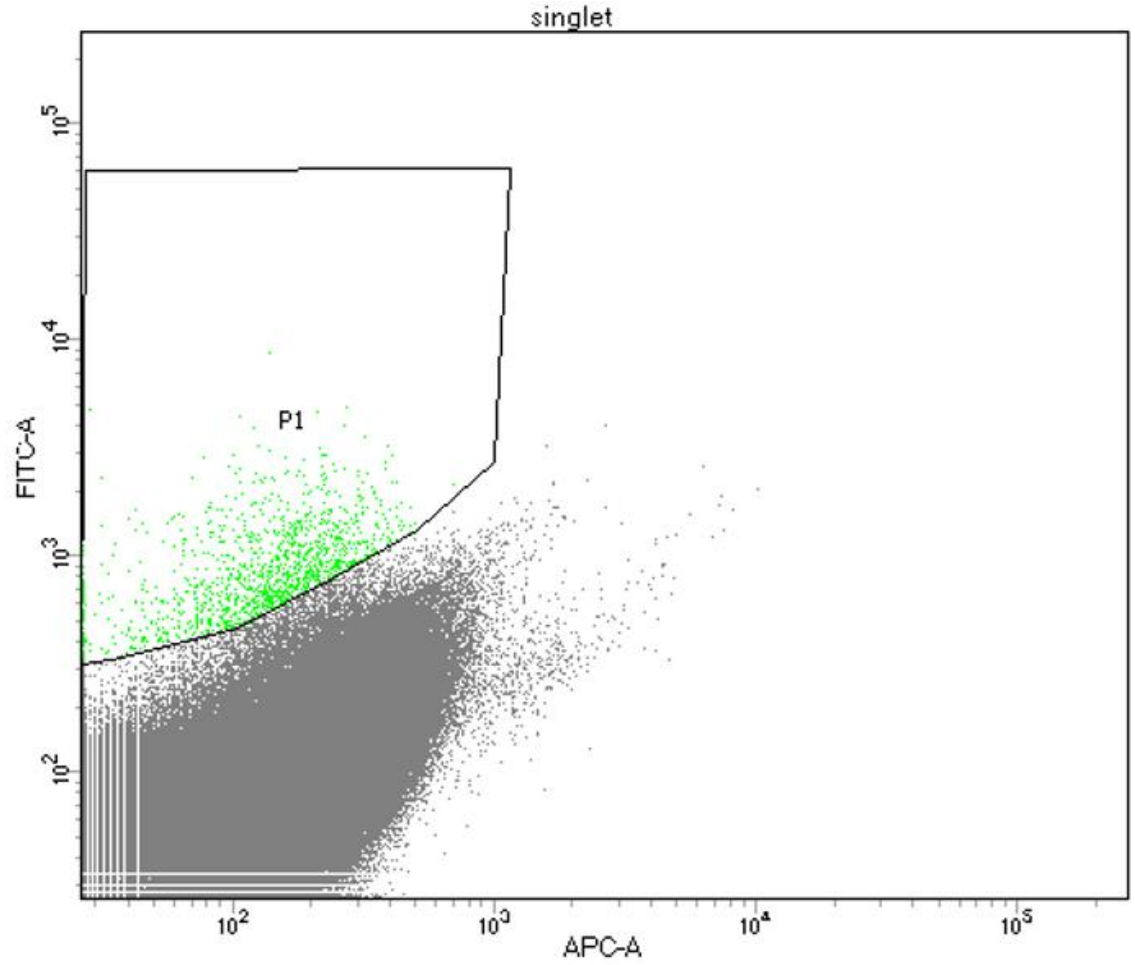
Healthy cells were sorted again, measuring for FITC fluorescence and APC fluorescence. The FITC fluorescence was used as a measure of GFP intensity and the APC fluorescence was a control that measured for auto-fluorescence in the cells. The non-transfected HEK cells, as expected, showed no fluorescence (see Figure 8). The co-transfected HEK cells showed GFP fluorescence at a very low efficiency (> 0.1% of cells were positive for GFP expression) (see Figure 9).



Tube: CA_06Sep17_001

Population	#Events	%Parent	%Total
All Events	27,720	####	100.0
scatter	12,255	44.2	44.2
singlet	9,282	75.7	33.5
P1	0	0.0	0.0

