ENGINEERING ANTIBODY THERAPEUTICS AND CELL CULTURE MODELS FOR NEURODEGENERATIVE DISEASES

by

Kyle P. McHugh

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering

Fall 2016

ENGINEERING ANTIBODY THERAPEUTICS AND CELL CULTURE MODELS FOR NEURODEGENERATIVE DISEASES

by

Kyle P. McHugh

Approved: _____
Abraham M. Lenhoff, Ph.D.
Chair of the Department of Chemical and Biomolecular Engineering

Approved: _____
Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

David W. Colby, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Wilfred Chen, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

April M. Kloxin, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Abhyudai Singh, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I have many people to thank for helping me get to this point, but I would especially like to thank my mother and grandparents for their guidance and helping me become the person I am today. I also want to thank my entire family enduring our conversations about science and always supporting everything I do. Their love and support have been integral to both starting and completing my graduate work. I would also like to thank all of my friends including those from before my time at the University of Delaware as well as all of those that have suffered with me through this whole process for all the great memories and fun times we've had. Their companionship has not only helped relieve some of the stress of graduate school but has made my time here an enjoyable and significant time in my life.

For the research I have been fortunate enough to perform and my growth as a scientist, I would like to thank my advisor Professor David Colby for his guidance and mentorship throughout all my thesis work. I would also like to thank my committee professors Wilfred Chen, April Kloxin, and Abhyudai Singh for their helpful conversations and constructive feedback. I would especially like to thank my fellow lab members including Olga Morozova, Abhinav Jain, Shy'Ann Jie, Ming Dong, Kyle Doolan, Stefanie Berges, and Elisa Ovadia for their support, commiseration, troubleshooting, teamwork, and friendship. For their help with my research and allowing me the opportunity to mentor fellow students, I would like to thank all of the undergraduate students and visiting students who have worked with me during my time at UD.

A lot of people have helped my research in varying capacities to varying extents and so I would like to thank Kathie Young in our department as well as the entire staff in Colburn Lab for handling so many of the administrative details that our department would not be able to run without. I would like to thank other professors including Terry Papoutsakis, Wilfred Chen, Bramie Lenhoff, and Millicent Sullivan for allowing me to use their lab space and equipment. Karen Black has made ordering and restocking our lab so much smoother and has provided numerous pleasant conversations for a welcome break from intense lab work. Many people have contributed to my success and shared in my failures over the years and I would not have been able to reach this point in my life without them. I dedicate this work to my family, all of the people mentioned here, and the dozens and dozens of people I don't have the space to thank.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Neurodegenerative diseases are typically characterized by selective neuronal subtype vulnerability and associated incorrectly processed and misfolded amyloidogenic proteins. These disease pathologies differentially affect anatomically specialized areas made up of dozens of different subtypes of neurons interconnected in complicated pathways. Currently, there are no established treatments to halt or even slow the progression of these diseases. To deconvolute these interacting effectors of neuronal dysfunction, we can develop neurodegeneration cell culture models to break up different components into intrinsic and extrinsic effects. Intrinsic effects consist of age-related dysfunction and neuron subtype identity and how the type of neuron mediates selective vulnerability. Extrinsic effects include neuro-inflammation factors, apoptotic factors released from nearby dying cells, and misfolded disease proteins such as tau protein as they travel through the brain. By understanding some of these effects in isolation we can identify points of therapeutic intervention and quantify disease biomarkers to track disease progression and therapeutic efficacy. In this work, we develop a high-throughput platform for neuronal reprogramming to culture specific subtypes of patient-derived neurons for studying cell type-specific intrinsic properties, we establish an improved culture model of tau protein disease strain fibrillization to study extrinsically-added misfolded protein, and we demonstrate the feasibility of engineering therapeutic proteins in a heterologous host for improved stability and expression.

Huntington's disease (HD) is a dominantly inherited, progressive neurodegenerative disease characterized by specific and extensive loss of medium spiny neurons during striatal degeneration. It remains unclear whether this cell-type selectivity is primarily due to intrinsic cell-autonomous effects or neuronal dysfunction in surrounding neurons. The ability to obtain relevant cell-types *in vitro* would be an invaluable resource to study these cell-type specific effects. We assembled and screened a library of transcription factors to reprogram human fibroblasts or induced pluripotent stem cells from HD and non-HD patients into neuronal subtypes important in HD. We also developed a combinatorial computational model to assist in the selection of experimental parameters to optimize library screening efficiency. The resulting identified sets of transcription factors may be used to develop a fast, reproducible *in vitro* model that accurately captures the major phenotypic characteristics of HD in order to further understand the intrinsic cell-type specific pathogenesis of the disease.

The microtubule-associated protein tau stabilizes microtubules in neurons for cytoskeletal support and cellular trafficking. Fibrillization and accumulation of tau is a defining characteristic of a group of neurodegenerative diseases called "tauopathies" in which tau forms distinct fibril strain conformations across different tauopathies which may contribute to their unique histological and clinical characteristics. Most current tauopathy cell models rely on the application of synthetic tau fibrils formed by induced fibrillization with the poly-anion heparin. While it has been previously demonstrated that heparin-induced fibrils are structurally distinct from brain-derived tauopathy fibrils, we have shown that synthetic fibrils are also significantly different when propagated in a cell culture model of extrinsically-applied tau strain

propagation. Additionally, while truncated or mutant tau has been useful to study the mechanics of synthetic fibril propagation, we have shown that full length wild type tau is better suited for propagation of disease-associated strains in this model. The cell culture model presented in this study may be broadly applicable to many other neurodegenerative diseases.

The protein-only hypothesis for prion-like diseases asserts that neurodegenerative disease-associated proteins such as the prion protein or tau protein can spread their misfolded disease conformation to natively folded protein by direct intermolecular interaction to template the misfolded structure. Understanding this interaction suggests a therapeutic approach in which a specifically-binding therapeutic may be able to disrupt the interaction of native and misfolded protein to halt the progression of the disease. We chose an anti-prion antibody ICSM18 that was previously engineered by our group to better understand the feasibility and fidelity of engineering therapeutic proteins in a heterologous host. CHO cells are the most common platform for protein secretion in the pharmaceutical industry, but *Saccharomyces cerevisiae* are an attractive engineering platform because of their eukaryotic expression machinery and rapid doubling time. We found that apparent affinity by yeast display may be significantly lowered by differences in expression machinery between the two expression systems but that improvements in stability in yeast translate directly to improved expression yields in CHO cells which indicates that yeast display may not provide exact affinity estimation but that complex protein therapeutics may still be efficiently engineered in a heterologous host.

Chapter 1

INTRODUCTION

1.1 Motivation and Goals

Neurodegenerative diseases such as Alzheimer's disease (AD) affect over 47 million people with over 7 million new cases each year according to the World Health Association. The healthcare burden of Alzheimer's disease and other dementias is over $600 billion as of 2010 and is expected to increase over time. Neurodegenerative diseases are typically characterized by the misfolding and amyloid-formation of disease-related proteins, such as beta-amyloid and tau protein in AD, which is associated with neuronal dysfunction and loss resulting in cognitive decline and eventually death [1]. There are currently no known therapies to halt or reverse the progression of these disease and so the development of disease models is crucial to better understand disease pathogenesis and identify therapeutic targets.

The brain is a complex organ comprised of dozens of distinct neuronal subtypes involved in particular neuronal pathways in different anatomically distinct regions of the brain. Different neurodegenerative diseases spread through the brain along varying synaptic pathways with differing cell type-specific vulnerability and so it is difficult to model these diseases in culture. Disease pathogenesis and neuronal dysfunction arises from a complex interplay between cell-type specific effects including subtype identity and age-related changes with effects arising from outside the cell such as misfolded protein, neuro-inflammation, and pro-apoptotic factors from nearby dying cells. It is therefore useful to study these effects in isolation in disease

models *in vitro* to better understand their specific effects to better identify therapeutic targets and disease biomarkers (**Figure 1.1**).

In the first part of this work, we establish a high-throughput platform of neuronal reprogramming to covert accessible human somatic tissue samples such as fibroblasts directly into specific subtypes of neurons to be able to study subtype-specific vulnerability in different neurodegenerative diseases. We initially focus on the subtype of neurons most and earliest affected in Huntington's disease (HD), medium spiny neurons (MSNs), to establish the platform. In the second body of work, we derived an equation for cooperative library screening and optimized experimental parameter selection to screening reprogramming transcription factor libraries as efficiently as possible.

In the next body of work, we look at the misfolding and fibrillization of disease-associates tau protein strains. First we survey the field to summarize the current state of the field and identify the major accomplishments. We also look at how our lab has made significant contributions to the field and areas in which we could further contribute. Then we discuss the development of an improved cell culture model of tauopathies to study the propagation of disease-associated tau strains in cell culture. Our improved model may yield more predictive results to better recapitulate the propagation of disease fibrils in tauopathy patients to screen potential therapeutics to prevent the fibrillization of tau strains.

In the last section of this work, we discuss the engineering of a protein therapeutic for prion disease. Prions propagate in the brain by direct interaction of the disease conformation of the protein with natively folded protein to induce a conformational change into the disease form. By blocking this interaction with an

antibody capable of specifically binding native prion protein, we can prevent the spread of the disease in the brain. We started with ICSM18 single-chain variable fragment (scFv) antibody mutants generated via directed evolution by a previous lab member to investigate mutational effects to better understand the differences in yeast and mammalian expression systems as well as the effect of stability improvement on mammalian expression.

Overall, towards studying cell type-specific disease effects we have established the groundwork for a high-throughput neuronal reprogramming platform, isolated and subcloned the largest current neuronal reprogramming factor library of over forty transcription factors, and developed a combinatorial model of library screening to guide selection of experimental parameters. Towards understanding the spread of misfolded disease-associated proteins in the brain, we have generated an improved model of tau protein fibrillization in culture to better recapitulate the conformational strains that propagate in various tauopathies. To better understand the development of protein therapeutics, we investigated the use of protein engineering in a heterologous host for stability and expression improvements in mammalian cells. The models and approaches developed in this body of work may help to better study and understand neurodegenerative disease to screen therapeutic candidates able to slow or halt the progression of neurodegenerative diseases in patients.

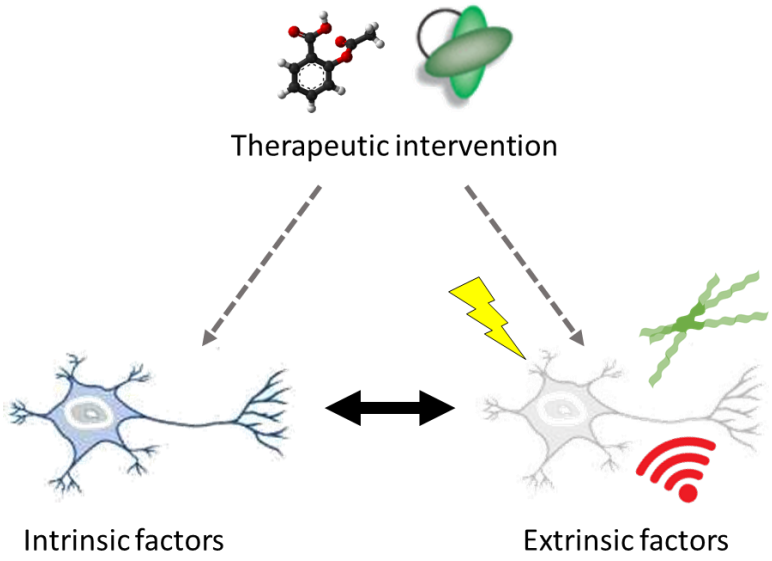


Figure 1.1 Deconvolution of the complex interplay of intrinsic and extrinsic factors in neurodegenerative disease models to identify points of therapeutic intervention and disease biomarkers.

1.2 Modeling Neurodegenerative Diseases in Culture

1.2.1 Intrinsic Neuronal Dysfunction

In neurodegeneration, intrinsic effects are the cell autonomous effects which arise from within the cell experiencing dysfunction. One of the aspects of intrinsic effects are stress and age-related changes. In the normal course of ageing, neurons experience several common types of dysfunction which make them more vulnerable to neurodegeneration such as calcium dysregulation, proteolytic dysregulation, and energy metabolism dysregulation [2-7]. These changes vary depending on location in the brain and type of neuron which accounts for the stressor-threshold hypothesis of neurodegeneration which suggests that neurodegeneration spreads as neurons become vulnerable enough due to age-related changes to become unable to handle the load of

dysfunction from surrounding neurons and misfolded protein [8]. The other major aspect of intrinsic neurodegeneration effects is which particular subtype a neuron is and how its particular properties make it more or less vulnerable in different neurodegenerative diseases. Different types of neurons may extend their neurites into more strongly-affected pathways or rely upon different neurotransmitters and receptors which may make them more susceptible to excitotoxicity. Different subtypes of neurons therefore react differently to different stimuli which means that a disease model of neurodegeneration should include neurons most affected in the pathogenesis of the disease as well as lesser-affected neuronal subtypes to test disease-specific dysregulation of cellular properties and screen for therapeutic intervention to correct these irregularities.

Obtaining primary human neurons is difficult because of the prohibitively invasive nature of biopsies and the long and complicated processes required to differentiate stem cells into particular subtypes of neurons. Cellular reprogramming provides a faster more efficient alternative which is able to transdifferentiate one particular differentiated cell type directly into a specific type of neuron without a pluripotent intermediate by forced expression of transcription factors. Since different subtypes of neurons have different levels of vulnerability in various neurodegenerative diseases, cellular reprogramming would be useful to establish reproducible cell culture models of neurodegenerative disease. For example, in Huntington's disease (HD) which is a genetic disease caused by an expansion of the poly-CAG tract in the huntingtin gene, medium spiny neurons (MSNs) in the caudate and putamen of the basal ganglia are most adversely affected throughout the early and late stages of the disease [9]. In Alzheimer's disease (AD), basal forebrain cholinergic neurons (BFCN)

of the hippocampus are most affected in the early stages of the disease before neurodegeneration begins to spread to other regions of the cortex [10, 11]. A high-throughput approach to neuronal reprogramming would help screen for a combination of transcription factors sufficient to transdifferentiate patient tissue such as fibroblasts directly into these neuronal subtypes to study HD and AD in culture and screen potential therapeutics.

1.2.2 Extrinsic Effects in Neurodegeneration

Extrinsic effects are another important group of cellular dysfunction-inducing phenomena which are independent of the affected cells and arise from a variety of sources. Neuro-inflammation caused by activation of the immune system by protein plaques for example causes the local release of pro-inflammatory factors which places a stress upon local neurons [12]. Local neurons in primary or secondary stages of apoptosis release pro-apoptotic factors which also induce cellular dysfunction in local surrounding neurons. One of the more salient characteristic features of neurodegenerative diseases is the presence and accumulation of improperly-processed amyloidogenic proteins associated with various neurodegenerative diseases. For example, histological analysis of brain tissue from AD patients at progressive stages of the disease first begin to show tau protein plaques closely tied to neuronal loss and later the accumulation of extracellular amyloid-beta plaques [11, 13]. Other diseases are associated with different disease-related proteins such as Parkinson's disease with alpha-synuclein, Huntington's disease with the huntingtin protein, and prion disease like Creutzfeldt-Jakob disease with the prion protein. These extrinsic effects place a burden upon the neurons which induces neuronal dysfunction that propagates across

the brain along local regions and distal synaptically connected pathways throughout disease pathogenesis.

A particular extrinsic effect of interest in the misfolding and conformational propagation of disease-associated proteins in neurodegenerative disease. In a group of diseases commonly referred to as tauopathies, the tau protein which normally has a microtubule-stabilizing function misfolds and forms characteristic long, twisted fibrils which takes on a gain-of-toxic function that adds an amyloid burden to the cell and induces neuronal dysfunction. This fibril conformations propagate by direct molecular interaction with native monomer to induce a conformational change and stack the misfolded monomers in a repeating pattern. These disease strains exit the cell via endosomal escape, exocytosis, or membrane destabilization during apoptosis and are capable of spreading to nearby neurons and glia by macropinocytosis or along synaptically connected pathways. We can study the propagation of disease-strains in culture by isolating tau fibrils from patient brain samples, propagating their structure *in vitro* with recombinant monomeric tau [14], and applying the amplified fibrils to cells expressing monomeric tau. These experiments may be performed in an immortalized cell line to study the cell line-independent propagation of these structures followed by the application of disease-strains to primary neurons to investigate the neuronal subtype-specific uptake and propagation of tau fibrils across different tauopathies.

1.2.3 Identification of Therapeutic Targets and Disease Biomarkers

By decoupling neurodegenerative disease models by their intrinsic and extrinsic effects, we can begin to understand different aspects of disease pathogenesis to identify potential therapeutic targets. Disease models in general may be used as a

proxy to determine which biomarkers or cellular processes are affected in the course of the disease so that the model can yield more predictive results about both safety and efficacy in the *in vivo* models. Since cellular processes are complex and signaling pathways overlap for many different functions including division, proteosomal degradation, energy metabolism, and transcriptional regulation, cellular models are useful for capturing these different levels of interaction to rapidly screen potential small molecule or protein therapeutics with a high level of confidence before expending resources on *in vivo* models. Expanding this concept further by applying cellular programming to patient-derived tissues, we can screen therapeutics in a more directed manner for personalized medicine.

Most neurodegenerative disease-related proteins misfold and interact with natively folded protein monomers to propagate the disease structure. This prion-like propagation of disease conformations in the brain during disease pathogenesis is carried out by direct intermolecular interaction between the disease conformation and the native conformation which suggests the development of therapeutics capable of disrupting this interaction. Antibodies are an attractive candidate because of their highly specific interaction and strong affinity for their target antigen. For prion disease which propagates by direct protein-based interaction, an antibody that binds the native form could disrupt this interaction and halt the progression of the disease. One issue with antibody therapeutics for neurodegenerative disease is the low amount of antibody capable of crossing the blood brain barrier with about 0.1% on average [15, 16]. Single-change variable fragment (scFv) antibodies address some of these issues since its smaller size allows for improved diffusion across the blood brain barrier and through interstitial space. Additionally, the single chain format facilitates expression

and protein engineering to improve therapeutic properties to make them suitable for use *in vivo*.

Chapter 2

MATERIALS AND METHODS

2.1 Culture of Mammalian Cell Lines

All mammalian cells were grown at 37C / 5% CO2 in humidified incubator. All reagents were purchased from Thermo-Fisher unless otherwise stated. Human embryonic kidney (HEK) cells (Invitrogen), N2a mouse neuroblastoma cells (ATCC), and primary mouse embryonic fibroblasts (MEFs) were cultured on tissue culture-treated plastic in high-glucose DMEM with 1x Pen-Strep, 0.1mM non-essential amino acids (NEAA), 1mM sodium pyruvate, 2mM GlutaMax, and 10% fetal bovine serum (FBS). HEK, MEF, and N2a cultures were passaged at 90% confluency using 0.05% trypsin or 1x Cell Stripper non-enzymatic cell dissociation reagent. Human induced pluripotent stem cells (WiCell) were cultured in mTeSR1 (STEMCELL Technologies) on tissue culture plates pre-coated with Matrigel for 1 hour. Human iPSCs were passaged mechanically after treatment with 1x Accutase for 3 minutes. Reprogrammed neurons and primary mouse embryonic neurons were cultured in Neurobasal medium containing 1x B27 supplement and 1x N2 supplement. Long-term stocks of cells were stored in liquid nitrogen at about 1 million cells in 1 mL of culture medium containing 10% dimethyl sulfoxide (DMSO) as a cryoprotectant and slowly frozen at -80C overnight in an isopropanol box before transferring to liquid nitrogen for permanent storage.

2.2 RNA Isolation and cDNA Synthesis

RNA was isolated using TRIzol and ethanol with silica columns (Zymo Research) and eluted with nuclease-free water for further use. For RT-PCR, RNA was reverse transcribed and quantified using the Bio Rad One-Step RT-PCR kit (Bio-Rad) with primers specific to the gene being measured. For transcription factor isolation, RNA was reverse transcribed into cDNA using the Bio Rad cDNA Synthesis Kit (Bio-Rad) and used directly in PCR for amplify desired transcription factors.

2.3 Transcription Factor Amplification, Subcloning, and Screening

Transcription factors were isolated from a variety of sources including cDNA libraries generated from human tissue cultures, MegaMan cDNA library (Agilent Genomics), and plasmid-isolated clones from Addgene or OpenBiosystems. Transcription factors were amplified using primers specific to the TF itself with restriction sites added for subcloning. All transcription factors were cloned into the TetO-FUW plasmid backbone (Addgene). TetO-FUW is a lentiviral vector containing 5` and 3`-LTR sequences and a psi-packaging signal for lentivirus packaging and integration. The multiple cloning site is under a minimal CMV2 promoter with a tetracycline responsive element (TRE) consisting of seven repeats of the tet operator sequence recognized by the TetR repressor protein required for tetracycline or doxycycline expression control. All PCR products and empty vector were digested with appropriate restriction enzymes (New England Biolabs) for 1-2 hours at 37C, run on a 1% agarose DNA gel at 120V for 1 hour and eluted with a gel extraction kit (Zymo Research) in water. Ligations were performed at a 1:3 ration of vector:insert overnight at 16C using T4 DNA ligase (NEB). Overnight ligations were transformed into chemically competent NEB5a *E. coli* cells (NEB) according to the manufacturer's

protocol and plated on LB-amp plates (100ug/mL ampicillin) overnight at 37C. At least four colonies were chosen and pre-screened by check digest before sequencing by standard Sanger sequencing (Eurofins MWG Operon). Positively confirmed clones were stored for long term use at -80C by combining overnight liquid culture in LB amp with 30% glycerol in a 50:50 ratio for a final concentration of 15% glycerol.

2.4 Cloning of Tau Expression Constructs into Lentiviral Vector

Tau constructs were generated by assembly PCR using primers specific to particular regions of tau or the entire tau gene with overlap to fuse the tau PCR with GFP or a hemagglutinin (HA) tag. Restriction enzyme sites were added to the termini of each PCR products by inclusion in the primers. Final PCR products and TetO-FUW were digested with appropriate restriction enzymes (NEB) for 1-3 hours at 37C, purified by gel electrophoresis in a 1% agarose gel at 120V for 1 hour, and extracted using a gel extraction kit (Zymo Research). Ligations were run at a 1:3 ratio of vector to insert with T4 DNA ligase (NEB) overnight at 16C and transformed into NEB5a cells (NEB) for screening and sequence confirmation. Positive transformants were stored at -80C in 15% glycerol for long-term storage or grown in LB-amp and plasmid DNA was isolated by MidiPrep (Zymo Research) for further experiments.