Figure 8: FITC vs. APC fluorescence of non-transfected HEK cells

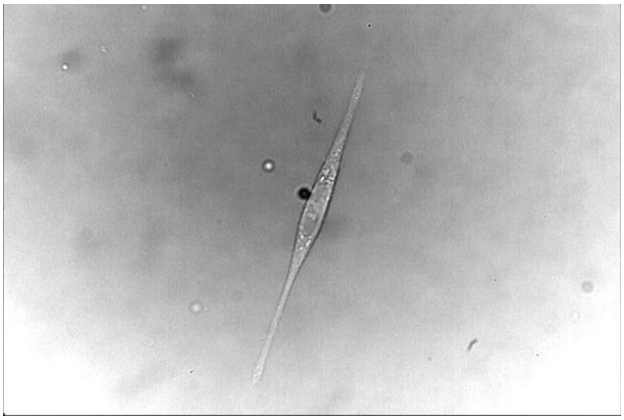


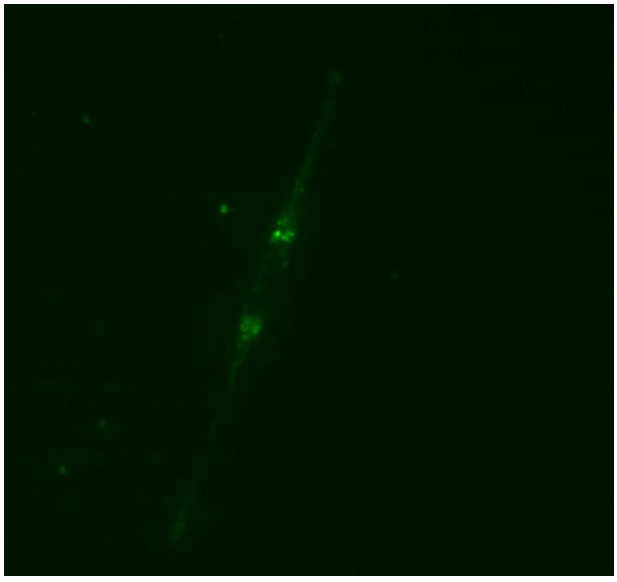
Tube: CA_06Sep17_002

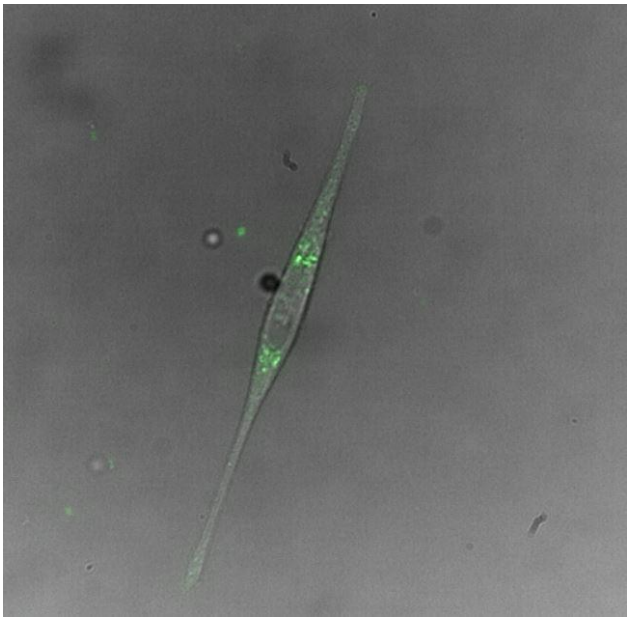
Population	#Events	%Parent	%Total
All Events	2,402,993	####	100.0
scatter	1,400,306	58.3	58.3
singlet	1,025,029	73.2	42.7
P1	1,260	0.1	0.1

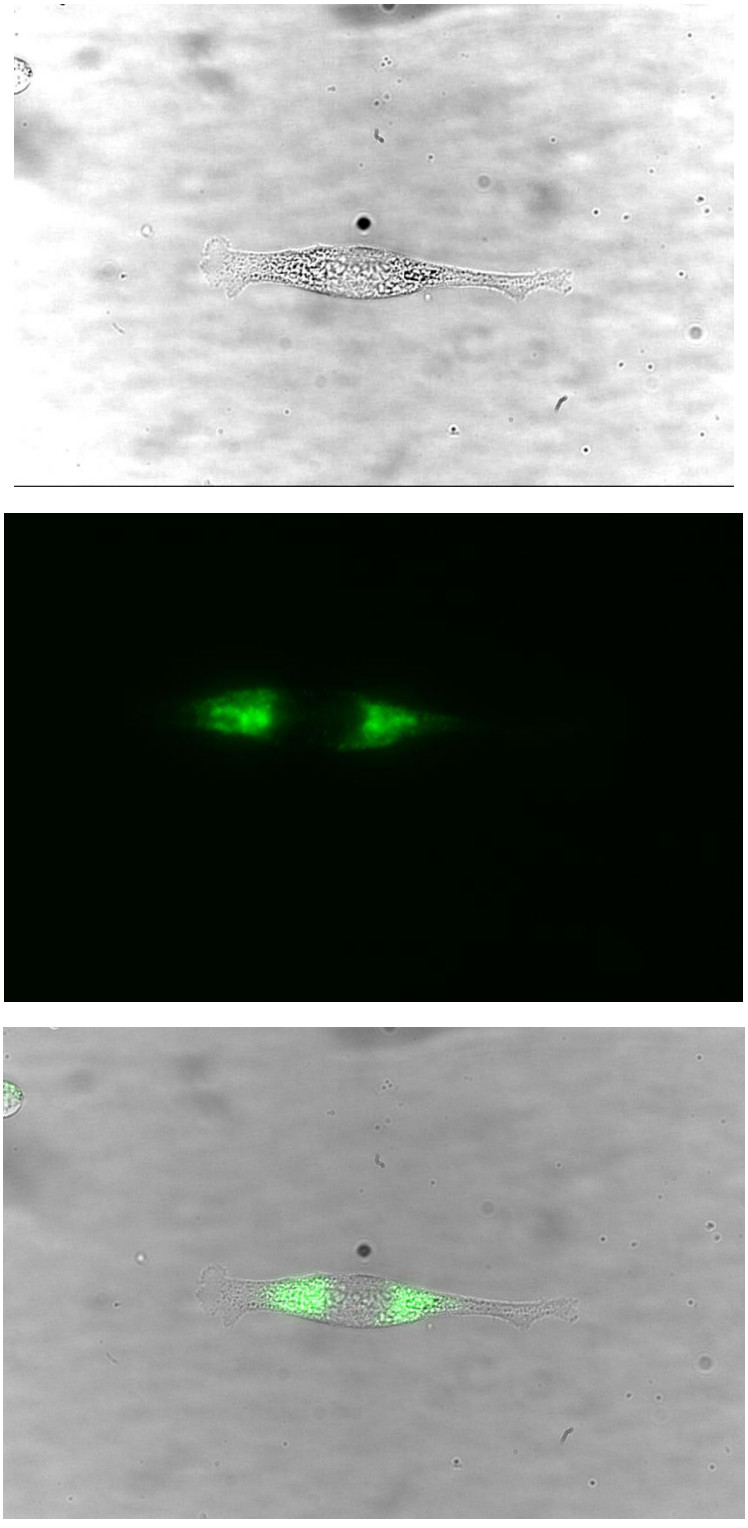
Figure 9: FITC vs. APC fluorescence of co-transfected HEK cells

The HEK 293 cells did not survive the FACS sort but the sort did verify the success of the pX330 and rescue vector co-transfection. Co-transfections were then exclusively carried out in SK-N-SH cells due to their large populations of PLCβ1. Co-transfection success of the SK-N-SH cells was determined after puromycin selection by visualization of the GFP with a confocal microscope (see Figure 10; for additonal images see Appendix 6).









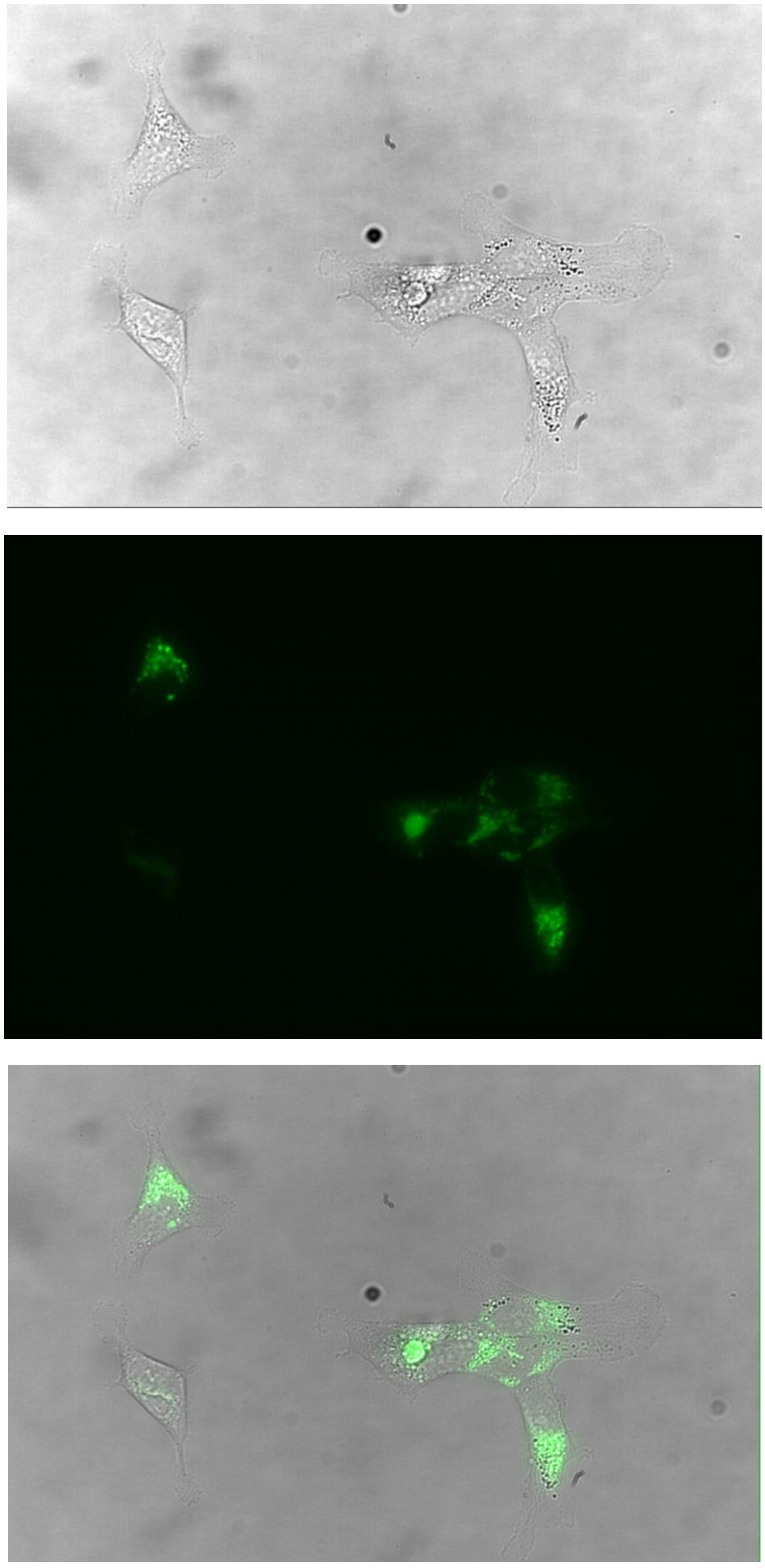


Figure 10: DIC, eGFP, and DIC/eGFP overlay images of PLCβ1 fluorescence

Discussion

In this experiment, we attempted to insert the eGFP sequence before the transcriptional start site of the PLCβ1 gene. It was decided that the eGFP sequence would be placed on the N-terminus of the PLCβ1 protein because sites for G protein activation, C3PO binding, nuclear localization and alpha-synuclein binding are mapped to the C-terminus. Past studies in the Scarlata lab indicated that the attachment of a GFP fluorophore to this enzyme would not impact its localization, activity or ability to be activated by G proteins (Dowal et al, 2006). The ultimate goal of this post-translational modification was to gain the ability to visualize endogenous PLCβ1 within live cells. With PLCβ1 having a variety of functions that exist beyond its well-known function in the phosphoinositide pathway, including a role in RNA interference and control of neuronal cell proliferation and differentiation, this ability to visualize PLCβ1 was paramount.

With the SK-N-SH cells now expressing an obvious GFP-tagged protein, it is crucial that the next phase of this experiment be the confirmation that what is fluorescing in the cells is PLCβ1. This can be done in a multitude of ways, the easiest being either western blot or reverse transcriptase polymerase chain reaction (RT-PCR). With western blotting, cells would be lysed and run on a polyacrylamide gel. The resulting protein configuration would be transferred to a membrane which would be blotted for PLCβ1. If the PLCβ1 is attached to the GFP, a band shift should be visible on the