2.5 Transient Transfection of Tau Constructs and Fibril Preparation/Addition

Tau fibrillization experiments were performed by seeding HEK cells on tissue culture plastic in complete DMEM with 10% FBS at about 20% confluency overnight followed by transfection with tau plasmid DNA. Tau construct transfection was performed by incubation of 0.15ug plasmid DNA/cm2 in 25uL serum DMEM/cm2 with 2:1 TransIT transfection reagent (Mirus Bio, LLC) (2uL per ug DNA) for 10 min

at room temperature before adding dropwise to the adherent HEK cells. Transfections were incubated overnight at 37C to allow for expression of the tau construct before fibrils were added.

Fibrils were prepared as previously described [14] by Olga Morozova using recombinant tau protein to propagate the tau strain conformations. Upon completion of fibrillization, tau fibrils were harvest by centrifugation at 21k x g for 45 min and resuspended in 1x PBS buffer to a final concentration of 0.1 mg/mL. Fibrils were pre-incubated at 1ug/well (24 well tissue culture plate) in 50 µL buffer with 1 µL Lipofectamine-3000 (Invitrogen) for 10 min at room temperature. The fibril transfection mixture was vortexed briefly before adding dropwise to HEK cells expressing various tau constructs and incubated for 36 hours before assaying by microscopy or lysis and centrifugation for structural assays.

2.6 Production of Lentivirus in HEK Cells

Lentivirus for mammalian expression of transcription factors, tau constructs, or secreted ICSM18 scFv and GFP reporter was generated by transient co-transfection of Human Embryonic Kidney (HEK) 293T cells with 2nd generation packaging plasmids (Addgene) and the pHAGE expression vector. HEK 293T cells were grown in Dulbecco's Modified Eagle's medium (DMEM) containing 4.5 mg/ml glucose, 10% fetal bovine serum (FBS), 100 U/ml penicillin, 100 ug/ml streptomycin, 1 mM sodium pyruvate, 1X NEAA, and 2 mM GlutaMax. Cells were cultured in DMEM at 37°C in a humidified atmosphere of 95 % air/5 % CO2 and subcultured every 3-4 days using 10 mg/ml trypsin. Briefly a 100mm dish of non-confluent 4 x 10^6 cells were transfected with 5 µg psPAX2 (Addgene, Cambridge, MA, USA, Plasmid 12260), 7.5 µg pMD2G (Addgene, Plasmid 12259), and 10 µg of lentiviral vector plasmid

(pHAGE) in 500 µl serum-free DMEM with 45 µl (2:1, ul transfection reagent : ug DNA) TransIT-293 (Mirus Bio) and incubated at room temperature (RT) for 10 min. Following the incubation, transfection mixture was added dropwise to HEK 293T cells and rocked to distribute. Lentivirus-containing supernatant was harvested 48 hrs post-transfection by filtration with 0.45 um syringe-tip filter (CellTreat) for subsequent experiments.

2.7 Lentiviral Transduction of Mammalian Cells

Lentivirus supernatant collected from HEK viral production and filtration was added directly to cells at an appropriate concentration determined by titration and qPCR to calculate IU/mL. Cells being transduced with viral supernatant were pre-treated with 6 µg/mL polybrene (SigmaAldrich) to increase viral uptake and lentivirus was added directly to the culture wells and incubated at 37C overnight to allow transduction to occur. Medium was changed on the second day to remove polybrene and feed cultures with fresh medium. Cell cultures for reprogramming were either assayed for fluorescence by flow cytometry, tested by RT-PCR for integration, or cultured for 1-3 weeks in neuronal medium for neuronal reprogramming experiments.

2.8 Determination of Multiplicity of Infectivity by Quantitative PCR

Lentiviral titer was measured by quantitative-PCR of genomic DNA (gDNA) relative to a titration of known gDNA concentration and a titration of known plasmid concentration. Q-PCR of gDNA was chosen because other methods typically have much higher variability. ELISA of viral coat proteins, qRT-PCR of packaged cargo RNA, and comparison to a fluorescent control have as much as one to two orders of magnitude difference from actual infectious units (IU), whereas measuring integrated

copies directly is a much more straightforward and reproducible method. Briefly, about 2.5 x 10^5 cells were dissociated using 0.05% trypsin (ThermoFisher) and gDNA was isolated using silica columns (Zymo Research). Quantitative PCR was run using a 2x Master Mix containing SYBR Green without reverse transcriptase (Bio-Rad) on a Bio Rad CFX96 RT-PCR cycler with primers against the gene of interest relative to a plasmid titration (10^3-10^7 plasmid copies) and primers against GAPDH for gDNA quantification relative to a titration of gDNA (10^2-10^5 genomic copies). Linear regression was performed for the plasmid and gDNA titrations to approximate integrated viral copies and genome copies so that a ratio may be taken to estimate the multiplicity of infectivity.

2.9 Breeding of Bi-transgenic Mice and Dissection of Embryonic Tissue

Bi-transgenic mice were generated from crossing two individual transgenic mouse lines. TauGFP mice (Jackson Laboratory)) were maintained as a homozygous line by breeding homozygous parental mice so that all littermates are positive for the tauGFP transgene. Drd1aRFP mice (Jackson Laboratory) were maintained as a heterozygous line by crossing with C57BL/6 wildtype mice and using a genotyping PCR on tail tip isolations to determine which littermates contained the drd1aRFP transgene. Bi-transgeneic mice were generated by breeding the homozygous tauGFP mice with drd1aRFP mice with an expected Mendelian ratio of 50% tauGFP-only pups and 50% bi-transgenic pups. Embryonic day 14.5 (E14.5) mouse embryonic fibroblasts (MEFs) and primary neurons were harvested following 14 days post-positive identification of a plug after setting up breeding pairs. Briefly, pregnant female mice were sacrificed according to standard protocol using carbon dioxide asphyxiation. Uterine horns containing the E14.5 embryos were dissected out and

rinsed with sterile PBS. Individual embryos were isolated, neuronal tissue from the brain and spinal cord were removed, visceral organs were removed, and the remaining fibroblasts were incubated in 0.05% trypsin for 5 minutes before dissociation by trituration. The resulting MEFs were cultured for one passage in complete DMEM with 10% FBS before freezing or use in further experiments. E14.5 primary neurons were isolated from embryo brains by incubation in trypsin and gentle trituration and cultured in complete Neurobasal medium (Invitrogen).

2.10 Detection of MEF Reporter Expression and Sorting by FACS

MEFs being used for reprogramming were seeded on tissue culture-treated plastic at about 50% confluency overnight followed by lentiviral transduction as described overnight with 2mg/mL doxycycline to induce TF expression. MEF cultures were then incubated in complete Neurobasal medium containing 2 mg/mL dox for 1-3 weeks and fed with fresh complete neuronal medium every third day which reprogramming was taking place. End point MEFs as dictated by individual experiments were either checked by microscopy or dissociated with trypsin and checked by flow cytometry for reporter fluorescence intensity and percent positive reprogrammed cells. Positive reprogramming samples were sorted by fluorescence activated cell sorting (FACS) on a FACS-Aria I or II (Christiana Care Hospital and the Papoutsakis Lab) for cells positive for reporter expression. Sorted cells were collected and RT-PCR was used to quantify factors either enriched or depleted relative to cells sorted for no reporter expression.

2.11 Quantification of Fibril Stability by Guanidine Titration and Dotblot

Dot blots were used to identify the presence of aggregated tau protein or for guanidine destabilization curves for tau fibril stability measurements. Parent fibrils generated from the tau fibrillization reaction [14] were centrifuged for 45 min at 21k x g and resuspended to 100 ng/µL in PBS. Tau fibrils (100 ng/rxn) were incubated in a range of guanidine-HCl as noted for 24 hrs at room temperature. Intracellular fibrils were harvested from HA-tau-expressing HEK cells 36 hours after seeding with parent tau fibrils or buffer control. HEK cells (~2.5 x 10^5 cell) were dissociated with 250 µl Cell Stripper non-enzymatic cell dissociation reagent to avoid proteolysis of tau fibrils. To the cell suspension, 50 µl of lysis buffer was added for 10min at room temperature for final concentrations of 0.1% sarkosyl, 1mM DTT, 1mM PMSF, 1x mammalian PI cocktail (SigmaAldrich), 5U Benzonase endonuclease (SigmaAldrich), and 1mM $MgCl_2$. The clarified lysate was centrifuged at 21k x g for 45min, the fibril pellet was resuspended in 100 µL PBS, and 10 µL of intracellular fibrils were incubated in a range of guanidine-HCl concentrations for 24 hrs at room temperature. Parent fibrils and intracellularly-derived fibrils incubated in GdnHCl for 24 hrs were diluted to 200 µL total volume with PBS to dilute the GdnHCl before running the samples on a 96 well Bio-Dot microfiltration apparatus through a cellulose acetate membrane with a 0.2 µm porosity. Cellulose acetate is non-protein binding and so only fibrillar tau is trapped on the membrane while monomeric tau flows through the membrane to waste. Membranes were blocked in PBS with 0.1% Tween-20 (PBST) and 1% bovine serum albumin (BSA). Blocked membranes with recombinant tau parent fibrils were probed with 1:100 N-terminal tau antibody 5A6 (Iowa Hybridoma Bank) and a mouse secondary HRP antibody. Membranes with intracellular HA-tau fibrils were probed with 1:5000 anti-HA Y11 antibody (SCBT) and a rabbit secondary

HRP. Probed membranes were incubated with chemiluminescence reagent and exposed to detect protein content. To measure fibril stability, chemiluminescence intensity at each concentration was quantified by pixel intensity and background was subtracted with a constant measurement area. Pixel intensities were fit to a sigmoidal unfolding curve and the guanidine concentration at which half of the fibrils were remaining was taken as the measurement of stability.

2.12 Structural Comparison of Tau Fibrils by Limited Proteolysis

Limited proteolysis was used to compare fibril strains of different conformations by selectively cleaving solvent-exposed regions of the protein in a regular manner with a protease resulting in a unique band pattern specific to a particular strain. Parent tau fibrils were used directly in the assay at a concentration of 1ug fibrils in 19uL buffer with 1uL 10mg/mL trypsin for 1 hr at 37°C. For intracellular fibrils, 2.5 x 10^5 HEK cells seeded with 1 µg tau for 36 hrs were dissociated with 250 µL Cell Stripper and lysed with 50 µL lysis buffer for a final concentration of 0.1% sarkosyl, 1mM DTT, 5U Benzonase endonuclease, and 1mM $MgCl_2$ for 15 min at room temperature. Clarified lysate was spun for 45min at 21k x g and resuspended in 19 µL PBS and 1 µL 10mg/mL trypsin was added and incubated at 37°C for 1 hour. Digested samples were stopped with 20 µL 2X SDS loading dye, boiled for 5 min at 100°C, and run on a small peptide 16% Tricine gel for 90min at 130V. SDS-PAGE gels were transferred to nitrocellulose for 45 min at 25V, blocked with PBST+1%BSA for 20min and probed with 1:5000 77G7 anti-tau antibody (BioLegend) and mouse secondary HRP antibody. Membranes were incubated with chemiluminescence reagent and exposed to image the digested tau bands. Both trypsin

and Pronase protease cocktail were tested with similar results and trypsin was chosen for subsequent experiments because the bands were slightly cleaner.

2.13 ICSM18 Synthesis and Plasmid Construction

ICSM18 scFv was synthesized as described [17]. The ICSM18 scFv and anti-fluorescein 4-4-20 scFv (obtained from Anne S. Robinson) were subcloned into the lentiviral mammalian expression vector pHAGE-CMV-dsRed-UBC-GFP-W (Addgene plasmid #24526) in place of the dsRed reporter by restriction digest cloning such that the final construct included an N-terminal Igκ light chain secretion signal and a C-terminal FLAG tag (DYKDDDDK). For yeast experiments, ICSM18 was cloned into pCTCON2 between the Aga2 yeast surface protein and a myc tag (EQKLISEEDL) for expression detection and transformed into EBY100 yeast using the lithium acetate method [18]. The protein was expressed on the surface by induction in galactose medium overnight. For yeast secretion experiments, ICSM18 was cloned into the integrating pITy vector [19], transformed into YVH10 yeast, and a clone was selected for further use. Secretion was guided by an alpha-secretion signal upon induction with galactose-containing medium at 20°C for 1-3 days.

2.14 Generation of Stable Antibody-Secreting CHO Lines and Antibody Purification

A Chinese hamster ovary (CHO-S) cell line adapted to serum-free and suspension culture was obtained from Kristen Valente and Kelvin Lee [20]. CHO cells were cultured in 125 ml shake flasks in serum-free culture medium (GE Healthcare HyCLone SH30549.01) at a seeding density of 2×10^5 cells/ml and passaged every 3-4 days. Stable cell lines for scFv secretion were generated by transduction of 1×10^5 cells with scFv lentiviral particles in 10 ml culture medium containing 6 ug/ml

polybrene (SigmaAldrich) and grown at 37 C / 5% CO2 with shaking for 7 days. Stably-transduced CHO cells were then sorted by FACS for the GFP reporter to enrich expressing cells. Median fluorescence intensity (MFI) of GFP reporter in CHO cell lines was assessed by flow cytometry (Accuri C6). Expression of scFv was carried out by culturing 1 x 10^7 transduced CHO cells in 20 ml for 7 days. Culture was then centrifuged at 3,000 x g for 5 min and filtered with a 0.22 µm filter and purified using anti-FLAG affinity chromatography (SigmaAldrich). Purified scFv was concentrated using Amicon Ultra-15 10 kDa-cutoff centrifugal filters (EMD Millipore UFC901024), assayed for purity by silver stain (Life Technologies 24612), quantified by BCA assay, and stored in TBS at 4°C for short-term use or -20°C for long-term storage. ScFv was detected using Western blot by running samples with 1x SDS Loading Buffer on to a 4-20% Tris-Gly gel (Pierce Biotechnology), transferring to nitrocellulose, and probing with 1:1000 anti-FLAG M2 (Sigma F3165) antibody and 1:10,000 anti-mouse-HRP.

2.15 Antibody Mutagenesis and Sorting by FACS with Yeast Surface Display

Mutagenesis library construction from the pCTCON2-ICSM18 yeast surface display vector was performed by Kyle Doolan as previously described [21]. The transformation and homologous recombination step led to a library of 3 x 10^6 members determined by serial dilution plating onto tryptophan-deficient plates. Sanger sequencing of selected clones provided an effective mutation rate of 6 amino acid mutations per member in the library. Yeast libraries were grown at 30°C overnight in tryptophan-deficient glucose medium, harvested by centrifugation, resuspended to OD 1.0 in tryptophan-deficient galactose medium, and grown from 18-24 hrs at 20°C to induce protein expression. Yeast labeling was performed as previously described [22].

Surface expression was labeled with 1:600 chicken anti-c-myc (Life Technologies A21281) and 1:600 anti-chicken Alexa Fluor 488 (Life Technologies A11039). Surface-displayed scFv was labeled with 10 nM fluorescently labeled recombinant prion protein. Labeled yeast were sorted by fluorescence activated cell sorting (FACS) on a FACS Aria II for multiple rounds and a second round of mutagenesis until desired properties were obtained.

2.16 Purification and Fluorophore Conjugation of Soluble PrP

The wild-type mouse prion protein gene (Prnp) encoding the mature form of the protein (residues 23-231) was obtained by removal of the 3F4 epitope tag from a PrP expression vector by site-directed mutagenesis (Addgene plasmid 1321) [23]. Site-directed mutagenesis was performed using the Phusion site-directed mutagenesis kit (Life Technologies F-541) according to the manufacturer's instructions and the following primers: [Phos]TTTGCTGGGCTTGTTCCACTG and [Phos]CCAAAAACCAACCTCAAGCATGTGGCAGGGGCTGCG (Eurofins MWG Operon). Wild-type mouse Prnp was cloned into *E. coli* expression plasmid pET11a. Recombinant protein expression was induced by addition of 50 µM IPTG to BL21 cells (NEB) transformed with the pET11a-MoPrP plasmid when the culture reached OD 1.0 for an additional 4 h. Cells were harvested by centrifugation and purified as described previously [24] to obtain lyophilized, oxidized recombinant prion protein.

For fluorescent labeling, lyophilized recombinant PrP was resuspended in water to 2 mg/ml and quantified by absorbance at 280nm. The soluble protein was rapidly added to a solution of Alexa Fluor 647 Carboxylic Acid, Succinimidyl Ester (Life Technologies A20006) in pH 8.3 sodium bicarbonate buffer for a final concentration of 1 mg/mL PrP and stoichiometric equivalents of Alexa Fluor 647 and

incubated rocking for 1 h at room temperature, protected from light. Labeled, precipitated PrP was recovered by centrifugation and solublized in refolding buffer consisting of 4M Urea, 0.25 mM oxidized glutathione (GSSG), and 0.1M Tris (pH 8) for 90 min [25]. Concentration and degree of labeling was checked by Nanodrop for absorbance at 280 nm and 647 nm. For further use, PrP in refolding buffer was rapidly diluted into 150 mM sodium phosphate buffer (pH 7.5) to a concentration of 0.01 ug/ml or lower to minimize precipitation.

2.17 Quantification of ScFv Affinity by PrP Titration on YSD ScFvs

Binding affinity of yeast surface-displayed scFvs were assessed by titration of AlexaFluor-647 fluorescently-labeled soluble prion protein. Clones were grown at 30 C overnight in tryptophan-deficient glucose medium, harvested by centrifugation, resuspended to 1 x 10^5 cells/ml in tryptophan-deficient galactose medium, and grown for 20 hrs at 20°C to induce surface-display expression. Yeast labeling was performed as previously described [22], using 1:600 chicken anti-c-myc (Life Technologies A21281) and then 1:600 anti-chicken Alexa Fluor 488 (Life Technologies A11039). Yeast labeled for c-myc expression were resuspended in 150mM sodium phosphate buffer (pH 7.5) and incubated in a titration of Alexa Fluor 467-labeled PrP concentrations in molar excess for 30 min followed by washing with PBS and analyzed on an Accuri C6 flow cytometer. The median fluorescence intensity (MFI) of c-myc-expressing cells was used for quantification and fit to a three-parameter binding curve by non-linear least squares fit.

2.18 Quantification of CHO-Expressed ScFv on N2a Neuroblastoma Cells

Mouse neuroblastoma cell line, N2a (American Type Culture Collection CCL-131) were grown to near confluence in a T-75 culture flask in complete DMEM/F-12 + 10% FBS and dissociated from the surface using non-enzymatic cell dissociation reagent CellStripper (Corning 25-056-CI). Cells were pelleted and washed with phosphate buffered saline (PBS) by centrifugation at 1,000 x g for 5 min and resuspension in PBS. FLAG affinity purified scFv was applied to 5 x 10^5 N2a cells in PBS/1% BSA dissociated from a flask using CellStripper to preserve expressed PrP, at concentrations ranging from 0.0316 nM to 100 nM. After a 45 minute incubation at 37°C cells were centrifuged at 16,000 xg for 1 min and washed in PBS+1%BSA. Cells were then incubated in 1:500 anti-FLAG M2 antibody in PBS+1%BSA for 20 min at 4°C. After incubation, cells were again centrifuged at 16,000 x g for 1 min and washed in PBS+1%BSA. Cells were then incubated in 1:200 anti-mouse Alexa-647 antibody in PBS+1%BSA for 20 min at 4°C. After incubation, cells were again centrifuged at 16,000 x g for 1 min and washed in PBS+1%BSA. Cells were resuspended in 200 µL PBS+1%BSA and run on Accuri C6 to measure fluorescence of bound scFv. Affinity was determined by fitting median fluorescence of expressing populations to a sigmoidal equilibrium binding curve.

2.19 Quantification of ScFv Stability by Differential Scanning Fluorimetry

SYPRO Orange (ThermoFisher) was used to assess thermal stability [26] of scFvs purified by anti-FLAG affinity chromatography from CHO supernatant by differential scanning fluorimetry (DSF). Briefly, 100 µg/ml scFv or buffer was mixed with 5X SYPRO Orange in a 25 µl volume. A RT-PCR cycler (Bio Rad, CFX96) was programmed to ramp the temperature at a rate of Δ0.5°Cper 30 sec, taking a

fluorescence reading every 30 sec with an appropriate LED/photodiode set for the excitation and emission of SYPRO Orange. Technical replicates were run in triplicate on two different preparations of scFv and buffer control samples were subtracted to remove background fluorescence. The temperature at the peak signal of the first derivative of the fluorescence transitions curves was taken as the melting temperature of the scFv.

2.20 *In vitro* Aggregation of ScFv

Aggregation propensity of purified scFv samples was tested by incubation at 3.5 ug/ml in PBS at 37°C without agitation for 24 or 48 hrs. Aggregated samples were tested by SDS-PAGE on a 4-20% Tris-Gly gel (Pierce Biotechnology) in 1x SDS Loading Buffer with or without reducing agent followed by Western blot probed with anti-FLAG M2 antibody (SigmaAldrich).

Chapter 3

DEVELOPMENT OF A HIGH-THROUGHPUT PLATFORM FOR NEURONAL REPROGRAMMING

3.1 Introduction

Many neurodegenerative diseases, including Alzheimer's and Huntington diseases, target specific subtypes of neurons in the brain. However, the ability to derive cell culture models of specific neuronal lineages needs improvement if we are to adequately mimic affected cell types and develop *in vitro* models from human-derived cell lines. We seek to control differentiation of human fibroblasts into targeted neural subtypes by engineering the transcriptional network of human fibroblasts. The knowledge gained through our work will facilitate the development of cell culture models of neurological diseases from patient derived cells.

Fibroblasts can be reprogrammed into induced Pluripotent Stem (iPS) cells [27, 28], enabling subsequent differentiation into neurons e.g. [29-31]. Reprogramming of fibroblasts directly into functional neurons by either overexpression of neuronal-inducing transcription factors [32, 33] has also been reported. Some specific neural lineages have been generated, resulting in useful human-derived disease models of Parkinson's disease [34], but much remains to be accomplished if we are to create, *in vitro,* the myriad of neuronal cells that make up the brain in high yield and purity (**Figure 3.1**). In particular, derivation of basal forebrain cholinergic neurons (BFCNs), the neuronal subtype primarily affected in

Alzheimer disease, and striatal medium spiny neurons (MSNs), strongly affected in Huntington disease, remains difficult and can only be achieved in limited yield.