membrane showing that PLCβ1 has increased in size. We could also blot for GFP and look for signal above where the typical bands for PLCβ1 would be. With RT-PCR, a primer for the GFP sequence would be used to determine the presence of the GFP sequence as well as, theoretically, the beginning of the PLCβ1 sequence upstream of the GFP sequence. This would ensure that what is visible in the cells is, in fact, fluorescent PLCβ1.

With the CRISPR system introduced to the lab, other mutations can be made to facilitate research into the new and existing PLCβ1 pathways. By simply making edits to the gRNA sequences and targeting the Cas9 endonuclease to the catalytic site of the PLCβ1 enzyme, single stranded (or double-stranded) oligonucleotides can be introduced to engineer a point mutation in the catalytic site. This point mutation would render the PLCβ1 enzyme inactive and allow for more in-depth study into the role of PLCβ1 with respect to its phospholipid hydrolysis. There is also a lot of experimental potential within the novel SK-N-SH cell line itself as now the endogenous levels of PLCβ1 do not have to fluctuate and accurate PLCβ1 localization can be determined. This novel cell line opens up a lot of exciting possibilities for the continued study of PLCβ1 but more experiments will need to be done to verify the accuracy of this CRISPR mutation.

Bibliography

Barrangou, Rodolphe, Christophe Fremaux, Hélène Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A. Romero, and Philippe Horvath. "CRISPR provides acquired resistance against viruses in prokaryotes." Science 315, no. 5819 (2007): 1709-1712.

Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu et al. "Multiplex genome engineering using CRISPR/Cas systems." Science (2013): 1231143.

Doudna, Jennifer A., and Emmanuelle Charpentier. "The new frontier of genome engineering with CRISPR-Cas9." Science 346, no. 6213 (2014): 1258096.

Dowal, Louisa, Paxton Provitera, and Suzanne Scarlata. "Stable association between Gαq and phospholipase Cβ1 in living cells." Journal of Biological Chemistry 281, no. 33 (2006): 23999-24014.

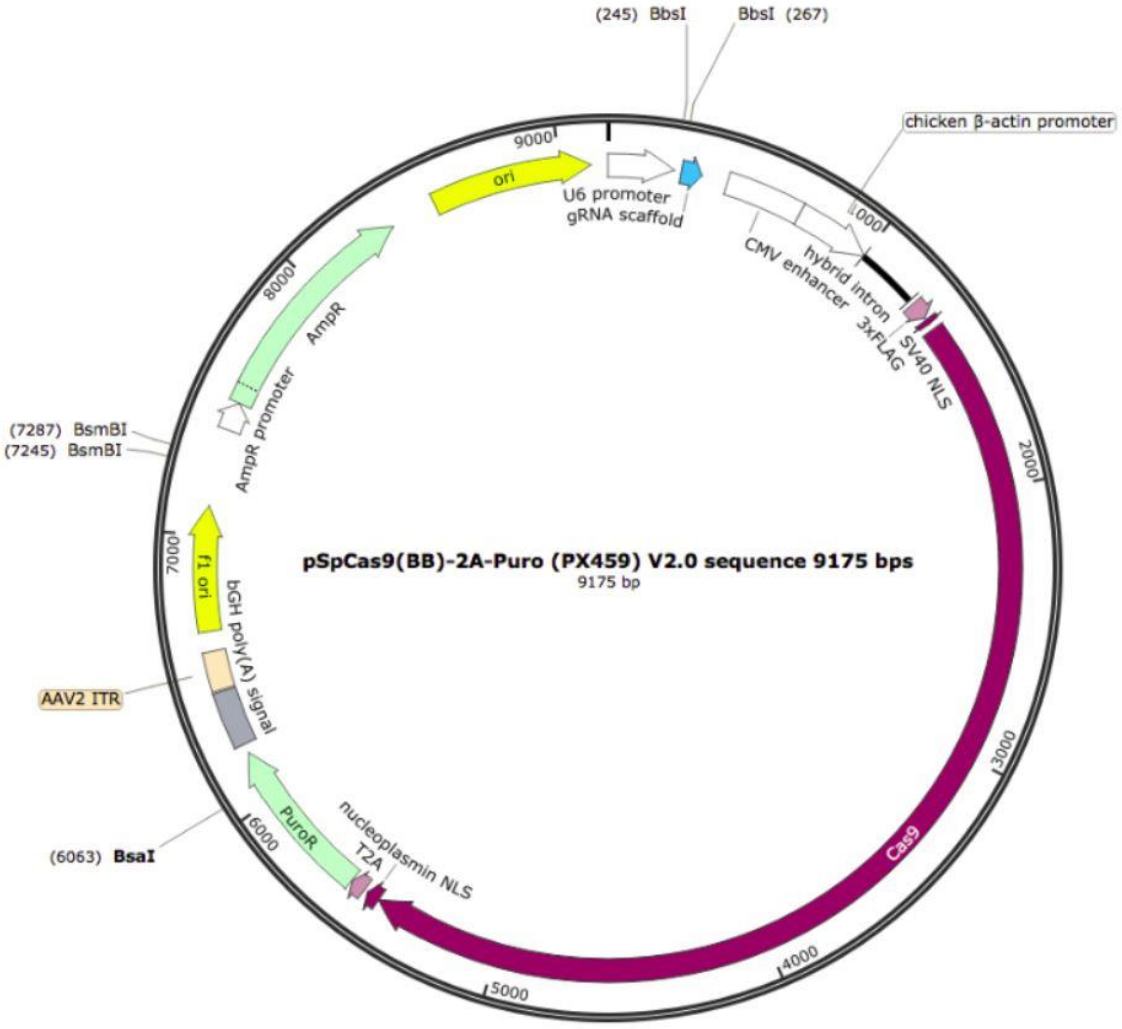
Garwain, Osama, Kaitlyn Valla, and Suzanne Scarlata. "Phospholipase Cβ1 regulates proliferation of neuronal cells." The FASEB Journal (2018): fj-201701284R.

Kadamur, G., & Ross, E. M. (2013). Mammalian phospholipase C. Annual review of physiology, 75, 127-154.

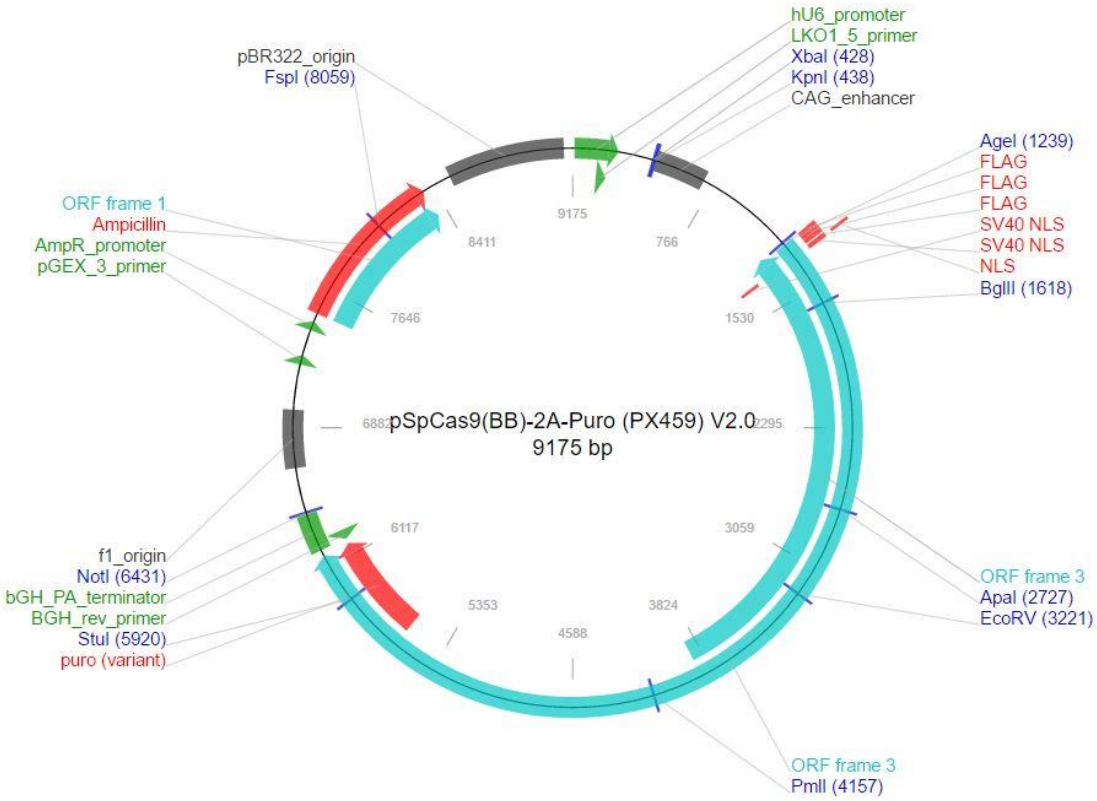
Scarlata, Suzanne, Osama Garwain, Leo Williams, Imanol Gonzalez Burguera, Barbara Rosati, Shriya Sahu, Yuanjian Guo, Finly Philip, and Urszula Golebiewska. "Phospholipase Cβ connects G protein signaling with RNA interference." Advances in biological regulation 61 (2016): 51-57.