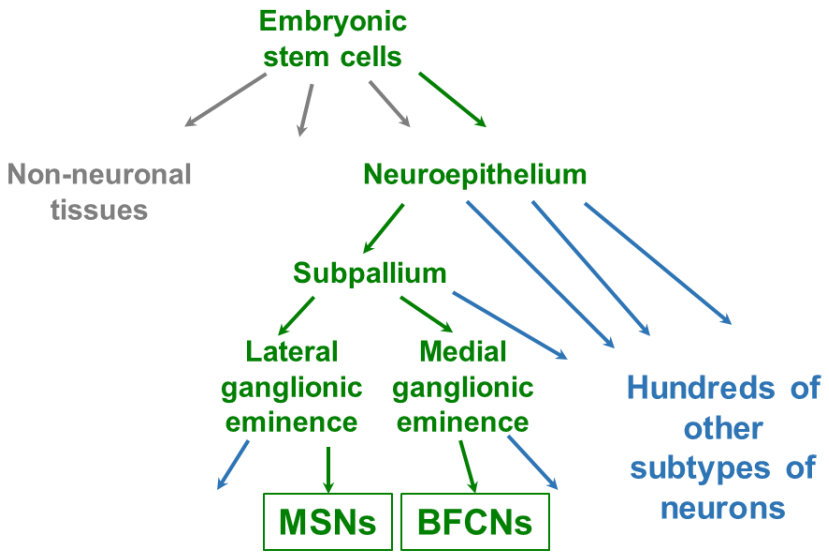


Figure 3.1 Neuronal differentiation pathway *in vivo* generates hundreds of individually identifiable neuronal subtypes. Medium spiny neurons (MSN) and basal forebrain cholinergic neurons (BFCN) for example, follow a highly specific, well-defined differentiation path.

A seminal study by Yamanaka and Takahashi in 2006 showed that the ectopic expression of several key transcription factors (TF) in mouse embryonic fibroblasts was sufficient to induce a state of pluripotency similar to that of embryonic stem cells [27]. This introduced the idea that a terminally differentiated cell type could be transdifferentiated into another cell type merely by expression of particular transcription factors. Cellular reprogramming by ectopic expression of transcription factors is a powerful approach to direct differentiation of patient derived cells into

desired cell types and subtypes that complements traditional culture medium-based approaches that rely on spatiotemporal presentation of specific soluble factors.

Cellular reprogramming is a process by which neurons may be derived from renewable cell types, such as fibroblasts and induced pluripotent stem (iPS) cells, obtained from living patients. For many years, *in vitro* neurogenesis research has focused on how to direct differentiation of stem cells to specific subtypes of neurons using soluble factors [35] which is a complicated long and expensive process of sequentially incubating cultures with particular patterning morphogens for specific periods of time (**Figure 3.2**). Though much has been accomplished using this approach, progress has been slow, leaving the complex diversity of neuronal subtypes relatively unattainable from ES/iPS cells. In 2010, neuronal cells were successfully derived directly from mouse embryonic fibroblasts using three specific transcription factors [32]. A year later, neurons were derived from human post-natal and adolescent fibroblasts using the same factors[33]. Transdifferentiation of fibroblasts into dopaminergic neurons and several other subtypes of neurons have also been achieved with different sets of transcription factors[36, 37]. These results indicate the potential for fibroblasts or iPS cells to transdifferentiate into different types of neurons by reprogramming.

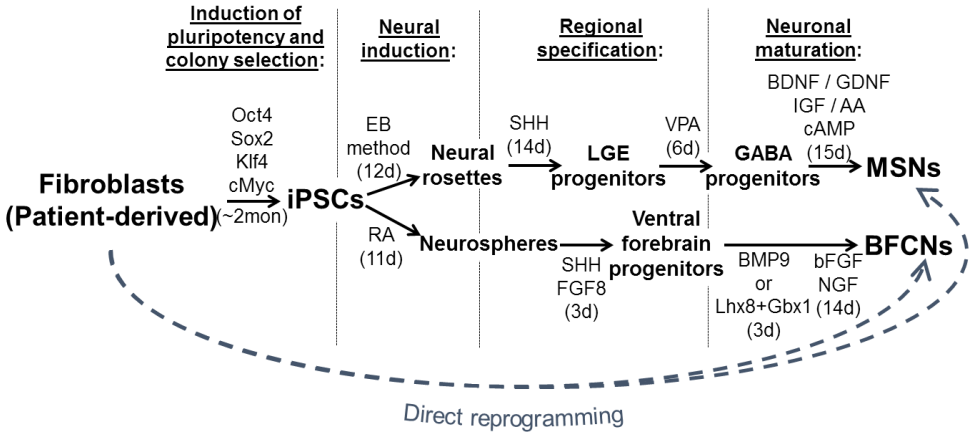


Figure 3.2 Current standard protocols for the derivation of MSNs [38] and BFCNs [39] from patient-derived fibroblasts. Direct neuronal reprogramming (dotted arrows) provides a faster alternative with fewer spatiotemporal inputs. LGE, lateral ganglionic eminence; MSN, medium spiny neuron; BFCN, basal forebrain cholinergic neuron.

HD *in vitro* models have been developed that use patient-derived HD iPS cells or neural stem cells derived from them[11-13]. However as one study noted, both proliferation and caspase activity were not significantly different between HD and non-HD controls while the difference was significant comparing NSCs, indicating the importance of lineage restriction in differential vulnerability to the expanded CAG mutation [13]. HD patients experience massive loss of medium spiny neurons (MSN) in the striatum and a widespread loss of cortical neurons at later stages while some neuronal subtypes remain relatively spared[14]. Therefore it is important to evaluate the effect of deriving distinct terminally differentiated neuronal subtypes from HD patient cells. As of the time of this work, MSNs and many other subtypes have not yet been generated *in vitro* in high yield by neuronal reprogramming. As the process of differentiating pluripotent cells into neurons is both expensive and time-consuming,

neuronal reprogramming provides a faster alternative to generate more neuronal subtypes *in vitro* to investigate HD pathogenesis and screen for therapeutics.

Alzheimer's disease *in vitro* models have also been developed using cellular reprogramming [40-43], however these models do not span a sufficient spectrum of neural subtypes to explain differential vulnerability observed in AD pathology. Some recent reports demonstrate the utility of directly reprogrammed AD fibroblasts [42] or disease-specific human iPSCs differentiated mainly into glutamatergic neurons [41] or neurons of an undetermined subtype [40, 43], but not the earliest and most affected subtype, forebrain cholinergic neurons [44].

Here we propose to develop methodologies for the direct reprogramming of patient-derived fibroblasts into diverse neural subtypes using a high-throughput platform in which a library of neuronal transcription factors is serially tested by random combinations and fluorescence activated cell sorting (FACS) to identify a subset of factors sufficient to reprogram the starting cell type into a particular subtype of neuron (**Figure 3.3**). Previous reprogramming efforts rely upon a low-throughput approach by systematically testing different combinations of factors individually until a desired cell type is obtained. Our primary objectives for establishing this platform were derivation of medium spiny neurons (MSNs) which are most affected in the early stages of HD followed by basal forebrain cholinergic neurons (BFCNs) most affected in early stages of AD. However, the approaches and resources we develop should be broadly applicable to the generation of other neural subtypes. The availability of a spectrum of patient derived cell culture models of neurodegenerative disease, each with distinct neural subtype identities, will enable researchers to identify cellular

features that correlate with disease susceptibility and lead to *in vitro* models suitable both for studying disease pathogenesis and screening for potential therapeutics.

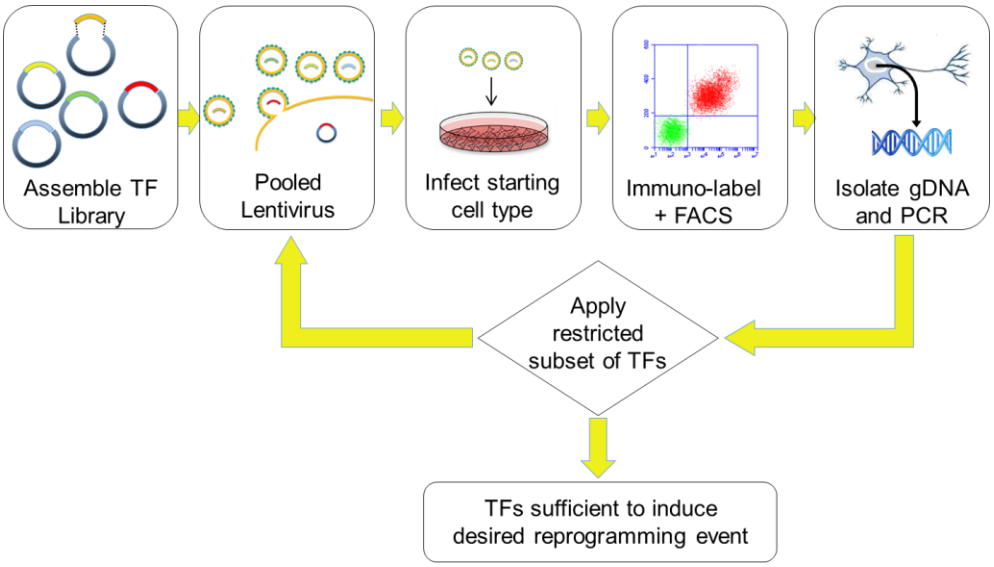


Figure 3.3 Schematic of high-throughput neuronal reprogramming relying upon an iterative process of random infection and selection for particular neuronal markers

3.2 Selection and Isolation of Reprogramming Library Factors

The goal of a high-throughput reprogramming library is to be as broadly applicable as possible to generate any subtype of neuron. Based on a search from literature and Gene Ontology for genes involves in the development and maintenance of neurons, 1033 neuronal genes were found. By cross-referencing this list with a list of known transcription factors identified by Gene Ontology, a reduced list of 183 neuronal transcription factors was generated (**Figure 3.4**). This is a large number of factors and so the list was reduced by focusing on factors known to be related to the *in vivo* development of medium spiny neurons (MSNs) most affected in Huntington's

disease and basal forebrain cholinergic neurons (BFCNs) earliest-affected in Alzheimer's disease as well as factors known to play an important role in neurogenesis in general. Reducing the list left a more manageable set of about 60 transcription factors (**Figure 3.4**).

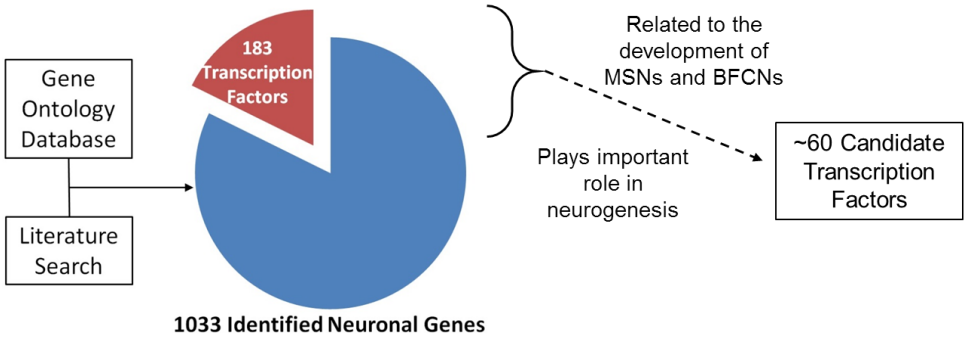


Figure 3.4 Neuronal genes were selected from the Gene Ontology (GO) Database and literature. From the total set of transcription factors involved in these pathways, factors were selected that are specifically involved in the differentiation of MSNs and BFCNs as well as the factors most commonly studied in literature for their role in neurogenesis

The delivery/expression platform for a reprogramming library and the degree of expression control are both important choices. We chose to subclone all factors into a lentiviral expression vector (TetO-FUW) for several reasons. Lentiviral delivery is both fast and highly efficient, resulting in the stable integration of the expression cassette in contrast to transient transfection. Additionally, we wanted to avoid a partially reprogrammed state in which sustained high expression of exogenous transcription factors may support an artificial and transient reprogramming state. We therefore chose a vector with an inducible tetracycline responsive element (TRE) which allows for expression control to turn off exogenous expression ensuring that a

new and stable transcriptional state is reached. The TRE is controlled by a reverse tet-transactivator (rtTA). A plasmid expressing the rtTA is co-transduced with the library to ensure dox-control. Doxycycline is a more stable molecule which is better able to induce tetracycline-inducible expression.

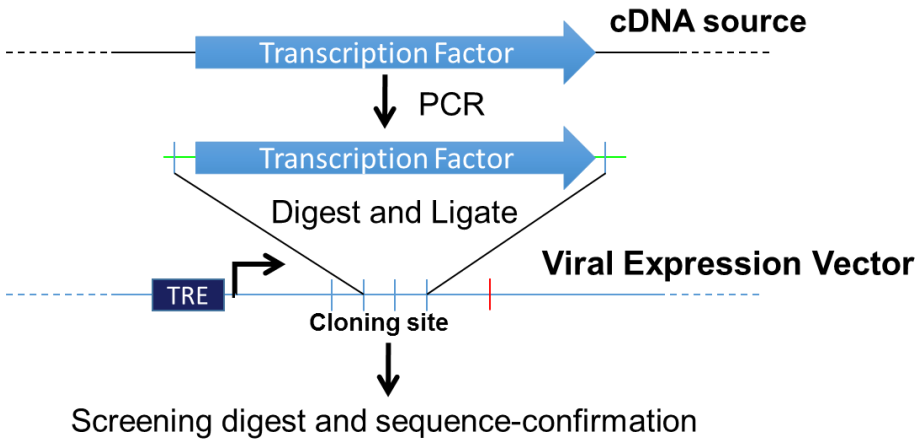


Figure 3.5 Flow diagram for transcription factor (TF) library generation. Amplified TFs are cloned into a lentiviral vector with dox-responsive expression and sequenced by standard Sanger sequencing to confirm proper integration

In brief, the process consists of isolating individual transcription factors from various cDNA sources including cDNA libraries generated from human tissue, commercially available MegaMan cDNA library (Agilent Technologies), and Addgene cDNA clones and subcloning them into the final expression vector (**Figure 3.5**). The resulting transformations are screened for positive transformants by a screening digest and confirmed by Sanger sequencing to ensure that each clone was properly integrated in frame. Amplifying transcription factors is complicated by a

variety of factors including their low abundance due to the nature of transcriptional control and also that they're typically >75% GC-content which makes them difficult to amplify by PCR. Some transcription factors require extreme DMSO concentrations (>10%) or assembly PCR after amplifying individual exons in situations in which the entire transcriptions factor will not amplify correctly. So far a total neuronal reprogramming library of over 40 TFs has been generated which is the largest current reprogramming library published (**Table 3.1**). While we have prioritized the selection of TFs for MSN and BFCN generation as well as general neurogenesis, we have shown the ability to subclone a much larger library to encompass the full objective of all neuronal transcription factors.

Ascl1	EMX2	LHX8	NEUROG1	NR4A2	SOX2
ASCL2	EN2	LMX1A	NEUROG2	OLIG2	SOX5
Brn2	ISL2	MEF2C	NKX2-2	PAX6	SP1
DLX1	LHX1	Myt1l	NKX6-1	PAX7	SP3
DLX2	LHX2	NeuroD1	NKX6-2	PHOX2A	TLX3
DLX6	LHX5	NEUROD2	NR2F1	PITX3	Zic1
EMX1	LHX6	NEUROD4	NR2F2	POU4F1	

Table 3.1 Current list of TFs successfully isolated and subcloned into the lentiviral expression vector representing a broad range of neurogenic factors as well as MSN-specific and BFCN-specific factors

3.3 Inducible Control of Expression of Exogenously Delivered Factors

Lentivirus of the rtTA fusion protein is co-transduced with library factors to control expression. The rtTA fusion consists of a mutant of the *E. coli* TetR tet-repressor protein to the VP64 viral transcriptional activator domain. The TetR protein typically binds the TRE region constitutively and is released upon binding of tetracycline or its analogue doxycycline. The mutations in the TetR part of the rtTA reverses its functionality such that it only binds the TRE upon addition of dox. The VP64 domain is one of a group of known transcriptional activators which promotes assembly of the transcriptional initiation complex when localized to a gene by the TetR domain. The TRE consists of seven repeats of the tet-operator sequence which is recognized by the rtTA. Since the TetO-FUW plasmid contains a basal CMV2 minimal promoter, there is some basal expression without dox-induction. Upon addition of dox, the rtTA binds the TRE and expression is significantly increased. To test this system, we used the fluorescent protein RFP. Cells that had been transduced

with RFP lentivirus and rtTA but without inducer show a modest increase in fluorescence due to basal RFP expression (**Figure 3.6**). After adding dox, the expression level increases by multiple orders of magnitude (**Figure 3.6**). After removal of dox, expression of the exogenous protein construct will drop back to basal expression as host cell proteolysis mechanisms naturally degrade the expressed protein.

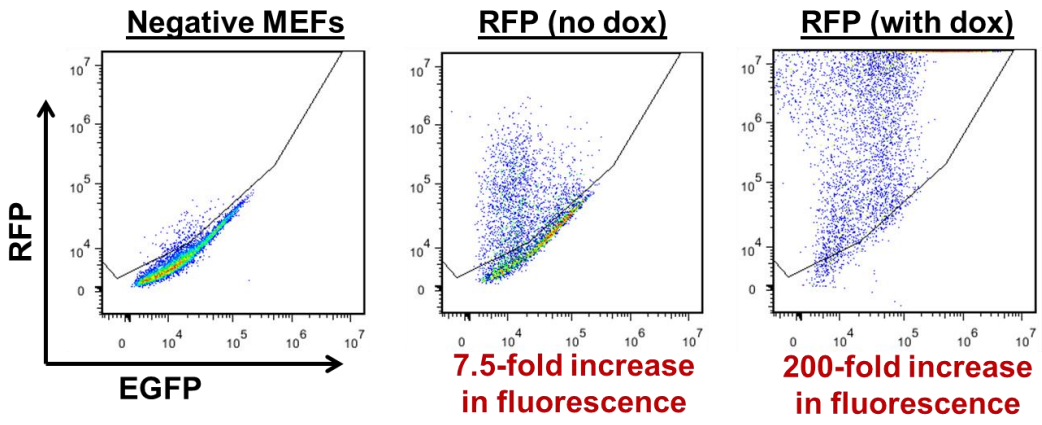


Figure 3.6 Demonstration of dox-inducible control of a control RFP expression construct. Flow cytometry plots of control MEFs without lentivirus (left) and MEFs infected with both RFP lentivirus and reverse-tet transactivator (rtTA) lentivirus without inducer (center) and in the presence of doxycycline inducer (right).

3.4 Generation of Bi-Transgenic Reporter Mouse Embryonic Fibroblasts

Cellular reprogramming is complicated by the necessity to appropriately distinguish between the original population of calls and the successfully reprogrammed cells. The most common option is to immunolabel the cells for particular markers of the desired subtype of neuron. We chose to take an alternate option using transgenic mouse embryonic fibroblasts (MEFs). An advantage of MEFs

is that they are more readily reprogrammed than adult primary human fibroblasts [45, 46]. We took a mouse line homozygously expressing enhanced green fluorescent protein (EGFP) under the tau locus which expressed EGFP in all neuronal cells and crossed it with a mouse line heterozygously expressing red fluorescent protein (RFP) under the dopamine D1a receptor locus which expressed RFP in cells expressing the dopamine D1a receptor (Drd1a) (**Figure 3.7**). Dopamine D1 receptor is expressed in the types of neurons most affected in Huntington's disease, MSNs. Neurons taken from the transgenic embryos show strong reporter expression (**Figure 3.7**). The MEFs harvested from these bi-transgenic mice therefore so not express either fluorescent reporter, but upon transdifferentiating into the appropriate neuronal subtype, they will begin to express both reporters which can be detected by fluorescent microscopy or flow cytometry.

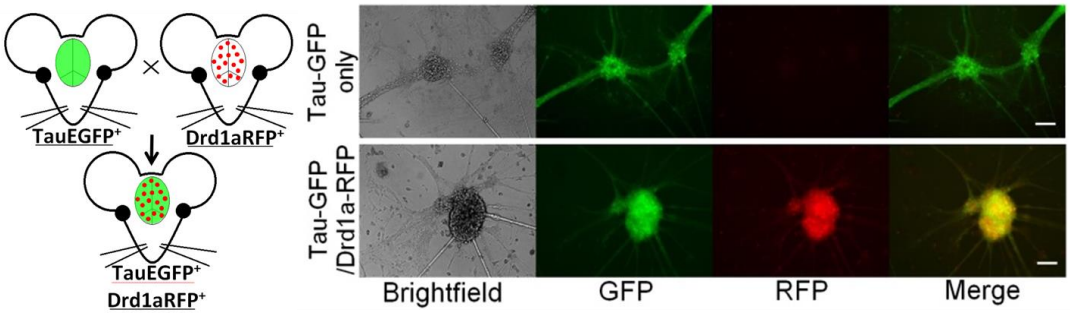


Figure 3.7 Schematic of transgenic mouse breeding (left) to obtain bi-transgenic mouse embryonic fibroblasts containing an EGFP reporter under control of the pan-neuronal tau locus and a RFP reporter under control of the dopamine receptor Drd1a locus. Microscopy of primary embryonic neurons (right) isolated from TauEGFP or bi-transgenic embryos.

3.5 Confirmation of Transcriptional Activity of Exogenous Proteins

Since transcription factors control gene expression in the cell by definition, we chose to make sure that the TFs expressed in our system were still functional after subcloning, producing lentivirus, and exogenously expressing in MEFs. The TF Ascl1 is considered a pioneer transcription factor because of its ability to induce expression of its target genes even in non-neuronal cells. In a positive feedback loop, Ascl1 is known to self-induce its own expression and another known target is Gadd45y. By qRT-PCR, we confirmed the expression of virally-delivered Ascl1, the endogenous Ascl1 gene, and Gadd45y (**Figure 3.8**). Since both endogenous Ascl1 and Gadd45y were induced in cells expressing Ascl1, we confirmed that transcription factors in the library should still be functional in our system.

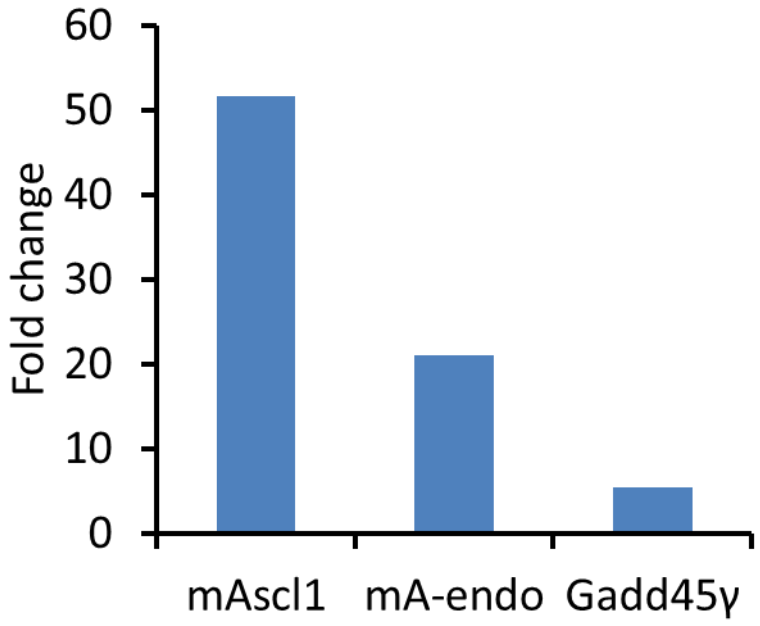


Figure 3.8 Exogenous expression of mAscl1 transcription factor in MEFs is able to induce expression of the endogenous Ascl1 (mA-endo) and a direct transcriptional target Gadd45γ by qRT-PCR. Fold change is calculated relative to un-transduced MEFs, normalized by expression of ActB housekeeping gene.