Shrivastav, Meena, Leyma P. De Haro, and Jac A. Nickoloff. "EDITING WITH HOMOLOGY DIRECTED REPAIR (CONT)." CRISPR 101: A Desktop Resource: 49.

Appendix



Appendix 1 - Plasmid Map of PX330 (Retrieved from Addgene)



Appendix 2 - Digestion Map of Plasmid PX330 (Retrieved from Addgene)

Scores of Single Hits

The actual algorithm used to score single offtargets is:

$$\prod_{e \in \mathcal{M}} (1 - W[e]) \times \frac{1}{\left(\frac{(19-\bar{d})}{19} \times 4 + 1\right)} \times \frac{1}{n_{mm}^2}$$

Within the first term, ***e*** runs over the mismatch positions between guide and offtarget, with M:

$$M = [0, 0, 0.014, 0, 0, 0.395, 0.317, 0, 0.389, 0.079, 0.445, 0.508, 0.613, 0.851, 0.732, 0.828, 0.615, 0.804, 0.685, 0.583]$$

representing the experimentally-determined effect of mismatch position on targeting. (Hsu et al, Nature Biotechnology 2013)

And terms two and three factoring in the effect of mean pairwise distance between mismatches (*d*) and a dampening penalty for highly mismatched targets.

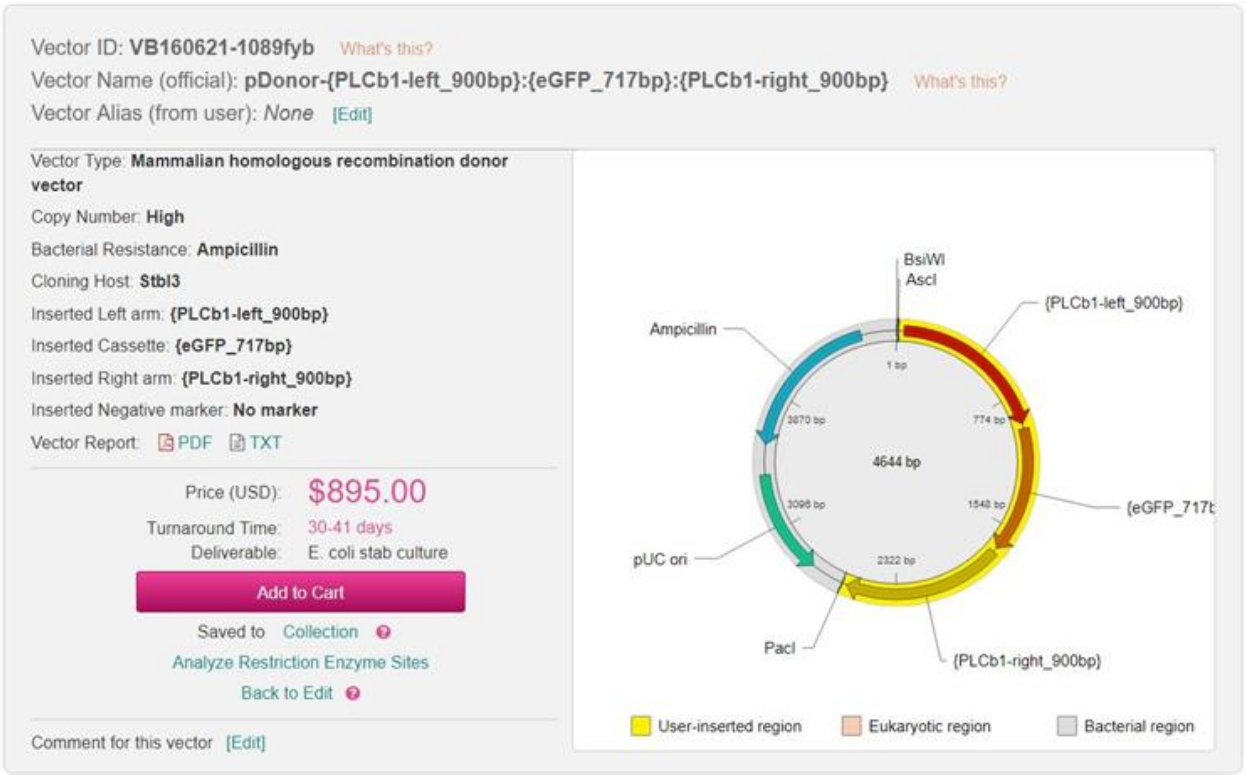
Aggregate Scores by Guide

Once individual hits have been scored, each guide is assigned a score:

$$S_{guide} = \frac{100}{100 + \sum_{i=1}^{n_{mm}} S_{hit}(h_i)}$$

and colored according to a broad categorization of guide quality which, taken into account with the presence or absence of marked genes in high-scoring offtargets indicate the relative (un)favorability of using a particular guide for specific targeting in the query region.

Appendix 3 - gRNA Scoring Formulas (Retrieved from crispr.mit.edu)