3.6 Addition of the Entire Transcription Factor Library to Bi-transgenic Mouse Embryonic Fibroblasts

To test the entire library of transcription factors, lentivirus was made with all factors combined in an equimolar mixture to produce virus of all factors simultaneously. The resulting lentivirus-containing supernatant was then added to the bi-transgenic MEFs at various concentrations or multiplicities of infectivity (MOIs) and cultured in neuronal medium. Since the viral transduction events are discrete and independent of each other, they infect the MEFs at random with approximately a Poisson distribution to capture a range of different combinations of library factors at random. Without addition of factors, no transgenic reporter fluorescence was detected

by flow cytometry (**Figure 3.9**). However, a small but significant population of tauGFP-positive cells were detectable by one week post-infection indicating the emergence of a neuronal subpopulation (**Figure 3.9**). By two weeks a more significant population of around 10% tauGFP-expressing cells was seen (**Figure 3.9**). A range of MOIs spanning from 2-20 factors per cell on average was tested without further increase in percent of cells expressing tauGFP and no significant drd1aRFP expression was seen in any experiments.

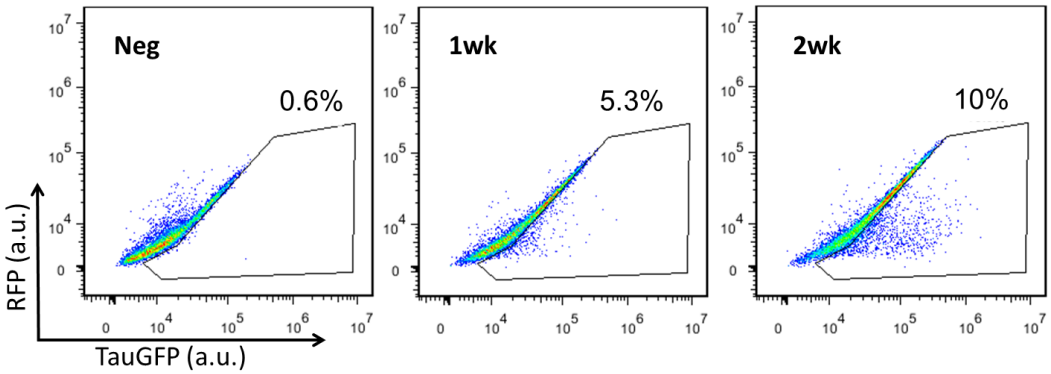


Figure 3.9 Transgenic E14.5 MEFs 1 week and 2 weeks post-tranduction with all library members results in a progressive increase in expression of the pan-neuronal tau-GFP transgene over time.

A second way to identify neuronal cells is by their morphology. MEFs have a flat, spread-out morphology which changes during neurogenesis to a morphology indicative of neurons which consists of a small, rounded cell body with long processes that extend multiple cell lengths from the cell soma. Different subtypes of neurons have slightly different specific characteristics that are occasionally able to visually differentiate them from each other. Many neurons have 1-3 main processes which

extend from the cell body. They are occasionally highly branched depending on whether the process is an axon or heavily arborized dendrites. MSNs have highly branched processes and are characterized by the presence of small spines along their dendrites. After adding all library factors to MEFs, some of the tauGFP-expressing cells also began to change their morphology closer to that of neurons (**Figure 3.10**) while the surrounding fibroblasts maintained their characteristic morphology and did not have detectable reporter expression. This indicates that we were successfully able to obtain newly reprogrammed neurons but not necessarily of the desired MSN subtype.

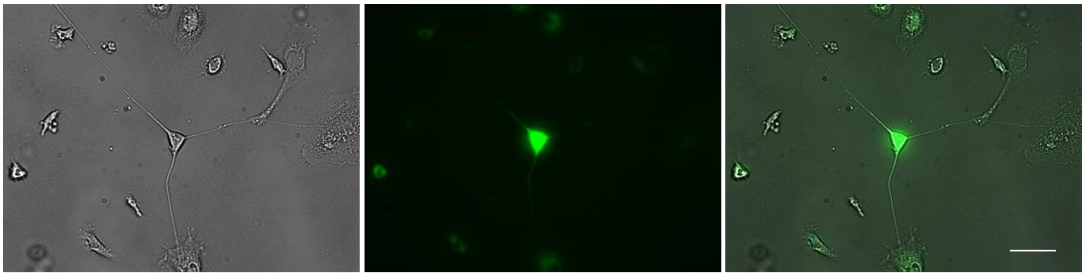


Figure 3.10 Microscopy of transgenic MEFs transduced with all library members 2 weeks post-infection. Changes in morphology and presence of tauEGFP reporter indicate positive neurogenesis. (scale bar = 100 μm)

3.7 Quantification of Integrated Factors in Sorted Subpopulations

As the last step of the iterative process of high-throughput reprogramming, sorted subpopulations of cells expressing the fluorescent reporters were tested for enrichment or depletion of library factors relative to cells negative for both reporters. RT-PCR was used for factor detection as opposed to quantitative PCR of sorted genomic DNA for increased sensitivity since one integrated factor in the genomic

DNA is capable of expressing a large number of RNA transcripts. Since no new population co-expressing both fluorescent reporters was found, RT-PCR was run on sorted populations with either increased tauGFP only (**Figure 3.11**) or increased drd1aRFP only (**Figure 3.12**). The results of the tauGFP RT-PCR showed a significant enrichment of know pro-neurogenic transcription factors including Ascl1, NeuroD1, and NEUROG2. Testing the tauGFP cells for pan-neuronal markers and those indicative of MSNs, we saw the highest increase in the MSN-specific DARPP-32 and the GABAergic marker GAD1. Pan-neuronal markers Tuj1 and MAP2 were moderately increased in the tauGFP subpopulation by about an order of magnitude. The subpopulation of cells with increased drd1aRFP fluorescence were also sorted by FACS and quantified by RT-PCR for transcription factor enrichment. Since these cells were not also positive for tauGFP expression, it is not surprising that pan-neuronal markers were not strongly affected. Tuj1, Ascl1, and MAP2 were not significantly enriched as expected for non-neuronal cells. However, we did see a significant enrichment in pro-MSN markers DARPP-32 and GAD1 as well as some factors associated with MSN development NEUROG2 and DLX6. Factors known to have an inhibitory effect on MSN development including NKX2.2 and LHX6 were significantly depleted. Taken together, these results indicate the feasibility of using this bi-transgenic MEF model for MSN reprogramming.

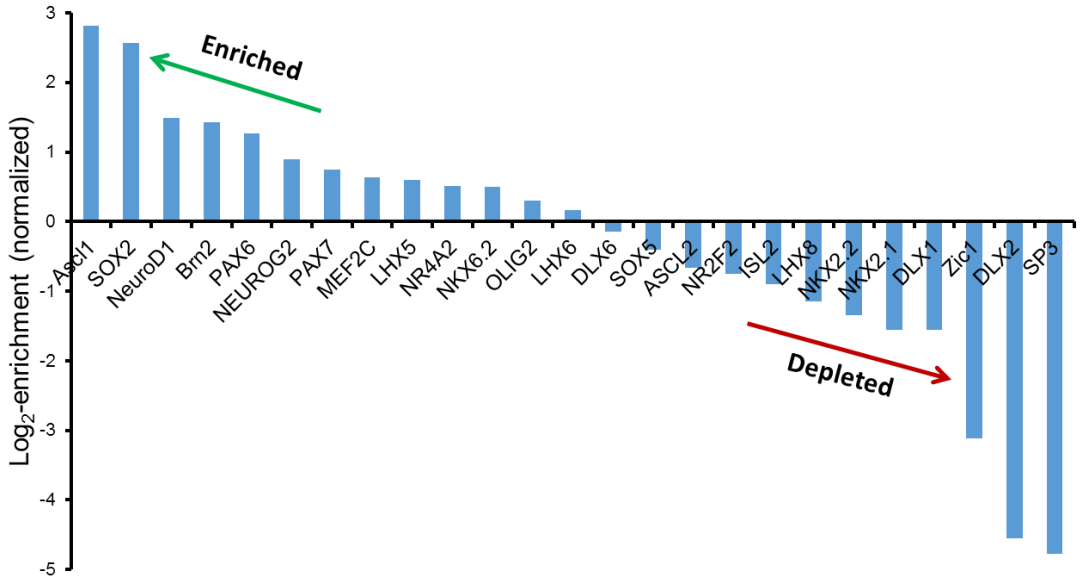


Figure 3.11 Fold-enrichment of TauEGFP+ library members by RT-PCR of FACS-sorted subpopulation of MEFs 1 week post-transduction with all factors at a multiplicity of infectivity of ~18 factors/cell. Enrichment is relative to MEFs sorted for no transgene expression.

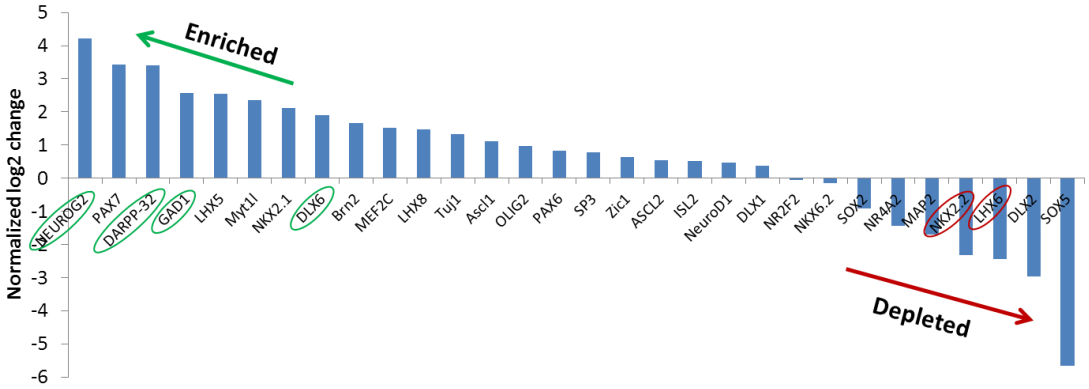


Figure 3.12 Fold-enrichment of Drd1aRFP+ library members by RT-PCR of FACS-sorted subpopulation of MEFs 3 weeks post-transduction with all factors at a multiplicity of infectivity of ~2 factors/cell. Sorted cells did not express TauEGFP reporter. Enrichment is relative to MEFs sorted for no transgene expression.

3.8 Secondary Sort of Library Members Using Restricted Subset of Transcription Factors

By comparing the results of the tauGFP sort (**Figure 3.11**) and the drd1aRFP sort (**Figure 3.12**), we identified a subset of factors enriched in both data sets. By repeating the reprogramming process using these fourteen factors, we hypothesized that we would see a greater fraction of cells expressing either of the transgenic reporters as well as a new subpopulation of cells expressing both reporters which would indicate the presence of positively reprogrammed MSN-like cells. After transducing MEFs with this smaller subset of factors and incubating them in neuronal medium for two weeks, the amount of cells positive for either reporter remained similar to the full library sort at around 3-5% and additionally, no new double-positive subpopulation of cells were observed. Cells that were positive for each individual reporter were sorted to check which factors were enriched or depleted and we found that all factors added were enriched in both subpopulations relative to the cells sorted with no reporter fluorescence (**Figure 3.13**). Finding all factors enriched is unsurprising since the experiment was set up such that factors found to stimulate expression of each reporter were specifically selected. Testing more values of MOI and potentially subcloning additional library factors may be required to successfully reprogram MEFs into MSN-like cells.

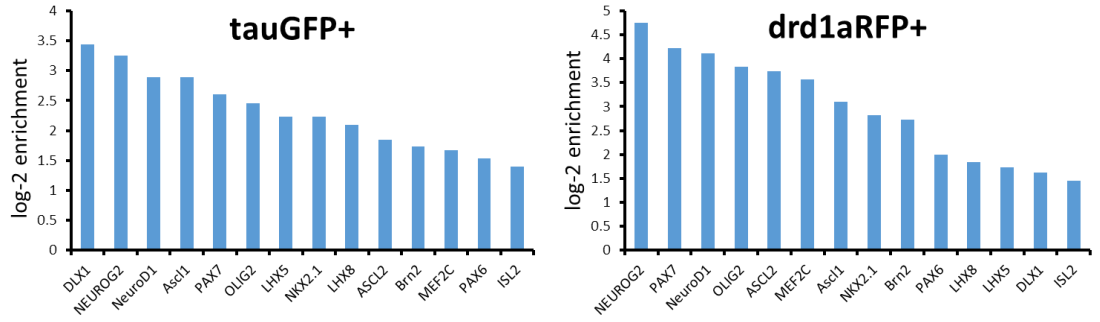


Figure 3.13 All TFs were enriched in sorted MEFs infected with a restricted subset of factors at a multiplicity of infectivity of 5 factors/cell for 2 weeks and sorted by FACS for either TauEGFP (left) or Drd1aRFP (right) expression. Sorted subpopulations accounted for 3-5% of infected cells and no population of MEFs was observed expressing both reporters.

3.9 Confirmation of Growth and Maintenance of Human Induced Pluripotent Stem Cells

Pluripotent stem cells are uniquely capable of differentiating into any cell type in the body. Typically this process unidirectionally follows a differentiation pathway involving the expression of particular transcription factors stimulated by spatiotemporal cues from the cell microenvironment to mediate this process. By forced expression of certain transcription factors, somatic fully-differentiated cells may be reverted to a pluripotent state to generate induced pluripotent stems cells (iPSCs) which are functionally similar to embryonic stem cells in that they are proliferative and pluripotent [47]. We decided to start with human iPSCs because of their pluripotent potential to differentiate into any cell type and try forced expression of transcription factors to directly differentiate iPSCs into specific subtypes of neurons rapidly and efficiently. The growth and maintenance of iPSCs is challenging because these cells require far more attention to feed and subculture them. They require pre-coating the growth surface for proper attachment with Matrigel and the cells are grown

as colonies because they need contact with neighboring iPSCs to maintain pluripotency. We have successfully demonstrated healthy iPSC growth with typical healthy morphology (**Figure 3.14**) as individual colonies which are mechanically dissociated for subculturing. To ensure that our iPSC cultures are being maintained in a pluripotent state, we used a live stain that detects the presence of alkaline phosphatase which is an indicator of pluripotency by fluorescence microscopy (**Figure 3.14**). All of our iPSC colonies were strongly positive for alkaline phosphatase indicating that our iPSC cultures are both healthy and pluripotent.

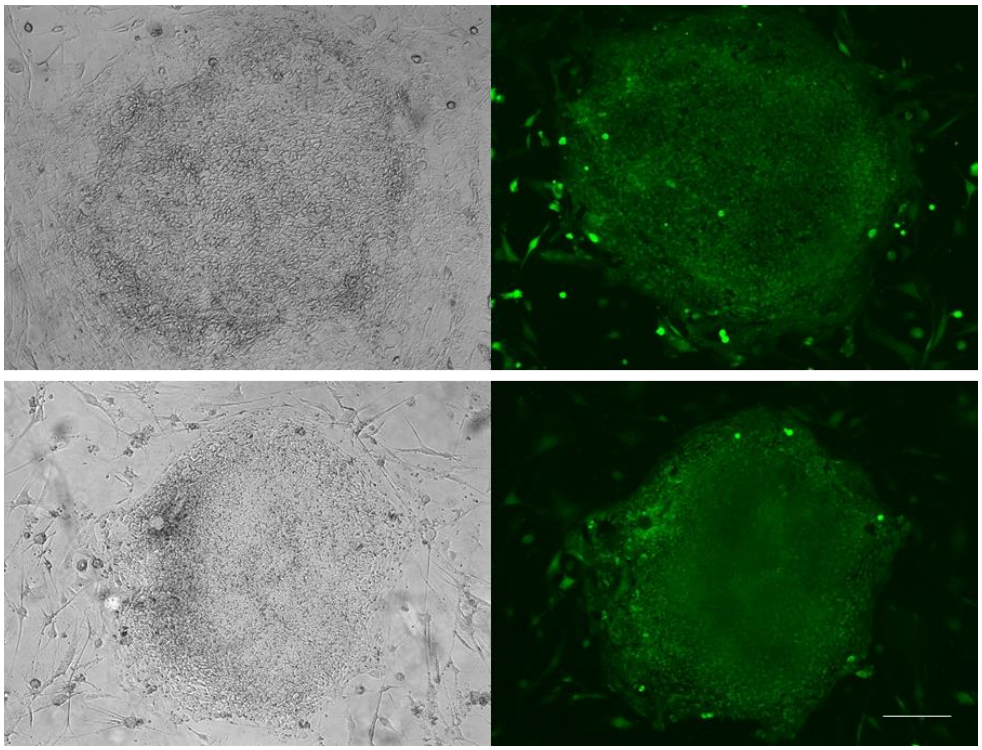


Figure 3.14 Growth of human induced pluripotent stem cells (iPSC) confirmed by brightfield (left) and corresponding Alkaline Phosphatase Live Stain (Life Technologies; right) 5 days after mechanical passaging. Strong alkaline phosphatase expression is an marker of pluripotency indicating healthy colony growth and maintenance of a pluripotent state in culture. Scale bar = 100µm.

3.10 Directed Differentiation of Induced Pluripotent Stem Cells by Exogenous Expression of Transcription Factors

For our studies with iPSCs, we chose to reduce the set of library factors down to 12 specific factors known to be involved in MSN development based upon a survey of literature related to MSN growth and differentiation to increase our chances of finding combinations of factors sufficient to induce MSN-like cells. Cultures of iPSCs were transduced with either no factors, the set of MSN-related factors, or MSN factors with 25% NeuroD1 and grown for over a week in neuronal culture medium before

fixing and immunolabeling for a neuronal marker Tuj1 and a MSN-specific marker DARPP-32 (**Figure 3.15**). NeuroD1 was added as a significant fraction of the viral load since it was shown previously that exogenous expression of either NeuroD1 of NeuroG2 alone was sufficient to induce rapid and efficient neurogenesis [48]. We hypothesized that MSN-related factors may be able to guide directed differentiation while NeuroD1 induced neurogenesis. Cultures in which no factors were added had mild Tuj1 expression since non-neuronal cells do express basal levels of Tuj1, but no neuronal morphologies were detected. Transduction of iPSCs with MSN-related cells did induce a significant amount of neurogenesis which indicates that at least some of the MSN-related factors were definitely pro-neuronal, but no expression of DARPP-32 was detected indicating that we did not generate any MSN-like cells. The addition of a high quantity of NeuroD1 lentivirus drastically increased the amount of neurogenesis as seen by increased Tuj1 labeling and increased elongated neuronal morphologies, but these cultures also did not have any detectable DARPP-32 expression. Since the *in vivo* differentiation and current *in vitro* differentiation protocols are a long process frequently lasting greater than a month, we hypothesize that longer incubation times or additional factors may be required to generate newly differentiated MSN-like cells.

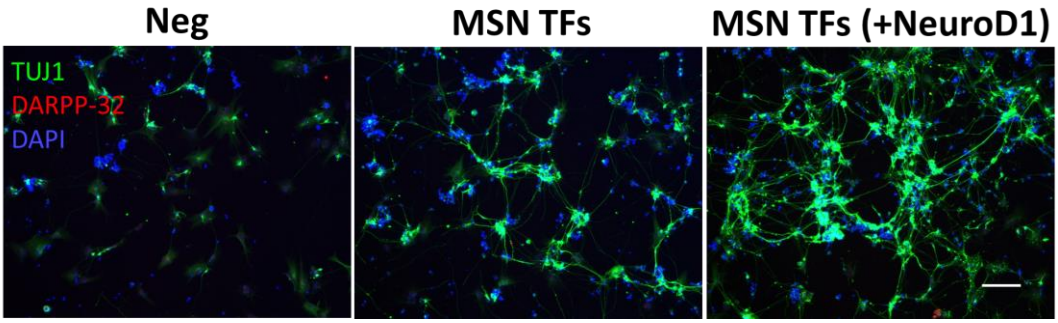


Figure 3.15 Attempted directed-differentiation of human iPSCs 10d post-transduction with a subset of MSN-related TFs in an equimolar ratio or MSN TFs with 25% NeuroD1 which is known to rapidly induce neurogenesis of pluripotent cells (Zhang Y et al. Neuron 2013). No expression of the MSN marker DARPP-32 was seen. (scale bar = 100 µm)

3.11 Discussion

In this study, we established the process and generated the resources for a high-throughput approach to neuronal reprogramming. This process will open up numerous possibilities to study the neuronal subtype-specific cell-autonomous effects in neurodegenerative diseases that have not been possible using traditional methods of differentiation with patterning morphogens. Guided small molecule differentiation of stem cells is a slow process that cannot easily be interrogated in a high-throughput manner since the small molecules are in solution and can interact with any cell in the culture. By using lentivirus at different multiplicities of infectivity, many different combinations of factors can be tested simultaneously in a single culture. Our library of transcription factors is the current largest set of reprogramming factors assembled for neuronal reprogramming which was intentionally targeted towards any subtype of neuron to make this process as broadly applicable as possible. We also demonstrated the ability to control expression of exogenously added factors as well as confirmed the efficacy of factors subcloned and transduced by measuring transcriptional targets. A

drawback of increasingly large libraries of factors is the astronomically large number of potential combinations that increase exponentially with factor set size. We have addressed this problem from a combinatorial perspective to maximize the efficiency of each experiment by guiding the selection of MOI for a given set of factors. By using known transcription factor interactions from literature, the total set of factors may be reduced to maximize the number of useful combinations that can be tested.

Our full attempt at library screening with bi-transgenic MEFs was able to successfully generate a subpopulation of neuronal cells. By RT-PCR we identified which library factors were enriched in the sorted reporter-expressing cells relative to reporter-negative cells and repeated the process. By the second iteration, it was found that all factors enriched in the first round were again enriched in the second round which is unsurprising since the tauGFP reporter identifies any subtype of neuron because tau is a pan-neuronal marker. We tested a small range of MOIs which did not result in a significant increase in the percent of reporter-expressing cells and did not result in expression of both reporters. Preliminary tests using small molecule chromatin-modifying inhibitors such as HDAC-inhibitors, methyltransferase inhibitors, DNA methyltransferase inhibitors, and SMAD inhibitors did not significantly increase reprogramming efficiency in our system. Additional combinatorial analyses we have performed suggests alternate values of MOI to test that may be performed in the future. Since there are known differentiation protocols of MSNs, we could also use RT-PCR or microarray to identify specifically which factors are present at different stages of differentiation in order to guide the selection of factors to include.

An alternate possibility for the lack of double reporter-positive cells is that MEFs are perhaps not epigenetically permissible to transdifferentiate into MSNs and so a full iPSC reprogramming may be required to reset some of these epigenetic marks. We therefore attempted a directed differentiation process by forced expression of transcription factors in iPSCs with subsets of transcription factors. In the time-scales tested, we did not observe the expression of a specific MSN markers DARPP-32, however most protocols for differentiation of stem cells into MSN-like cells require in excess of 30 days. Perhaps testing longer times and a broader range of MOIs may reveal more DARPP-32-positive cells. Another approach would be to augment the library factor size with more known MSN-related transcription factors such as Ebf1, Egr1, Gbx1/2, and Gsx1/2. In general, we need to better understand the specific markers, morphologies, and behaviors of particular subtypes of neurons to be able to properly identify and study them in culture. While much is known about different neurons in different brain pathways and some of the markers they tend to express, we need to take advantage of the increasingly available 'omics' methodologies to discriminate between neurons on a single-cell transcriptome and proteosome level to most effectively guide the development of cellular models.

Ultimately, we were able to obtain neurons in high yield from iPSCs which would be useful to establish disease-specific assays from patient-derived cells. By applying our high-throughput approach to different specific neuronal subtypes, these assays may be expanded to investigate the subtype-specific effects involved in these neurodegenerative diseases. Our preliminary results on some of these assays show promise for multiple points of comparison including neuronal primary and secondary apoptosis with AnnexinV and propidium iodide labeling, proteasome inhibition with

various destabilized GFP variants, lysosomal trafficking dysregulation with a lysosomal dye such as Lysotracker, and caspase activation with cleavable fluorescence dyes. These assays may be performed on neurons derived from disease patients relative to control patient tissue for different subtypes of neurons to better understand the cell-autonomous effects of Huntington's disease and Alzheimer's disease as well as many other different types of neurodegenerative diseases. In additional to better understanding the pathogenesis of these diseases, an appropriate cellular model that accurately captures the disease-related changes in the brain may be used to identify and test therapeutics targets as well as identify disease biomarkers to assess the efficacy of therapeutics which would be an invaluable tool for pharmaceutical development. This high-throughput platform has had all of the groundwork and proof-of-principle experiments performed in addition to the large pool of resources generated to be able to effectively derive any desired subtype of neuron from somatic patient tissue to study neurodegenerative disease in culture.

Chapter 4

STATISTICAL MODELING OF COOPERATIVE LIBRARY SCREENING FOR CELLULAR REPROGRAMMING

4.1 Introduction

Pooled library screening is an efficient way to test multiple combinations of factors that would otherwise be too inefficient to test exhaustively. As the number of library factors increases, the number of unique library members increases exponentially, which becomes impractical to screen exhaustively for large sets of factors. Pooled screening is especially important for situations in which multiple library elements may be simultaneously required or there are potentially cooperative effects between factors such as in cellular reprogramming or assembly of genetic circuits. Analysis of the combinatorial aspects of library screening may be useful for informing library construction and design of experiments.

A model example for application of library screening analysis is cellular reprogramming, in which one or more transcription factors are exogenously expressed in terminally differentiated somatic cells of one type to reprogram them into induced pluripotent stem cells, progenitor cells, or another type of somatic cells [27, 32]. Small pools of less than twenty factors are often screened to obtain a subset of factors sufficient to reprogram one cell type into another. The most common approach of exhaustive single and pairwise screening excludes many possible multiplexed combinations. Additionally, the number of factors ultimately required is not necessarily known initially and the potential inclusion of detrimental factors may not

be disregarded for processes as complex as transcriptional modification in the case of cellular reprogramming. While some studies have shown successful reprogramming between two specific cell types by exogenous expression of a single transcription factor [49-52], most studies involving transdifferentiation across germ layer boundaries require several [32, 46, 53]. For example, reprogramming fibroblasts into specific sub-types of neurons has often required several different factors and occasionally more, up to 7 transcription factors simultaneously [45].

As transcriptional control of cell fate is complex, the number and identity of the transcription factors sufficient to perform a particular reprogramming event is unclear. Therefore high-throughput screening of a pool of candidate transcription factors may yield the desired cell type more rapidly. The following analysis assumes that at least one instance each of some desired subset of library factors is sufficient to induce successful reprogramming. This analysis also accounts for the potential presence of some independent subset of detrimental factors that may either inhibit or misguide the correct reprogramming event. In the most lenient case, the detrimental subset is a null set and the minimal presence of the necessary factors is sufficient. In the most stringent case, all factors not required for reprogramming may inhibit the process such that all factors are important and selection of a proper concentration of viral particles is very important. Screening too few library members or poor average number of factors per cell may yield no successful library members, failing to provide any new information about the pooled library factors. By taking a naïve approach to the size of each set of desired and detrimental factors to consider all possible experimental situations, a large pool of factors may be screened more efficiently. The following analysis shows how a single screening experiment may be split into multiple

subdivisions with different parameters to take into account the potential sizes of these sets of desired and detrimental factors most effectively.

4.2 Derivation of Equations

For a pooled set of *n* factors applied to a population of cells to generate a library, we have derived the likelihood of obtaining a desired outcome defined as members containing at least one each of subset *J* of the *n* factors, but not potentially detrimental subset *Z* exclusive of subset *J*. For a given library member receiving *k* random factors selected in any order with potential repetition, there are n^k total combinations given by **Eqn. 4.1**:

$$n_1 \cdot n_2 \cdot n_3 \cdot \ldots \; n_k = n^k \quad (4.1)$$

The total combinations n^k may be expanded as a summation of combinations containing *i* unique factors from zero to *n* total factors. This will enumerate the number of ways to choose *i* unique factors from *n* possible factors multiplied by the number of surjections from a *k*-element set onto an *i*-element set. Effectively, the surjective function describes the number of ways to map *i* unique factors in *k* indistinguishable spaces where each of the *i* factors is represented at least once. The total expanded combinations may then be restricted to only combinations containing desired subset *J*, but not detrimental subset *Z*, and normalized by the total combinations n^k to obtain the likelihood. The likelihood, $P_k(n,j,z)$, of finding a particular combination of *k*-factors containing at least subset *J* but not subset *Z* may be simplified to a single summation (**Eqn. 4.2**, extended derivation in **Supplementary Info**).

$$P_k(n,j,z) = \frac{\left[\sum_{x=0}^{j} \binom{j}{x} \cdot (-1)^x (n-z-x)^k\right]}{n^k} \quad (4.2)$$

Additionally, for the most stringent case where all factors that do not belong in subset *J* fall in subset *Z* such that *j*+*z*=*n*, this equation simplifies as expected to the number of ways to select *j* unique factors at least once each (**Eqn. 4.3**):

$$P_k(n,j,z) = \frac{\left[\sum_{x=0}^{j}\binom{j}{x}\cdot(-1)^x(j-x)^k\right]}{n^k} = \frac{Surj(k,j)}{n^k} \quad (4.3)$$

For the application of discrete *n*-element libraries, we may assume a Poisson distribution of *k*-element containing cells. Any distribution may be used in place of a Poisson distribution depending on the biological system being interrogated such as situations in which every library member has exactly *k* elements. The total fraction of library members containing at least one each of subset *J* but not subset *Z* is found by summing the pairwise product of the combination likelihood $P_k(n,j,z)$ and the Poisson distribution $P_\lambda(k, MOI)$ over all *k* for a given multiplicity of infectivity (*MOI*, average value of *k* per library member)(**Eqn. 4.4**).

$$Total\ positive\ fraction = \sum^{k}[\,P_k(n,j,z)\cdot P_\lambda(k,MOI)]$$

$$= \sum^{k}[\,P_k(n,j,z)\cdot\frac{MOI^k e^{-MOI}}{k!}] \quad (4.4)$$

By maximizing this function, we find a value of *MOI* at which the highest fraction of positive members will be obtained. Directly related to the total positive fraction, the minimum number of library members to screen is the reciprocal of the total positive fraction, which is effectively the number of members to screen before one successful member. A desired minimum number of successful members follows as multiples of the minimum library size (**Eqn 4.5**).

$$Min\ library\ size = \frac{Desired\ successful\ members}{\sum^{k}[P_k(n,j,z)\cdot\frac{MOI^k e^{-MOI}}{k!}]} \quad (4.5)$$

4.3 Fraction of Successful Library Members is a Function of MOI

We have derived the likelihood of finding a particular desired subset from a total set of potential library factors for a given library member containing a specific number of factors. For example, consider a pool of five genes where two genes are sufficient to induce reprogramming, one gene has no effect, and there are two genes which may prevent the reprogramming event from occurring where the factors are initially indistinguishable from each other (**Figure 4.1**). The factors are applied to a population of cells to generate a cellular library in which different library members may have differing numbers of factors depending on the multiplicity of infectivity (*MOI*) which is the average number of factors per library member. For this analysis, we assume the number of library factors are Poisson distributed among library members which is used to describe the distribution of discrete events occurring independently over a defined interval of time such as viral infectivity [54, 55]. A positive library member is defined here as containing at least one instance of each element in the desired subset of factors and none of the factors from the potentially detrimental subset of factors (**Figure 4.1**). The likelihood of obtaining a positive library member in the members with a particular number of factors is then given by the likelihood equation (**Eqn. 4.2**). Using a Poisson distribution, the pairwise product of the fraction of members receiving specific numbers of factors and the likelihood of success for members with that many factors results in the overall fraction of successful library members in the entire library (**Eqn 4.4**). By maximizing this function to obtain maximum total successful fraction, an optimized MOI is obtained which is the ideal concentration of library factors to add to the population of cells (**Figure 4.1**). Once a subset of successfully reprogrammed cells is identified, they may be isolated by a variety of methods including immunolabeling and fluorescence-activated cell sorting

(FACS), laser-capture, or magnetic activated cell sorting (MACS) and probed by quantitative PCR or RNA-seq to determine which factors were enriched in the reprogramming process.

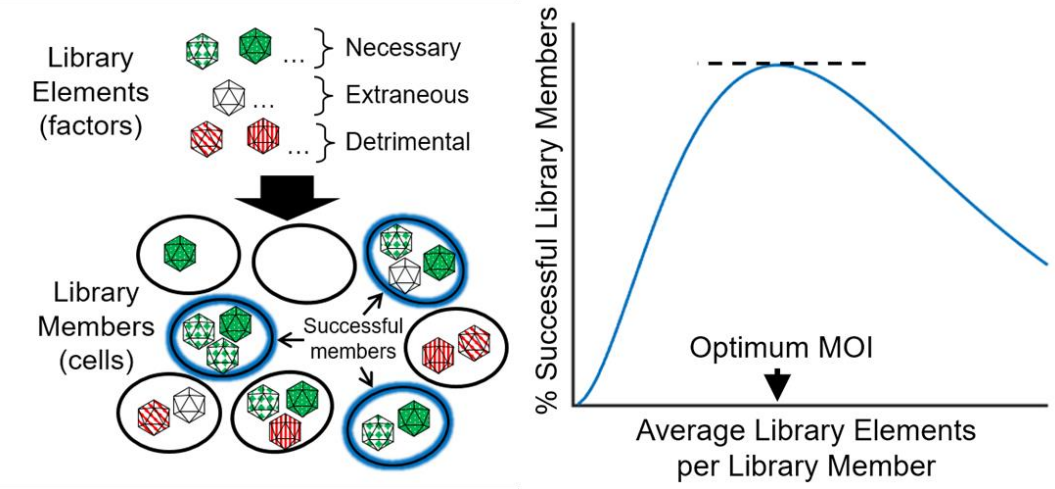


Figure 4.1 Schematic of cooperative library screening in which a mixture of factors that are either necessary, extraneous, or detrimental are added randomly to cells to generate a library. Successful library members are defined as members containing at least one each of necessary factors and none of the detrimental factors. The fraction of successful members is a function of the amount of factors added and there is a maximum successful fraction of an optimum multiplicity of infectivity (MOI).

The potential phase space of differing numbers of desired and detrimental factors is bounded by the total number of unique library factors since the desired and detrimental factors are independent of each other. This phase space may be conveniently visualized similarly to a ternary diagram (**Figure 4.2**). Each pair of potential combinations of desired and detrimental factors has its own total positive fraction of library members as a function of *MOI* with a corresponding optimal MOI

for each pair. By surveying published reprogramming studies for their total number of factors, the number of ultimately necessary factors, and estimation of the maximum number of detrimental factors, we found that most of this phase space of necessary and detrimental fractions is possible. The corresponding optimum *MOI* for each maximum positive fraction of library members provides a useful estimate of which *MOIs* to screen given prior knowledge of the relative number of desirable and detrimental factors to be screened (**Figure 4.2**). Situations in which many factors will be required simultaneously, such that necessary factors far outnumber detrimental factors, suggests that high values of *MOI* will be more effective for pooled screening. Conversely, pools of library factors in which it is known that many of the factors will be detrimental may require screening much lower *MOIs* to maximize the potential successful fraction of library members.

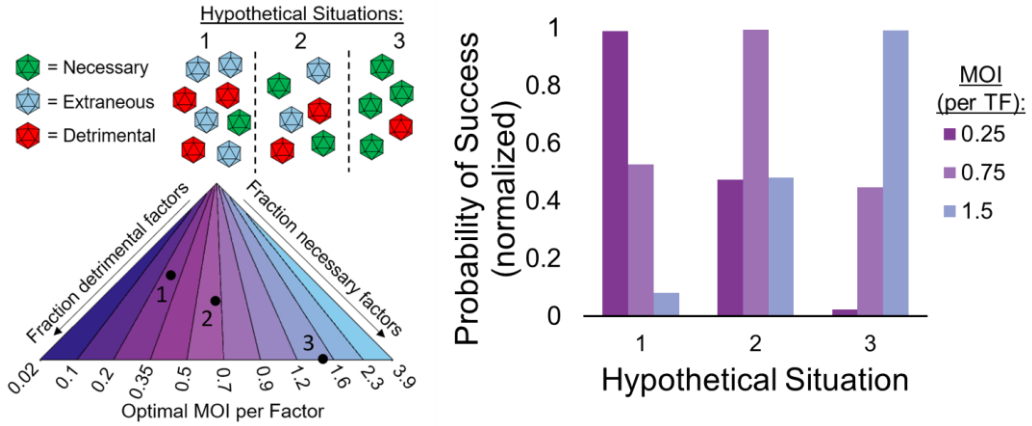


Figure 4.2 Hypothetical cases with varying fractions of necessary and detrimental factors to illustrate the significant effect of MOI choice on the probability of finding successful library members. The fraction of necessary and detrimental factors dictates a phase space (left) in which a wide range of MOI will provide the maximum fraction of successful library members. For three different choices of MOI, the probability of success normalized by the maximum successful fraction can vary greatly (right).

If we take a few hypothetical combinations of necessary, extraneous, and detrimental factors, we can quickly identify where on the ternary diagram of necessary and detrimental factors each hypothetical case falls (**Figure 4.2**). Each of those cases corresponds to a specific value of MOI at which the highest possible fraction of successful library members may be found. Since we may not know from the start the identity of which factors may be necessary or detrimental, a typical experimental approach may be to select a single MOI to test. However, for particular values of MOI to screen each of these hypothetical cases, we found that the probability of success is strongly dependent on choice of MOI (**Figure 4.2**). This means that random choice of MOI will likely result in a failed screening result since some of these combinations of necessary and detrimental factors may already have a very low maximum value of successful library members.

4.4 Generalized Approach to Library Screening for Optimization of Experimental Parameters

An important consideration for pooled library screening is that the number of effective factors or detrimental factors is not necessarily known from the start. In this case, a balanced *MOI* that considers the possibility of any number of desired and detrimental factors may be most effective. Relative success of the potential maximum positive library fraction may be defined as the ratio of the actual positive fraction for a particular *MOI* given by (**Eqn. 4.4**) to the maximum possible positive fraction. By maximizing the lowest point of the relative success, a balanced *MOI* is obtained (**Figure 4.3**). The corresponding relative success of the optimum value of MOI for a single experimental subset show that a single *MOI* does not effectively account for situations in which the number of the desired or detrimental factors is large as

evidences by regions of the plot where the coverage is close to zero (**Figure 4.3**). By splitting the screening experiment into multiple samples with different *MOIs*, the coverage of potential maximum positive library fraction is greatly improved (**Figure 4.3**). As an alternative way to show how splitting an experiment into multiple subsets is able to increase the probability of identifying positive library members, we can look at the lowest value of relative success for a given number of experimental subsets (**Figure 4.3**). The minimum relative success rapidly increases which indicates improved coverage of the experimental phase space to optimize screening efficiency.

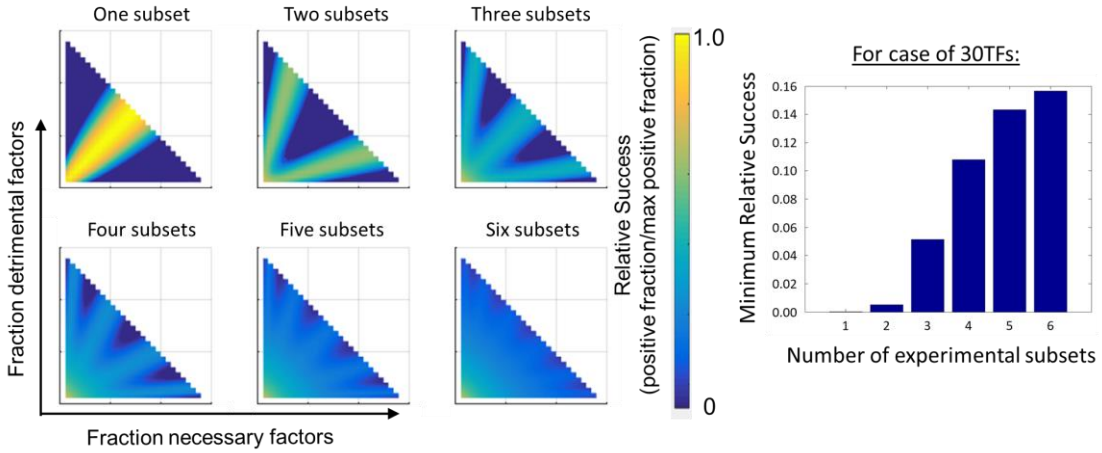


Figure 4.3 The fraction of necessary and detrimental factors for a given screening experiment is unknown, but splitting experiments into smaller subsets can better cover the potential phase space. The graphs on the left show the ratio of positive fraction of library members to the maximum fraction of library members for a given number of experimental subsets optimized on the lowest point of the entire phase space of necessary and detrimental fractions. The graph on the right is an example with 30 TFs of the lowest value of relative success for one to six experimental subsets.

Additionally, the optimum values of *MOI* for multiple *MOI* screening experiments also converge to specific values of MOI per unique library factor (**Figure 4.4**). By splitting a library screening experiment into multiple equal subdivisions with different MOIs, the phase space of different potential numbers of desirable and detrimental factors is better covered, improving the likelihood of experimental success. This analysis also indicates that continually increasing the number of MOIs does not continually increase the efficiency of library screening. This is most evident in the examples where higher numbers of MOI are being screened for low total unique library factors where the different values of MOI converge upon common values (**Figure 4.4**). Since the values of MOI per unique library factors are normalized by the total number of unique library factors, these same values may be applied to large pools of factors to screen.

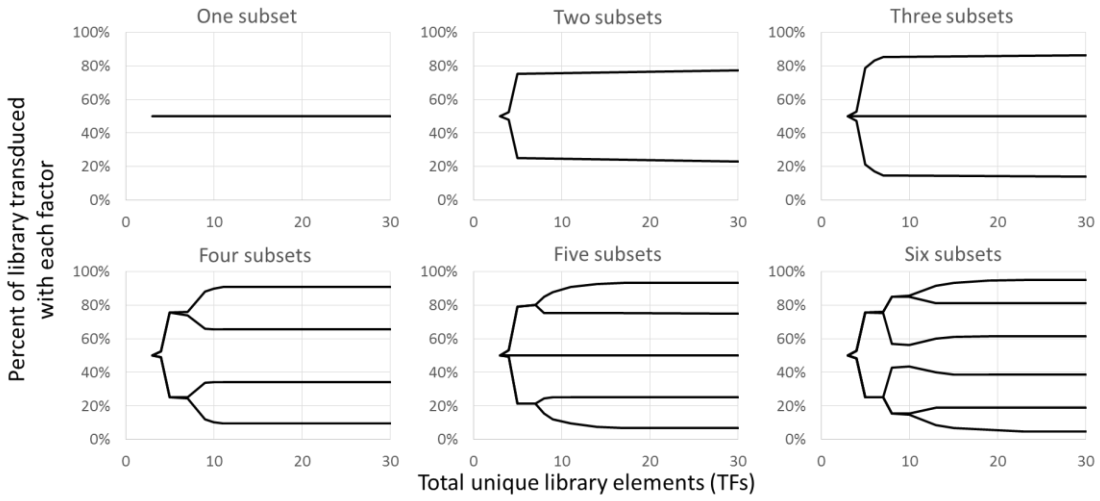


Figure 4.4 Experimental values of viral titers suggested by optimizing different numbers of experimental subsets across the entire phase space of necessary and detrimental fractions for different total numbers of factors. Viral titers are given of percent of library transduced with each factor which may be easily converted to MOI by the cumulative Poisson distribution.

4.5 Discussion

We have demonstrated a generalized approach to cooperative library screening in which multiple elements may be simultaneously required and some elements may be detrimental to obtaining a successful library member. Our analysis yielded several important points about pooled library screening. First, the choice of MOI or average number of factors per library member is very important since the relative success is highly sensitive to value of MOI. If the total number of necessary and detrimental factors is known, the optimum value of MOI may be quickly calculated or estimated from the data presented here (**Figure 4.2**). If the fraction of necessary or detrimental factors in a large set to be screened is unknown, we have demonstrated a method to optimize the choice of MOI or multiple MOIs to most efficiently test the experimental

phase space. From this analysis, we have reported the values of optimum MOIs which have been normalized to the total number of unique factors to make this analysis generally applicable to any library screening.