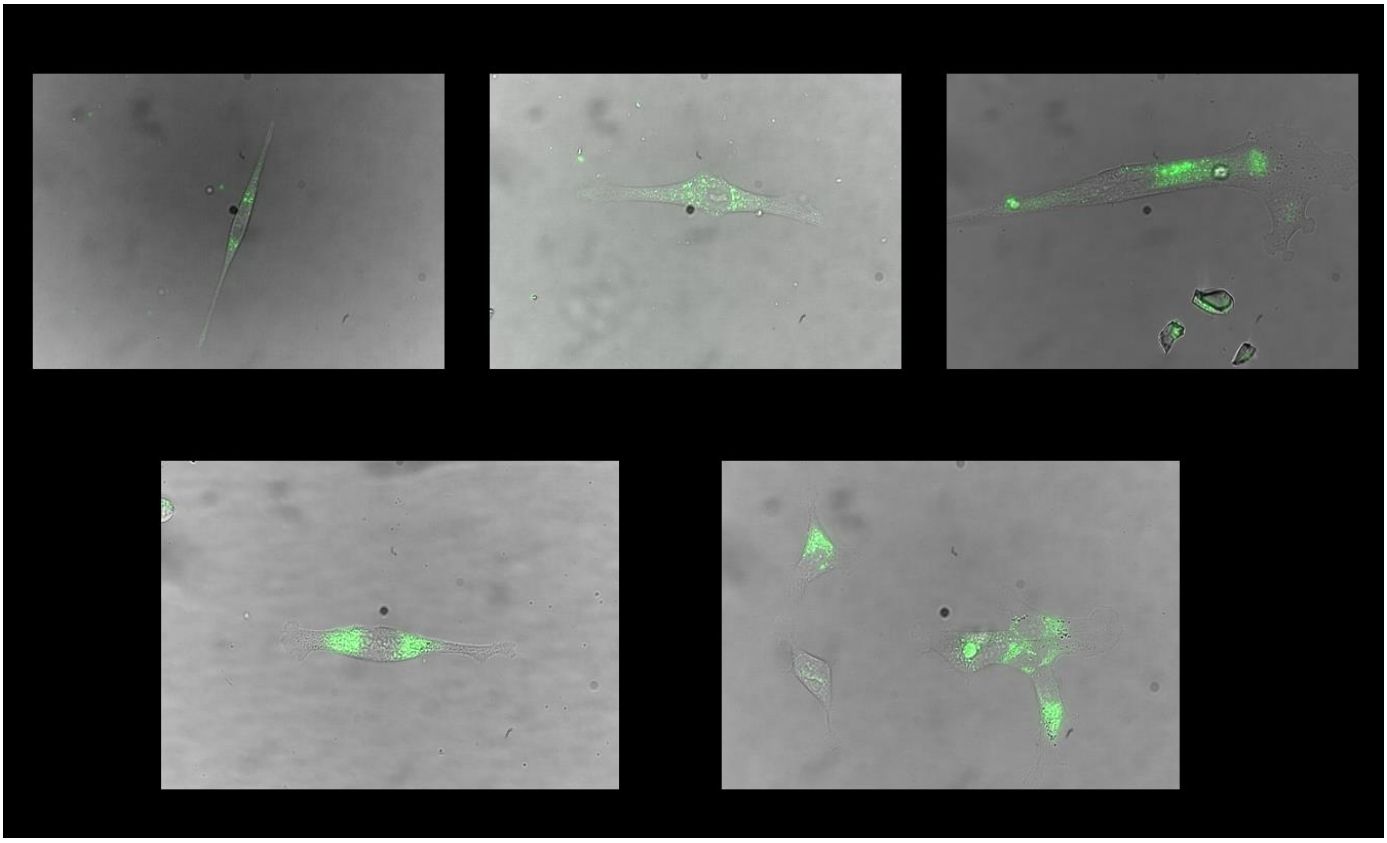
Appendix 4 - eGFP Rescue Vector Design

CAAACTCTCAGGCGCGGCCAGGTTTCCTCAAAGAGCCTCTCAACTGCGCTCTTCTCTGCTGATAAATGAT
AATTCTCTTCAAAAAGCTGGTTTGTGGCTGGCGCTGGAGAGCTCCAGCTAGAGCCAAAAGCCAGGGGTTG
CCTCTTTGGTTCCTCCCGATGGGCCCCTTTTCCCAGCCAGTTGGGATCCAAAGACGCTCCGGAGCTCGAG
CCCAAATTCCCAAGAAAACCCCAAAGAGCGTAAGGCTCCTGGGGACCCCTAGCACCTCTGCAACCAGGGC
ACGGGGCATGCTCAGAGCGAATGCGGGAGGCAGGCGTTGGCTGCAATGGGGCTCGGGAGCCAGGAGGTGC
CGCTGCGTCCCCGCTTCCTGGCGCCTGACCCAGCGTGCAAAGTTGGGCTGGGAGCCGCGGGGGGAAGGAA
GGACCCAGCGAGCGAGGGTGGCGGGGAAGGGGGGCGGGGGGGAGGAGGAGGGAGGAGGGAAAAGAGCGGT
TAGGGGGGGCGCCGGGCGTGTGTCACTCGCTCTCTCCCTCTGTGTATAGAGGATGTGCTGAATGGTGCGC
TTTGAGGCGGCGGCGGCGGAGGAGCAGAATCCGCCGCGACTGGCAGCCTCGGCTGACCGGCTCGGCTTCT
CTTCGCCTTCCGAGGCTCCTCATCCACCGCGGGCTCCAGACCTCGCGTCCCGCCCGGGGCATGGCCGGGC
GCTGCGCCCCCGCGCGCTCTGCCTGCTGAGCGGCGCCGGAGGGAGGTGCGGAGGCCGGGAGGCCGGGGAG
GCCGGCGGGGAGCAGAGTCGAGCGCCTCCGGAGCAGAGAAAGGAGCCCGCGCCCCGCGCCCCGCGCCCCG
CGCACGGTCCCCAGTCCCTGCCGCGCTCGCCCGGGCCGCCCGGAGCCCAGATGAGCTCAG – Left
arm

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG
TTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTG
CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGC
TACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTC
GAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC
ATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCC
CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATC
ACTCTCGGCATGGACGAGCTGTACAAG

ATGGCTGGGGCTCAACCCGCAGTGCACGCCTTGCAACTCAAGCCCGTGTGCGTGTCCGACAGCCTCAAGA
AGGGCACCAAATTCGTCAAGTGGGATGATGTAAGTATTGGGGCGGCCCGAGTCGGGGCGCTGGCTCGGGC
ACCGGGCAGGGCGGGCGTCGTGGGGGTGGGGCAAGGGGCGCGTTATGCAATGGGCGCACTGGGAGCGGGC
AGGGGCAGCCTCGGGCGCACAGGTTGGCATCTGCCAAAGCGGATGTCCAAGGGCAGAAGCTTTGCGCGCG
CTCCTGTTTCATCGGGCTTCAGTAGTTGCCATCCTTTCTGGGTCTGGCAGGCGCCCCCTGAGGGTCTTAC
ATCTTTTGGGCCAGTGGAGAGGGAGGGGCAAGCTTTGCCTGGTGACAGGTCCTGAGAGTTCGGAAAGTAG
GGAGTGTGAAGTTGGGACCAACTGGAAAGCGCCTGGTGTGAACAAATCGCGTCCCTCTCCGCCTGCGGAC
GCACCCCCTATTACCTCTTTTGCCCGATTTCAAGAAGGAAGGACCTGAAGTTCAAAACACTTTAGTGATT
GCGAACCCTCTTGGCGCTTTTGTGGGAGCCAAGGCGGCGGAGGACACAACGGTCCAGCCAACAGGTGGTC
GTTGATGCCAAACTTCAGCGCCTGGCTCCGGAGTGAGGGTCCCTGAAAAGGGCTGCGGGGACACATCTGG
GGAGGGCGTGAACAGAATGGGGGCCCTTGAGGGGTCCTGGAGGGCTCATGGGCCTGCAGGCAAGTGGGGA
GGGCTGACTTTTGAGAACTGAGCTGATTGGAGTCTTCTGCCTGAAATAGACAGGTTGAGAGTTTGCCCCA
CGATTTACCTGCACCGTTTGCCAGGTAGGGCAGAGCCTAGCCCAGCGCCGGGGTCTCTAA – Right
arm

Appendix 5 – eGFP Rescue Vector Sequence (edited PAM sequences highlighted in green)



Appendix 6 – Composite Images of eGFP-PLCδ1 in SK-N-SH Cells