Another consideration of this work is that the input of *a priori* knowledge of a given system, such as an estimated number of desired factors or the total number of factors to be screened, restricts the phase space of potential combinations and informs both the minimum library size to be screened and the optimal *MOI* of factors to be applied which may be calculated from the equations derived here. For situations in which there is no prior knowledge of the number of desired and detrimental factors, we have calculated the optimal ratios of MOI to screen for a single or multiple MOIs to maximize the likelihood of experimental success. The analysis and equations presented here may assist in the experimental design and generation of libraries of transcription factors for cellular reprogramming. Similar analyses may be used to also screen other types of libraries in which cooperativity may be an important factor such as libraries of siRNA, CRISPR sgRNAs, or synthetic gene networks.

Chapter 5

TAU STRAINS AND THEIR PROPAGATION IN EXPERIMENTAL DISEASE MODELS

5.1 Introduction

Misfolding of the microtubule-associated protein tau into deposits of neurofibrillary tangles is a distinguishing feature of a group of neurodegenerative diseases collectively referred to as tauopathies. Tauopathies are movement disorders and dementias that result in a progressive loss of neurons and cognitive impairment and include Alzheimer's disease (AD), Pick's disease, corticobasal degeneration, progressive supranuclear palsy, agyrophilic grain disease, chronic traumatic encephalopathy, and tangle-only dementia, among others [56]. The fibrillization and accumulation of tau is common among these diseases, however the histopathological progression and clinical presentation differs between them [57], suggesting the possibility of distinct strains of tau fibrils with unique properties that would account for their stereotypic characteristics. Currently there are no reliable methods to diagnose patients with many of these diseases [58] and no therapeutics available to halt or even slow disease progression.

Figure 5.1 Splice variant isoforms of tau. The six different isoforms of tau generated by alternative splicing of exons 2 and 3 for the N-terminal region, which encode the N1 and N2 segments of the protein (labeled here in orange and red, respectively), and of exon 10 in the microtubule binding region, which encodes the second repeat sequence (R2, shown in green) [59]. Terminology used for describing the isoforms varies among researchers; the more descriptive names listed to the left of each isoform are used throughout this chapter. Historical names are listed to the right, along with the number of residues present in each isoform. The microtubule binding domain is composed of 3 or 4 repeat sequences (blue and green), which have significant homology but are not identical. These domains form the core of tau fibrils, and when expressed recombinantly as truncated protein products, they are referred to as 3R and 4R; alternately they are sometime identified as K19 and K18, respectively.

Native tau mainly functions by binding to and stabilizing neuronal microtubules to facilitate cellular structure and intracellular transport of biomolecules [60, 61] though tau may also play a role in many cellular pathways [62, 63]. In the unbound state, tau is an inherently disordered, highly soluble protein [64, 65]. The microtubule-binding region near the C-terminus reversibly associates with axonal microtubules in a phosphorylation-dependent manner [66]. There are 6 different splice variants or isoforms of tau (**Figure 5.1**) that are expressed in humans and their relative proportions in neurofibrillary tangles often varies among different tauopathies [67].

Alternative splicing produces an N-terminal region that may have zero, one, or two variable polypeptide segments while the microtubule-binding region may contain either three or four repeat sequences with moderate homology. Roughly equal amounts of 3R and 4R isoforms are found in fibrils isolated from the brains of patients with AD [68], chronic traumatic encephalopathy [69], and tangle-only dementia [70]. The 4R isoforms are predominant in fibrils found in progressive supranuclear palsy [71], corticobasal degeneration [72], and agyrophilic grain disease brain tissue [73, 74], while Pick's disease inclusions are predominantly composed of 3R isoforms [75].

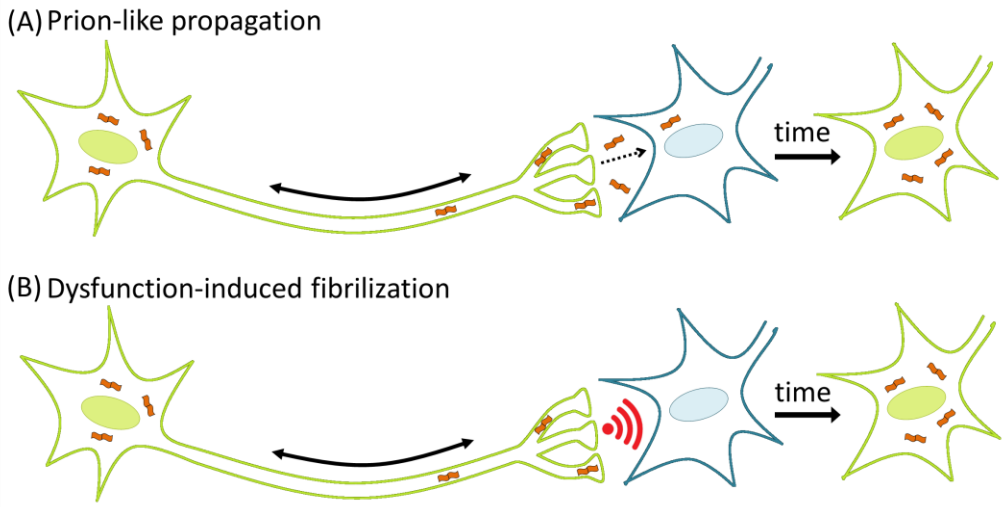


Figure 5.2 Competing hypotheses for the causative mechanism of disease pathogenesis. (A) The "prion-like" hypothesis of tauopathies suggests that strains of tau fibrils pass from dysfunctioning neurons (green) into healthy neurons (blue) and recruit native tau, resulting in dysfunction of the healthy neuron over time. (B) An alternate hypothesis is that dysfunctioning neurons induce a state of stress in healthy neurons through signaling, causing tau fibrillization as a downstream effect, e.g. through disruption of protein homeostasis or induction of apoptosis [8]; this hypothesis is akin to the amyloid cascade hypothesis in AD. The observation that tau strains are conserved throughout the brain as they spread supports the first hypothesis.

Prions are infectious agents composed of alternatively folded protein with the ability to transmit disease through recruitment and conversion of normally folded protein into the alternatively folded conformation [76]; additionally, structurally distinct prion strains, composed of protein with identical polypeptide sequences, induce distinct phenotypic changes in host animals [77, 78]. The mechanism by which initial spontaneous misfolding of tau occurs is still not fully understood, however increasing evidence suggests that propagation of fibrillar tau may be "prion-like" in its progression not only at the molecular level, but at the cellular level as well (**Figure 5.2**) [79-82]. Neurodegenerative diseases in general appear to follow a pattern of neuronal connectivity in their progression [83-85] which may signify the propagation of a pathogenic species. Thus, the prion-like propagation of tau strains may provide an opportunity for developing therapeutics to prevent their intercellular spread.

5.2 Prion-Like Phenomenon in Tauopathies

Shortly after the discovery of prions, the most common tauopathy, Alzheimer's disease, was hypothesized to share a common underlying mechanism for pathogenesis [86]. While most tauopathies are sporadic in etiology [87], genetic forms of these diseases caused by destabilizing point mutations have been discovered (**Figure 5.3**). The fact that such mutations are sufficient to initiate neurodegeneration suggests that tau fibrillization can initiate disease pathogenesis [88-90]. The first extensive classification of the histopathological stages of AD demonstrated a distinct pattern of the spread of tau tangles in the brain along neuroanatomical connections [91], suggesting either protein-based disease progression or progressive cellular dysfunction in series of neurons (**Figure 5.2**). The hierarchical spread of tau pathology has also been documented in other tauopathies [92]. Templated conformational propagation of

recombinant tau into structurally distinct strains [14, 93], transynaptic spread of tau fibrils and pathology in mice [94, 95], and the spread of tau fibrils following inoculation of exogenous fibrils [96, 97] all provide supporting evidence for the prion-like spread of tauopathies (**Table 5.1**). Serial propagation of strains has also been demonstrated in mice [98, 99] analogous to that of prion inoculation [100].

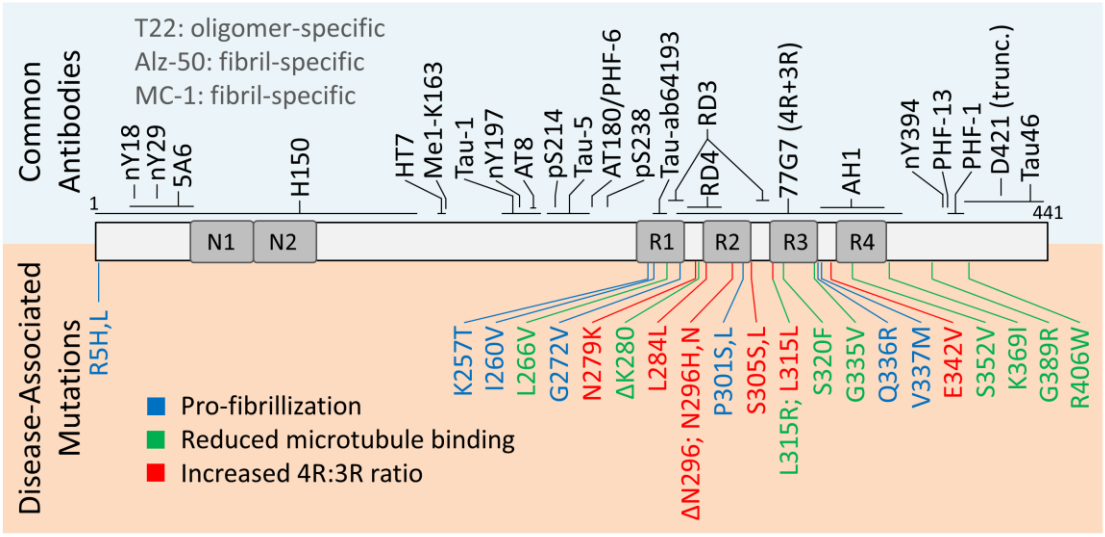


Figure 5.3 Commonly used tau antibodies and disease-associated mutations. (Top) Antibodies are mapped to the locations of antigens used to raise them, or to their epitopes if known. Antibodies T22, Alz-50, and MC1 are conformation-specific as indicated. The antibodies RD3 and RD4 are specific to the isoforms they are named after. Several antibodies are specific for post-translational modification, including phosphorylation, nitration, and methylation. (Bottom) Disease-associated mutations are mapped to their locations on the gene. Mutations are associated with increasing fibril formation (blue), reducing the binding affinity for microtubule binding affinity (green), or increasing the 4R:3R splicing ratio (red). Some mutations such as L315L are silent gene mutations that effect tau mRNA splicing (Bottom updated and adapted from [90]).

One distinguishing feature of *bona fide* prions is their transmissibility from one individual to another. An epidemiological study of the prevalence of several neurological diseases associated with protein misfolding in patients receiving injections of human growth hormone isolated from cadaver pituitary glands concluded that, with the exception of prion diseases, such diseases are not iatrogenically transmitted between patients in this fashion [101]. However, intraperitoneal injection of tau fibrils induced tauopathy in the brains of mice in another study [102]. Regardless of whether tauopathies are transmissible, such diseases do share key characteristics with prions at the molecular and cellular level.

Model System	**Evidence of strain-like behavior**	**References**
Biochemical	Synthetic tau strains faithfully propagated using recombinant protein	[93]
	Conformational properties of both mouse and human brain-derived fibrils faithfully propagated using recombinant protein	[14, 103]
Cell Culture	Synthetic strains form unique puncta morphologies and retain biochemical properties	[99]
	Human brain-derived fibrils form unique puncta morphologies	[99]
Transgenic mice	Strains have differential patterns of neuropathology, including brain regions and cell types affected	[98, 99, 104]
	Tauopathy-derived strains exhibit histopathological features of their human counterparts when passaged to mice	[98, 104]
	Fibrils maintained structural and biochemical features during three serial passages in mice	[99]

Table 5.1 Summary of evidence for prion-like propagation of tau strains

5.3 Formation of Synthetic Tau Strains from Recombinant Protein

Analogous to the diversity of conformational strains found in prion disease [77, 78, 105-108] fibrils isolated from brains of patients diagnosed with different tauopathies also exhibit a range of conformations (**Figure 5.4A**) [109-112]. Tau protein natively exists as an inherently disorganized structure consisting mostly of random coil, adopting a range of monomeric structures over a broad energy potential landscape [61, 113, 114]. The energy potential landscapes of amyloidogenic proteins contain intermediate states which bind to amyloids, leading to elongation of fibrils [115, 116]. This conversion is accompanied by a loss of random coil content and an increase in beta-sheet content [65, 115, 116].

Full-length tau produced recombinantly forms synthetic fibrils [117, 118], a process which is induced by polyanionic cofactors including sulfated glucosaminoglycans (e.g. heparin), RNA, or free fatty acids [119-122]. Investigation of the protein regions responsible for driving synthetic fibril formation indicated that the third repeat domain and adjacent sequences are sufficient to form fibrils [123]. In the presence of polyanionic inducers, the mechanism of fibril formation follows nucleated monomer-addition kinetics [124-127]. Such studies have provided insight into the molecular mechanism leading to spontaneous formation and stabilization of fibrils in the early pathogenesis of tauopathies.

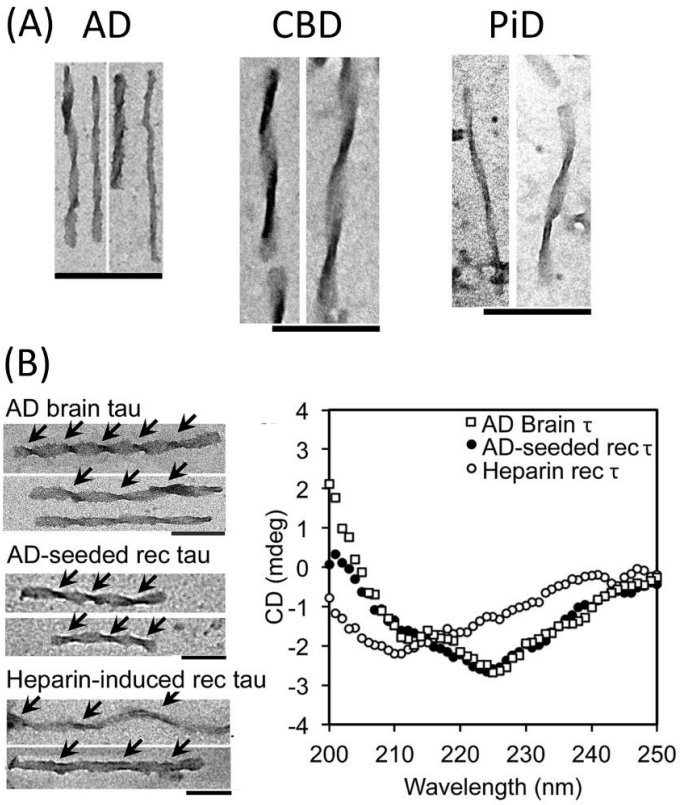


Figure 5.4 Tau strains and their propagation in cell-free systems without post-translational modification or cofactors. (A) Electron micrographs of the distinct structures of tau fibrils isolated from AD, corticobasal degeneration (CBD), and Pick's disease (PiD) brain homogenates. (B) Monomeric recombinant protein seeded with tau fibrils isolated from AD brain homogenates retains the structural features of the parental seed as assessed by electron microscopy and circular dichroism. For comparison, recombinant tau protein induced to form fibrils by heparin have distinct features. Scale bars represent 200 nm in (A) and 100 nm in (B); arrows indicate period of twist; images in (B) reprinted with permission from [14], copyright 2013 American Chemical Society.

Distinct fibril strains have been shown to propagate by sequestering the same recombinant tau substrate. Two different fibril conformations formed from wild-type and mutant tau in the presence of an inducer were subsequently used to seed

monomeric wild-type tau; the tau structures formed maintained the structural characteristics of the original seeds [93]. Other studies have shown that two different induced fibrils from 4R tau isoforms had distinct properties, and only one of the conformations was able to seed a 3R tau substrate [128, 129]. In addition to potential roles played by genetic polymorphisms and splice variants, thiol-disulfide interactions may contribute to fibril diversity [130].

5.4 Seeded Amplification of Brain-Derived Tau Strains

While polyanionic fibril induction has enabled insights into disease processes using recombinant protein, the underlying mechanism of fibril formation in the presence of such inducers appears to follow a different fibrillization mechanism than most amyloidogenic proteins [131], and the fibril strains formed are both structurally and biochemically different from their brain-derived counterparts (**Figure 5.4B**) [14, 103]. In contrast to these synthetic tau strains, full-length wild-type tau seeded by fibrils isolated from AD brain without heparin faithfully inherited the seed properties including fibril dimensions and secondary structure, independent of phosphorylation (**Figure 5.4B**) [14]. Hyper-phosphorylated tau has a reduced ability to stabilize microtubules so it has been implicated in tauopathy pathogenesis, though its role in fibril formation is a subject of debate [14, 132-134].

Additionally, recombinant tau seeded by P301S mouse brain-derived fibrils had the seeding potency and biochemical characteristics of the parent brain-derived seeds with and without hyper-phosphorylation [103] confirming the potential for tau strains to template conformational change independently of post-translational modifications [14].

5.5 Tau Fibril Uptake and Strain Propagation in Cell Culture

Cell culture models have been used to study the mechanisms by which tau fibrils become internalized, propagate, and transfer between cells. Synthetic tau fibrils, formed by misfolding tau in the presence of polyanionic cofactors, can be directly taken up by cells, seeding fibrillization of endogenous tau [135-137]. AD brain-derived tau also seeded intracellular fibrillization in HEK cells and SH-SY5Y cells [138]. Tau fibrils enter primary mouse neurons without lipid-based protein delivery reagents and stimulate conversion of endogenously expressed tau [81, 139-141].

Tau fibril strains also propagate in cell culture. Fibrils of 4R tau induced several biochemically and morphologically distinct synthetic tau strains in HEK cells [99]. These synthetic strains propagate in HEK cells as the cells divide and infect naïve HEK cells while maintaining the cellular and molecular properties characteristic of each strain. Tau fibrils purified from AD, agyrophilic grain disease, corticobasal degeneration, Pick's disease, and progressive supranuclear palsy brains elicited morphologically distinct patterns of aggregation in HEK cells [99]. Finally, tau fibers purified from end-stage P301S mouse brain seed conversion of tau in HEK293 cells about fifty times more efficiently than biochemically distinct, synthetic fibrils composed of recombinant tau with the same sequence [103].

5.6 Transmission of Tauopathies in Mice

The tendency of tau fibrils to be taken up by neurons and recruit endogenous tau into newly formed fibrils has also been demonstrated in mouse models of tauopathies. One mouse model expresses P301S mutant full-length human tau and begins to spontaneously develop tau tangles and neurodegeneration [142]. Intracerebral injection of P301S mouse brain homogenate into ALZ17 mice, a mouse

model expressing a wild-type human tau [143], seeded fibrillization of tau in ALZ17 mouse brains [96]. Since the ALZ17 mice do not develop tauopathy in their lifetime, the P301S mouse brain homogenate was inferred to have contained a transmissible species. Likewise, an injection of end-stage P301S brain homogenate into presymptomatic P301S mice accelerated fibrillization before the P301S mice normally develop tauopathy [144].

The induction of tau pathology in mice by synthetic fibrils made from recombinantly expressed tau [97] further supports the protein-only hypothesis. A larger dose of synthetic fibrils made from P301L 4R tau seeded tau fibrillization in mice expressing full-length P301L tau and additionally demonstrated selective neuronal loss in the hippocampus relative to wild-type mice injected with the same quantity of induced fibrils [145]. In addition to the structural differences between fibrils from AD brains and synthetic fibrils induced by heparin [14], brain-derived and synthetic fibrils have different seeding potency in mouse models*:* enhanced seeding of P301S brain homogenate relative to heparin-induced fibrils was observed [103]; likewise, AD brain-derived seeds fibrillized tau in wild-type mouse brains more potently than heparin-induced fibrils [146].

Mice expressing mutant tau only in the entorhinal cortex induced tau fibrillization in distal, but synaptically connected regions of the brain [94, 95]. Exogenously administered fibrils consistently induce a time- and dose-dependent pattern of transynaptic spread in mice [97, 99, 144, 145]. Patient brain-homogenates from six different tauopathies seeded tau pathology similar to their human counterparts in mice expressing wild-type 2N4R human tau [98] except for Pick's disease in which tau fibrils were confined to the injection site. Pick's disease is a

predominantly 3R isoform tauopathy, so there may be a seeding barrier between 3R and 4R tau isoforms [129, 136]. Brain homogenate from corticobasal degeneration and AD patients also induced similar histological progression in mice expressing full-length P301S tau; the number of hippocampal neurons decreased significantly in the AD cohort, but not in the corticobasal degeneration mice [104]. Two synthetic strains successfully propagated biochemical features after multiple sequential inoculations in mice [99]. Tau fibrillization was also propagated through serial inoculation of P301S mouse brain homogenates in ALZ17 mice, and brain homogenate from tangle-only dementia and agyrophilic grain disease patients in wild-type mice[98], although it was not clear whether the specific patterns of tau fibrillization were faithfully propagated as well. The intercellular transfer of tau fibrils in the brain and potential alternate routes of administration indicates that antibody therapies or vaccination may be an effective approach to treat or prevent tauopathies [147]; some groups have demonstrated success with antibodies capable of attenuating disease pathology [148-150].

5.7 Mechanisms Implicated in Cell-to-Cell Transfer of Tau Fibrils

Unlike disease proteins that are trafficked to the cell surface like prions or Aβ, tau is normally located in the cytosol and tau fibrils localize to the perinuclear space or somatodenritic compartment [151]. Monomeric tau is released into the interstitial space related to neuronal activity [152, 153] and is present in CSF [154]. It is unclear whether the release of pathogenic tau from neurons is related to the normal pathway of physiological release or unconventional secretory mechanisms [155, 156].

Experimentally, tau fibrils have been demonstrated to transfer between cells in cell culture [135, 157, 158] and transynaptically *in vivo* [94, 95, 158, 159].

Transynaptic spread in mouse models and both anterograde and retrograde trafficking of misfolded tau in neurons [141] strongly indicates the importance of synaptic connectivity in tauopathy pathogenesis. Alternatively, tau may be released into the interstitial space by the death of neurons containing tau fibrils. However, tau spreads in a hierarchical pattern following synaptic connectivity to distal regions of the brain [91, 94-98]; moreover tau levels do not appear to correlate with lactate dehydrogenase or tubulin release [160, 161]. Therefore it is unclear whether cell death is a dominant mechanism of intercellular transfer.

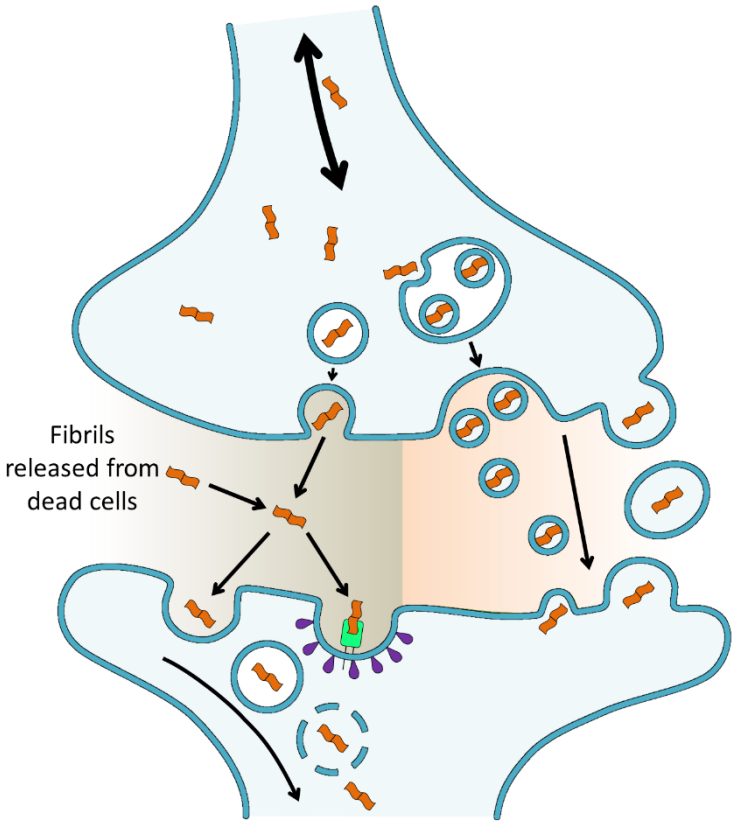


Figure 5.5 Hypotheses for the intercellular transfer of tau seeding species. Tau fibrils (orange) are transported both retrogradely and anterogradely along axons, as indicated by the large double-headed arrow at top, allowing fibrils to traverse neurons to presynaptic and postsynaptic termini [141]. Fibrillar tau has been observed outside of the cell in free form or in extracellular vesicles [155]. Shown to the left, naked fibrils may enter the extracellular space by release from dying cells or vesicular release by exocytosis potentially related to synaptic activity. Free fibrils may then be taken up by fluid-phase endocytosis or receptor-mediated endocytosis (receptor and vesicular coat proteins indicated in green and purple, respectively). Endosomal tau must then escape into the cytosol before it can recruit native tau into fibrils. Shown to the right, membrane-associated tau originates from exosomal/microvesicular packaging and release [162]. Once the membrane-associated tau reaches a post-synaptic terminal, it may be taken up by exosome/microvesicle fusion with the cell membrane.

The detailed steps by which pathogenic uptake of tau fibrils occurs are also not fully understood, although several potential mechanisms have been identified (**Figure 5.5**). Empirically, synthetic and brain-derived tau fibrils enter the cytosol of immortalized cell lines [135, 137, 138] and neurons in culture [139, 141, 163], and in mouse brains [96-98]; uptake also appears to be related to synaptic connections [91, 94, 95]. Lipid-based transfection reagents are used to promote fibril uptake [136, 137, 163]; however studies specifically addressing the direct uptake of free tau fibrils in multiple cell types agree that a dominant mechanism is macropinocytosis [138, 140, 141]. Macropinocytosis is an actin-driven endocytic pathway in which the cell membrane encapsulates and pinches off a large intracellular vesicle containing extracellular fluid and local biomolecules [164]. Early work identifying markers of macropinocytosis seemed to suggest non-specific uptake [135, 138, 141] however, a recent study showed tau fibrils actively stimulate macropinocytosis in a dose-dependent manner [140]. Once internalized into intracellular membrane vesicles, the endosomal escape of tau fibrils must occur before they can interact with endogenous tau. Heparin sulfate proteoglycans have been identified in neuritic plaques [165] however it was recently demonstrated that blocking cell surface heparin sulfate proteoglycan interactions strongly inhibits fibril uptake in culture and in mice [140] similar to the heparin-mediated pathways by which some bacteria and viruses enter cells [166]. Exosomal membrane fusion has also been suggested as a possible uptake mechanism [162]. The intercellular transfer of tau may also occur by a combination of these mechanisms [167], although it is not yet clear which is dominant in disease progression.

5.8 Role of Tau Fibers in Neuronal Dysfunction and Death

The contribution of the tau fibrillization process to neurodegeneration is unclear, as is that of the fibers themselves [168], with several lines of evidence questioning a direct contribution. Some degree of tau pathology was found in all individuals examined over age 75, increasing steadily with age, in the absence of cognitive impairment (n=26) [169]. Neuronal death in tauopathy patients greatly exceeds the number of neurofibrillary tangles accumulated [170]. Suppression of mutant tau expression, by addition of doxycycline in dox-off transgenic mice expressing P301L or ΔK280 mutant human tau and after cognitive deficits and tau pathology have begun to develop, leads to improvement in cognitive function despite the persistence of neurofibrillary tangles [171-173]. Behavioral impairment and neuronal loss occur in some tauopathy animal models in the absence of detectable insoluble tau [174-176]. Mice expressing full length mutant tau exhibit neuronal dysfunction before the presence of detectable, mature neurofibrillary tangles [172, 177]; further, fibrils from end-stage P301S mice intracerebrally injected into mice expressing full-length human tau develop clear histological tau pathology without obvious neuronal loss [96, 144]. These findings must be taken into account in dissecting tau's role in neurodegeneration.

The toxicity of tau fibrillization may be directly related to the prefibrillar conversion of tau rather than the presence of mature tau neurofibrillary tangles themselves. It has been hypothesized that destabilization of protein quality control systems by age-related dysfunction or external stress (as in chronic traumatic encephalopathy) by a stressor threshold effect [8] may allow this toxic conversion process to occur more rapidly or accumulate more efficiently, making upregulation of protein chaperones a potential therapeutic target [178-180]. As tau also has a

physiological role in many cellular pathways [62, 63], rapid sequestration of tau by fibril elongation may elicit dysfunction by a sudden loss-of-function of tau [181, 182]. Post-translational modifications including phosphorylation may not be required to template existing strains [14], however, they may influence the spontaneous origin of strains or affect the interactions of tau in these cellular pathways to cause dysfunction [133, 134]. Tau-induced microglial activation [183] or receptor activation [184, 185] may indirectly cause neuronal dysfunction in surrounding cells (**Figure 5.2B**).

5.9 Conclusion

The identification of unique fibril structures and evidence for propagation of strains in tauopathies provides a possible explanation for the distinct histological and clinical characteristics of tauopathies. Although significant progress has been made, additional research is needed to firmly establish the dominant mechanisms of neuronal release and uptake, the role of synaptic connections and activity on cellular transfer, cell type-specific vulnerability, and the conditions or cofactors responsible for the spontaneous origin of tau strains. Accumulating evidence suggests that tau fibrillization progresses within and between cells via templated conversion, similar to prions, Aβ, and other neurodegeneration-associated proteins [186, 187]; however, elucidation of strain-specific contributions to toxicity and cellular dysfunction is complicated by the possibility of mechanisms involving cell-autonomous effects and those arising from interactions with surrounding cells.

There are currently no therapies available to slow or halt the progression of tauopathies. Understanding the mechanisms by which tau strains propagate and cause or contribute to neurodegeneration will likely facilitate the discovery or design of

therapies [188-190] or immunization strategies aimed at particular conformations of tau [191-195].

Chapter 6

IMPROVED CELL CULTURE MODEL OF TAUOPATHY STRAIN PROPAGATION

6.1 Introduction

The microtubule-associated protein tau is involved in many cellular functions in addition to its commonly known role in stabilizing microtubules [60, 61]. Fibrillization and accumulation of tau is a defining characteristic of tauopathies and increasingly recognized to be involved in many other neurodegenerative diseases. Tau fibrillization is known to correlate strongly with disease pathogenesis suggesting a causative effect of amyloid formation [13]. In the native state, tau is a highly soluble intrinsically disordered protein which misfolds and forms distinct fibril strain conformations across different tauopathies (**Figure 6.1**) which may explain their unique histological and clinical characteristics. Different tau strains in different tauopathies such as Alzheimer's disease (AD), corticobasal degeneration (CBD), and Pick's disease (PiD) propagate in the brain in a prion-like manner targeting different regions and neuronal subtypes depending on their fibril strain conformation. Once monomeric tau misfolds into a particular strain conformation either spontaneously, by inherited pro-fibrillization mutations, or by direct interaction with misfolded tau protein, it begins to recruit additional natively expressed monomeric tau protein to propagate the misfolded structure in a prion-like manner (**Figure 6.1**). It is not known what cellular mechanisms initiate this process [196] or whether these fibril strains may be infectious in nature [197, 198], so a cellular model of tau strain fibrillization would

be useful to study how disease-associated tau strains propagate inside cells and how to inhibit this process.

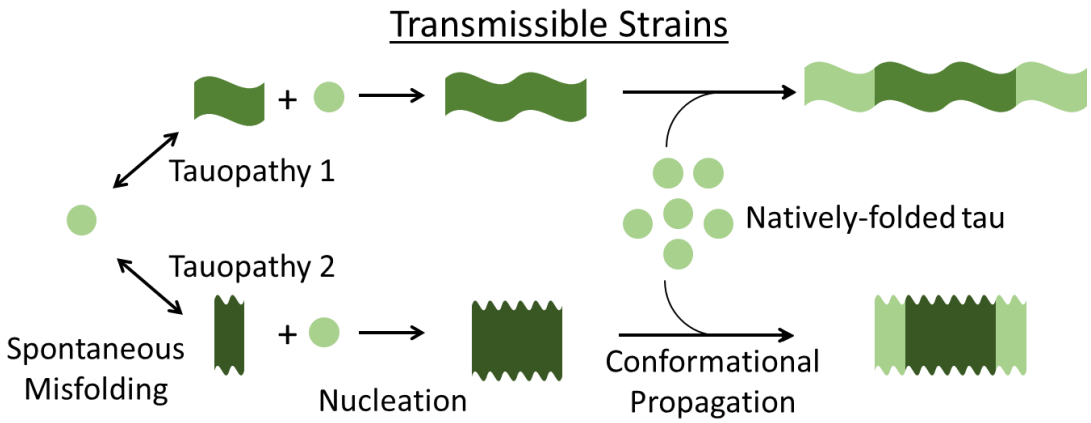


Figure 6.1 Tau protein is an inherently disordered protein which may misfold spontaneously, by inherited mutations, or by introduction of misfolded seeds. Different tauopathies are characterized by different conformational strains of tau protein which sequester native monomeric tau in a prion-like manner by direct protein-protein interaction to propagate these distinct fibrillar strains in cells and through the brain.

Culture models of tau fibrillization that currently exist typically rely upon non-physiological factors to promote tau fibrillization. Regarding the tau fibrils added to cells, synthetic fibrils induced to form tau amyloid by the poly-anion heparin are most commonly used. The resulting heparin-induced fibrils are significantly different from brain-derived fibrils by multiple ultra-structural analyses including a different width and period of twist by electron microscopy and different secondary structure content by cicrular dichroism [110-112, 199]. Additionally, upon incubating recombinant purified tau protein with brain-derived fibrils, the brain fibrils seed nucleation and

elongation of new fibrils which are structurally similar to the brain-derived seeds and also significantly different from the synthetic fibrils [14]. Brain-derived fibrils have been previously added to cells expressing tau in culture for the purpose of detection and so without determination of structural propagation [99, 138, 200, 201]. Regarding the choice of expressed tau construct full length human tau is occasionally used, but most models rely upon the expression of the microtubule-binding fragment for its propensity to aggregate easily as well as some common mutations identified in hereditary frontotemporal dementia patients such as V337M and P301L. The first study to clearly demonstrate prion-like strain behavior in culture and subsequently in mice used synthetic 4R fibrils to induce fibrillization before screening cells for fibrils which were able to propagate the same strain conformation to daughter cells [99]. They did not publish results showing strain behavior of disease fibrils.

In mouse models, synthetic fibrils have been used to demonstrate the cell-to-cell propagation of fibrils to distal regions of the brain [97, 202]. Further, brain-derived fibrils have been propagated in mouse brains with some evidence of disease-specific spread through the brain [98, 104]. However, in these studies no cytotoxicity was seen and there was no determination of structural differences between disease fibrils or brain-propagated fibrils. Cell-specific cytotoxicity in the brain was not observed until large doses of synthetic 4R fibrils were injected [145]. The ability of tau fibrils to spread their conformation to distal synaptically-connected regions of the brain suggests therapeutic approaches such as immunization or delivered antibodies may be feasible. Better discrimination of the conformational differences between disease-related strains would help understand the pathogenic differences in how these disease spread through the brain and target particular types of neurons. Understanding

how the conformational differences affect intracellular seeding will help develop biomolecules or small molecules capable of interrupting the process.

The goals of this project were to determine whether synthetic fibrils are sufficient for culture studies of tauopathy and which expressed tau substrate is suitable to faithful structural propagation of disease strains. Our disease model attempts to recapitulate the fibrillization process in the brain by uniquely applying disease-associated fibrils to cells expressing full length normal human tau protein to study the structural propagation of these disease strains. The "protein-only" hypothesis of tauopathies is becoming increasingly accepted and so we chose to use the immortalized human embryonic kidney (HEK) cells to investigate the cell-independent propagation of tau strains. We use a variety of orthogonal assays to detect the propagation of tau strains intracellularly and compare their structure to the parental disease-associated fibrils added to the cells including detection of aggregation with fluorescently-tagged tau substrate, chemical destabilization of fibrils, and limited proteolysis. We clearly show that synthetic fibrils are significantly different from disease-associated fibrils in culture and also that the cellularly-propagated fibrils are similar to their parental seeds. We also show that the shortened microtubule-binding fragment of tau may not be sufficient to propagate disease strains. This the first known demonstration of structural propagation of disease-associated tau strains intracellularly which may serve as a more-suitable model of tau fibrillization in culture for screening therapeutics and identification of therapeutic targets and disease biomarkers.

6.2 Components of Cell Culture Model of Tau Strain Propagation

To study tau fibrillization in a cellular environment, our culture model consists of human embryonic kidney (HEK) cells expressing different tau constructs. Parent

strains of tau fibrils are added to the cells so that their structure can propagate inside the cell and sequester monomeric soluble tau into newly formed fibrils (**Figure 6.2**). The primary expressed constructs consist of full-length, non-mutated human tau protein tagged with either enhanced green fluorescent protein or the hemagglutinin (HA) epitope tag (YPYDVPDYA) at the N-terminus. The N-terminus was chosen to minimize steric hindrance since the N-terminal portion of the tau protein typically extends from the fibrils while the microtubule binding region (4R) usually remains buried in the core of the fibrils. Occasionally other constructs were used such as just the 4R fragment fused to GFP to compare our results to other commonly used tau models. The GFP-tau construct is useful because its easy detection by microscopy allows real time monitoring of fibril formation in the cells. When parent seed fibrils are added to the cells, bright localized puncta may be seen by fluorescent microscopy which indicate the sequestration of solubly expressed GFP-tau into newly formed fibrils. In case the large GFP portion of the fusion protein has an effect on structural propagation, we use a HA tag to detect cellularly expressed tau as opposed to exogenously added fibrillar tau. The newly formed fibrils may be harvested and purified by lysis and centrifugation for structural studies to compare newly formed fibrils to their parent seeds.

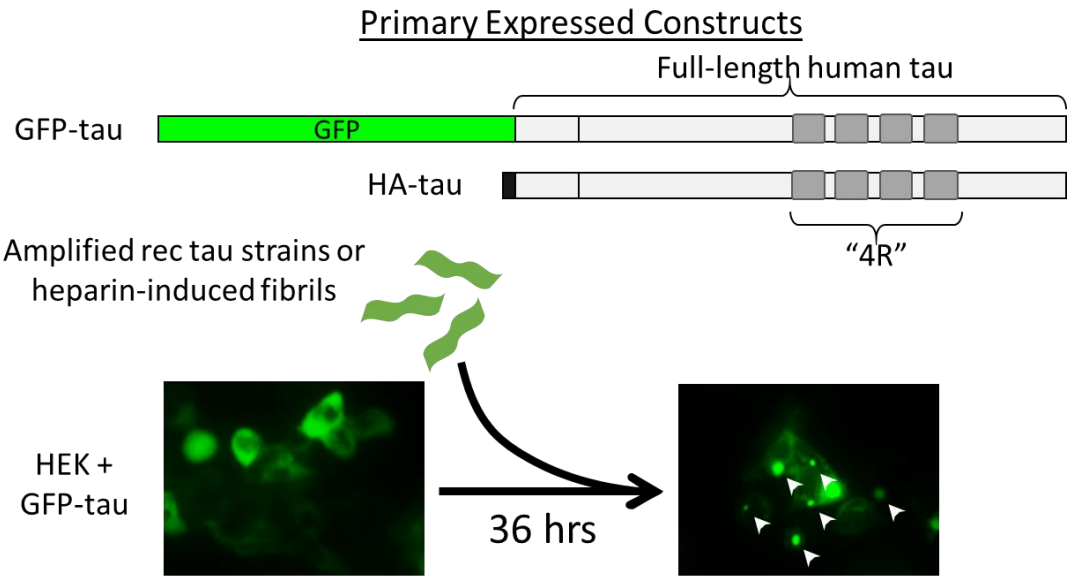


Figure 6.2 Model system to study tau fibrillization in cell culture. Primary expressed contructs include full-length non-mutated human tau protein with an N-terminal EGFP tag or a hemagglutinin (HA) epitope tag (YPYDVPDYA). Human embryonic kidney (HEK) cells are transfected to express monomeric tau protein, disease-associated or synthetic tau strains are added and 36 hours later, GFPtau cells are assayed for detection of puncta (white arrows) indicating formation of insoluble tau and newly formed fibrils from HAtau-expressing cells are harvested by lysis and centrifugation for structural assays.

The parental tau strain seeds are produced using recombinant tau by either forced induction of fibrillization using heparin (synthetic fibrils) or by seeding with fibrils isolated from human disease brains (disease-associated fibrils). All fibrils used in this study were generated by collaborator Olga Morozova. Before fibril addition, HEK cells are seeded in a tissue culture plate overnight, a tau construct is transiently transfected for a second night to allow for the construct to express, and then fibrils are pre-incubated with the lipid-based transfection reagent Lipofectamine 3000 before adding directly to the HEK cells. We found that fibrils added without transfection

reagent rapid associate with the cell surface but minimal uptake and propagation occurs. Upon adding the fibrils with Lipo-3000, puncta rapidly begin to form within 8 hours and reach a maximum by around 36 hours. Addition of the transfection reagent alone or with monomer tau is not sufficient to induce fibril formation, so newly formed intracellular fibrils are specifically induced by parent seed fibrils added to the cells.

Stable inducible HEK cell lines were generated by lentivirus containing GFP-tagged tau constructs in the TetO-FUW dox-inducible vector and sorted by FACS for populations of cells expressing uniform concentrations of the tau constructs. The reasons for generating stable cell lines were to eliminate variability in experiment-to-experiment expression levels as well as cell-to-cell expression levels since transient transfection may result in a distribution in which cells may be expressing high or low amounts of tau protein. Another reason for generating these lines was so that expression of the tau monomeric substrate could be controlled for studies of fibril persistence as well as threshold concentration for fibrillization in the cell. Our preliminary results found that low but detectable expression of GFP-tau with addition of tau fibrils was not able to form new puncta. After induction with dox and subsequently higher tau expression, fibril addition was able to generate new puncta. This indicates that there is a minimum concentration required for tau fibrillization in the cell which may vary between different strains of tau and required further investigation.

6.3 Both Synthetic and Disease-Associated Fibrils are Capable of Seeding GFP-tau in Cells

Synthetic fibrils induced by the poly-anion heparin are commonly used in the tau field for fibrillization studies and it is known that these synthetic fibrils can induce GFP-tau to fibrillize in cells. Without the addition of seeds, no new puncta are detected, indicating that full length GFP-tau remains monomeric in the cell and does not spontaneously fibrillize. Upon addition of either synthetic or disease associated fibrils including those generated from the brains of patients with Alzheimer's disease (AD), corticobasal degeneration (CBD), and Pick's disease (PiD), both synthetic and all disease-associated fibrils caused GFP-tau to form de novo fibrils in the cell (**Figure 6.3**). The total amount of puncta formed in either case is similar and so this result alone does not tell us whether synthetic or disease-associated fibrils are a more appropriate choice of tau strain seeds in a cellular model of tauopathy. We therefore perform additional assays to obtain structural information to compare synthetic and disease-associated fibrils in culture.

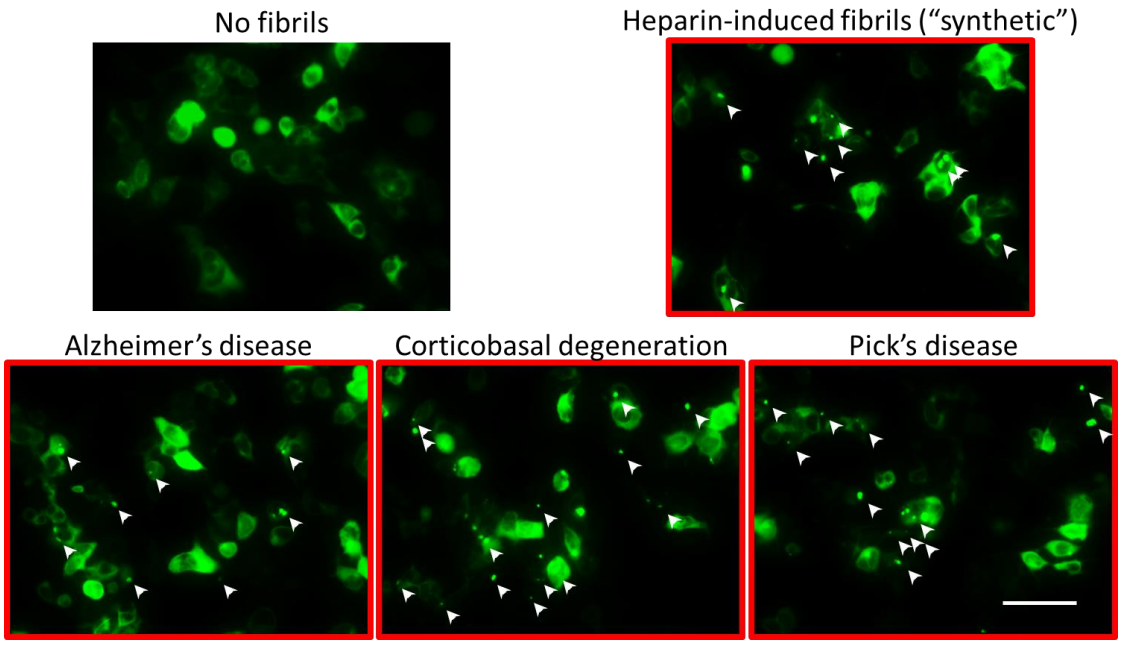


Figure 6.3 Both synthetic and disease-associated fibrils are able to induce new fibril formation in HEK cells expressing GFPtau. No puncta are observed in cells without the addition of tau fibrils. (scale bar = 100 µm)

6.4 Quantification of Structural Stability of Parent and Intracellular Fibrils by Destabilization with Guanidine

A typical method to quantify protein stability is to measure intrinsic tryptophan fluorescence with increasing concentrations of chemical denaturants such as urea or guanidine. However, the full length tau protein does not have any tryptophan residues and only has 6 hydrophobic residues in total which is why monomeric tau protein is highly soluble. We chose to use an alternate method to measure fibril stability that takes advantage of the large size scale differences between monomeric and fibrillary tau protein. Tau fibrils formed with recombinant tau or tau fibrils isolated from seeded HEK cells were subjected to a range of guanidine concentrations to destabilize the fibrillar structure and then each sample was run through a 96 well vacuum-assisted

dotblot apparatus through a cellulose acetate membrane with 0.2um porosity. Cellulose acetate does not bind proteins, so fibrillar tau protein remains on the surface of the membrane where it may be probed similarly to a Western blot while monomeric tau travels through the membrane to a waste container. Parent fibrils formed *in vitro* from recombinant tau are probed with the anti-tau antibody 5A6 to quantify remaining fibrils at each concentration of guanidine and the intracellularly-formed fibrils are probed with the anti-HA antibody Y11 to specifically detect HA-tagged tau protein that was expressed inside the HEK cells. The cellulose acetate membrane is then probed with a horseradish peroxidase secondary antibody and exposed with a chemiluminescent reagent to image the blots. The blots are then quantified by measuring the pixel intensity of each dot and the resulting intensities are fit to a sigmoidal unfolding curve and the half-concentration of guanidine is calculated at the point which half of the remaining fibrils are left as a metric of fibril stability.

One challenge encountered during the operation of this filter-retardation assay was the purification of intracellular fibrils from other cellular components. HEK cells containing newly formed fibrils had to be dissociated with a non-enzymatic cell dissociate reagent to prevent proteolysis of tau fibrils after lysing. Once the cells were in suspension, several detergents were tested including TritonX-100, Tween20, and sarkosyl. Only sarkosyl was able to sufficiently dissociate the cell membranes uniformly. Tau fibrils are sarkosyl-stable which is why sarkosyl is frequently used to purify fibrils from brain samples. To further destabilize cellular proteins and membranes, the reducing agent DTT was added since tau fibrils are not held together by disulfide bridges which further clarified the cell lysate. Genomic DNA was greatly increasing the viscosity of the lysate which would prevent centrifugation of tau fibrils

and so an endonuclease Benzonase was added to cleave cellular DNA. However, Benzonase requires magnesium as a cofactor and the cells were dissociated in a chelating agent so additional magnesium chloride was added to ensure activity of the Benzonase. A protease inhibitor cocktail and PMSF were added to prevent any cellular proteases from cleaving the tau protein. Once the cells are homogenized in this lysis cocktail for 5-10 minutes, the lysate is still not able to be run through the dotblot apparatus without clogging the cellulose acetate membrane and so cell lysates are centrifuged at 16k*g for 30-45 minutes to pellet the insoluble tau fibrils and remove the majority of cellular components, thereby crudely purifying the fibrils. The resulting fibrils are then able to be resuspended in buffer and titrated in a range of guanidine concentrations overnight before running them on a cellulose acetate dotblot.

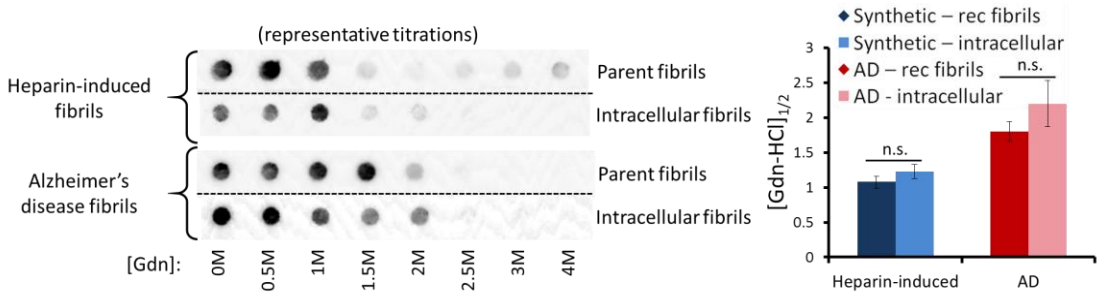


Figure 6.4 Filter retardation assay was used on parent fibrils propagated in vitro with recombinant tau and intracellular fibrils propagated by seeding with the parent fibrils over a range of guanidine to measure fibril stability. The concentration of guanidine at which half of the fibrils remained was measured by fitting the fibril signal to a standard unfolding sigmoidal curve. Parent fibril seed were probed with 5A6 anti-tau antibody and intracellularly formed fibrils were probed with an anti-HA antibody to specifically detect newly formed fibrils.

Both synthetic and Alzheimer's disease parent fibrils and intracellularly seeded fibrils were run on the filter retardation assay and qualitatively, it appears that parent fibrils have a similar unfolding curve as the intracellular fibrils they seed (**Figure 6.4**). Additionally, parent and intracellular synthetic fibrils appear to be significantly different from the parent and intracellular AD fibrils. After quantifying the results of multiple blots from multiple biological replicates, the guanidine half-concentration of parent fibrils was not significantly different from their intracellularly-seeded counterparts, but synthetic fibrils were significantly different from AD-associated fibrils (**Figure 6.4**). The two important conclusions from this result is that by structural stability, it appears that structure is propagated by synthetic and disease-associated strains in cells expressing an identical monomeric tau construct. More importantly, the synthetic fibrils are significantly different from the disease-associated fibrils and so a cell culture model of tauopathy fibrillization using brain-derived fibril strains may yield more predictive results for testing potential therapeutics then a culture model relying upon fibrillization of synthetic seeds.

6.5 Comparison of Fibril Structure by Limited Proteolysis

Limited proteolysis is a technique used to gain structural information of protein fibrils by taking advantage of the regular repeating assembly of protein amyloids. Individual tau monomers are assembled into tau strains in slightly different conformations which leaves different region of each monomer solvent exposed to allow proteases to cleave the peptide chain at regular intervals resulting in a peptide "fingerprint" of remaining fragment sizes. So different conformations will have a different band pattern allowing a comparison of different tau strains. In brief, fibrils are incubated in buffer in the presence of a protease for a period time followed by

standard SDS-PAGE on a small-peptide gel and transfer to a nitrocellulose membrane to probe with an anti-tau antibody. The tau antibody 77G7 was chosen because it binds the tau protein at the beginning of the 4R region which should be buried in the fibril structure and therefore protected from protease cleavage. Antibodies targeting other regions of the protein may not be able to detect any protein fragments because of complete cleavage by protease. Both the Pronase protease cocktail and trypsin were tested to see which protease performed better, and trypsin was found to generate similar but slightly cleaner band patterns.

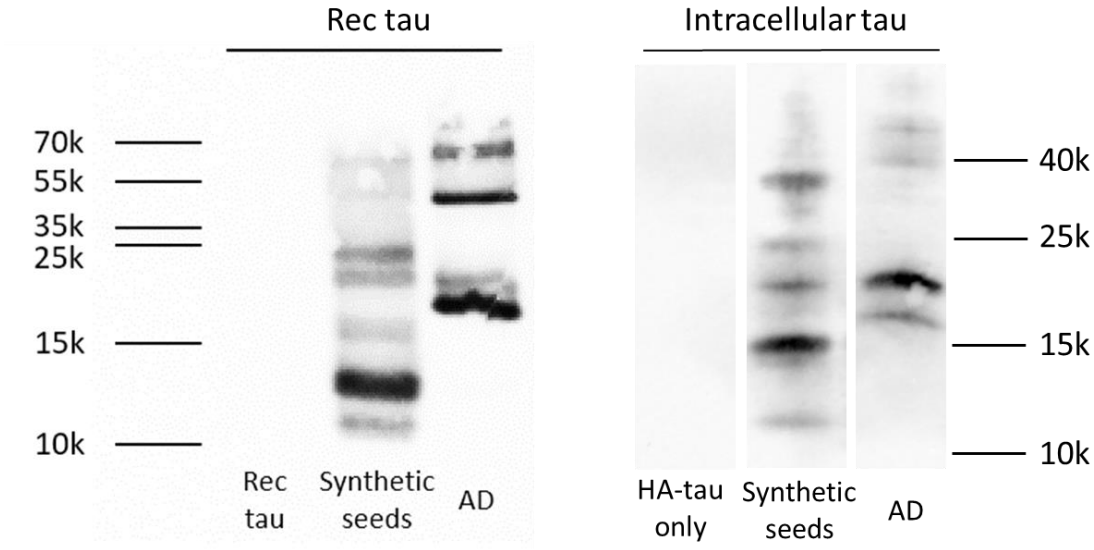


Figure 6.5 Limited proteolysis with trypsin was used to compare fibril structures by the "fingerprint" of bands remaining after proteolysis and Western blot for parental fibril seeds with a tau monomer control and intracellularly formed fibrils with a control of HEK cells expressing un-seeded HAtau. Blots were probed with 77G7 antibody.

When monomeric recombinant tau is exposed to trypsin, no detectable bands were seen indicating complete digestion since the entire monomer is solvent-exposed (**Figure 6.5**). However incubation of synthetic and Alzheimer's disease fibrils with trypsin results in clear but significantly different band patterns confirming the significant differences between synthetic and disease-associated fibrils. By performing a mock-isolation of tau protein from HEK cells expressing HA-tagged tau without adding fibril seeds, no detectable bands are seen similar to monomeric recombinant tau indicating that the intracellularly-expressed full length tau protein is not spontaneously fibrillizing. By isolating newly formed fibrils from HEK cells expressing HA-tau and seeded with recombinant synthetic and disease-associated fibrils and subjecting them to trypsin treatment, similar band patterns were seen compared to the parent fibrils (**Figure 6.5**). It is important to note that cells seeded with tau fibrils but not expressing tau monomer did not have detectable bands after limited proteolysis which supports that any bands detected are specifically from newly formed intracellular fibrils. While the band patterns of parent seeds and intracellular fibrils are qualitatively similar, there are some differences which may have arisen from non-specific association of cellular components with the newly formed intracellular fibrils which would have slightly changed which regions are fully solvent-exposed for access by protease. The important conclusion is that the synthetic and disease-associated fibrils propagated in the cell are significantly different from each other further supporting that disease-associated fibrils are likely a more representative choice for a cell culture model of tauopathy.

6.6 Intracellular Fibrillization of Full Length Tau and Truncated Tau 4R Fragment

The 4R region of tau which is responsible for reversible binding and stabilizing of microtubules is known to comprise the core of tau fibrils [123]. This region, when expressed by itself, will spontaneously fibrillize and therefore has been frequently used for cellular and recombinant studies of tau fibrillization. However, its suitability in a cellular model of tau fibrillization has not been tested with disease-associated strains of tau propagated with recombinant tau. To compare full length and truncated 4R tau in culture, we expressed either GFP-tau or 4R-GFP in HEK cells and added either synthetic or disease-associated fibrils to check for intracellular fibrillization. Without fibril addition, no puncta were seen in HEK cells expressing either 4R-GFP or GFP-tau. When synthetic seeds were added, both 4R-GFP and GFP-tau were able to robustly form puncta in HEK cells with similar efficiency (**Figure 6.6**). For studies of detection and quantification of misfolded tau, 4R-GFP would be an appropriate expressed substrate for synthetic heparin-induced seeds.

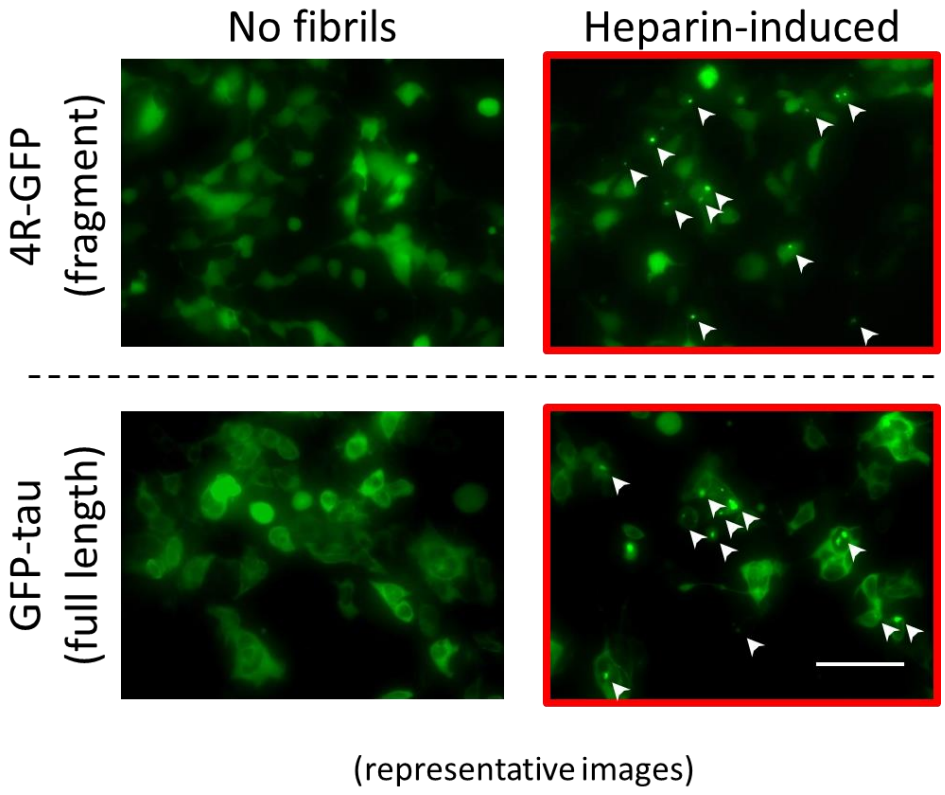


Figure 6.6 Synthetic seeds induced by heparin are able to induce puncta (white arrows) formation in both the microtubule-binding fragment of tau 4R-GFP and full length GFP-tau. No puncta were observed without fibril addition (scale bar = 100 µm)

As a point of comparison to the samples in which synthetic seeds were added, disease-associated tau strains including those generated from brains of patients with Alzheimer's disease, corticobasal degeneration, or Pick's disease, were added to HEK cells expressing 4R-GFP or GFP-tau. Similar to the synthetic fibrils, all tested disease-associated fibrils were able to fibrillize full length GFP-tau in HEK cells (**Figure 6.7**). However, in sharp contrast to the synthetic fibrils, none of the disease-associated strains were able to generate puncta in 4R-GFP cells. Other studies have had to use

variants of 4R-GFP with pro-fibrillization mutations in order to permit puncta formation with fibrils isolated from human brains [99]. The differences in fibril conformation across the synthetic and disease-associated structures must account for the ability for synthetic but not disease-associated strains to fibrillize 4R-GFP monomer. A panel of tau constructs were generated which contained various regions of the tau protein with a GFP tag at either end for further investigation of which regions of the tau protein are necessary for fibril elongation with different fibril strains.

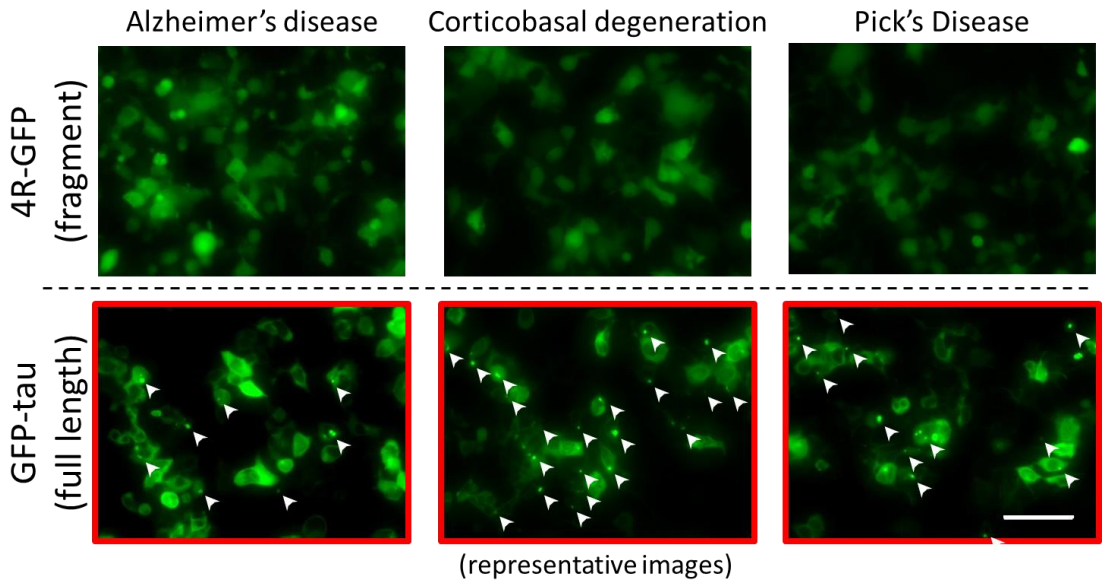


Figure 6.7 Disease-associated fibrils including Alzheimer's disease, corticobasal degeneration, and Pick's disease seeds are able to induce puncta (white arrows) formation in full length GFP-tau, but not 4R-GFP-expressing HEK cells. All experimental replicates were performed with 4R-GFP and GFP-tau simultaneously and puncta formation was not observed in any 4R-GFP wells. (scale bar = 100 µm)

6.7 Discussion

We have demonstrated with our model the structural propagation of disease-associated fibrils intracellularly. This has important implications on the protein-only hypothesis of tauopathies since conformation-specific strain properties may be conserved as disease fibrils propagate through the brain during disease pathogenesis. The potentially infectious nature of tau strains is still not commonly accepted and so using an improved model of tau strain fibrillization in culture may help better understand the mechanisms of how fibrils enter the cell and interact with monomeric native tau to induce new misfolding and aggregation of tau protein. Pending data that tau strains are truly infectious in the way in which prions are infectious, tau fibrils have still been shown to exhibit a toxic gain-of-function upon fibrillization and so determining how to inhibit the process of intracellular fibrillization could potentially halt the progression of various tauopathies in human patients.

Our model has confirmed that synthetic fibrils are significantly different from disease-associated fibrils on a structural level, but we have further shown that synthetic fibrils also propagate differently than disease-associated strains in culture. This indicates that synthetic fibrils induced by heparin to spontaneously aggregate may not be a suitable choice for studying tau fibrillization in culture. Additionally, the demonstration of faithful propagation of conformation from parent seed fibrils to newly formed intracellular fibrils shows that disease-specific fibrils may be a better choice of seed for studying particular diseases and their conformation-specific effects.

The choice of tau expression construct as a monomeric substrate for intracellular seeding was found to be important and that full length normal human tau may be the best construct. Since the unassisted rate of uptake for disease-associated fibrils may be low, other groups frequently use truncated or mutant tau to promote

fibrillization for the detection of disease fibrils. However, if we are to study structural propagation we found that truncations and mutations may not properly propagate structure. Upon comparing synthetic and disease-associated fibrils added to HEK cells expressing either full length tau or the 4R fragment of tau, we found a stark contrast between full length and 4R tau substrates. Both substrates were able to fibrillize into newly formed fibrils when seeded by synthetic fibrils, but only full length tau was able to fibrillize into new fibrils when seeded by disease-associated fibrils from patients with AD, CBD, and PiD. This indicates that the minimal segment of tau protein that interacts with fibrils to misfold and elongate the fibrils may be different for disease-associated fibrils and synthetic fibrils. By generating inducible stable HEK cell lines expressing a consistent basal intracellular concentration, we found that un-induced lines were unable to induce fibrillization and upon dox-induction and subsequent expression increase we saw robust fibrillization. This indicates that there is a minimum concentration of monomer substrate required for fibrillization to occur which may give some insight into the kinetics of fibril formation that may differ between diseases.

Regarding pro-fibrillization mutations, we saw an unexpected result with respect to what is referred to as the asymmetric seeding barrier. Typically, synthetic seeds generated from wild type tau substrate are capable of inducing fibrillization in monomeric wild type and mutant tau, but seeds generated from mutant tau only fibrillize mutant substrate. This seeding barrier held true for wild type synthetic seeds which were able to fibrillize both wild type GFP-tau and P301L GFP-tau, and seeds generated from a mutant stable synthetic strain [99] were only able to induce fibrillization in P301L GFP-tau as expected. Seeds propagated from PiD patient-

derived fibrils were only able to induce fibrillization in wild type GFP-tau but not the mutant variant indicating that these pro-fibrillization mutations may bias the conformational possibilities such that disease-associated fibrils are unable to elongate with a mutant monomer substrate.

Through the course of this study we encountered some interesting results which have contributed to new lines of research for which we have been able to lay the groundwork and obtain some preliminary data. The stark difference in intracellular fibrillization between 4R and full length tau for disease-strains led to the generation of a panel of tau constructs which contain differing regions of the tau protein adjacent to the 4R region to determine what the minimum elongation region is and how it may differ between tauopathies. The observation that basal expression of the tau construct was not sufficient to allow intracellular fibrillization until inducing higher expression suggests a study in which the inducer concentration is varied to determine quantitatively what the minimal fibrillization concentration is in a cellular environment and how that may differ between tauopathy strains. Determination of the minimal fibrillization concentration will give some insight into the kinetics of fibrillization to better understand the conditions in the brain which permit the spread of these strains to other neural cells. These results also suggest that the assays performed here should be repeated for other tauopathy strains to determine which assays are able to discriminate between different strains most effectively for diagnostic purposes. Since all experiments were performed in HEK cells to study the cell-independent propagation of these structures, it would be an important next step to see how these strains behave in neuronal cells and determine whether different conformational strains have different degrees of uptake and neurotoxicity. Overall, we

have demonstrated that disease-associated tau strains can propagate in a cellular environment and that synthetic strains and shortened or mutated expression monomer substrates may not be representative of *in vivo* propagation of tauopathy. Our improved model may yield more predictive results for screening potential therapeutics and better understanding the pathogenesis of tauopathies.

Chapter 7

HETEROLOGOUS-HOST ENGINEREERING OF AN ANTI-PRION PROTEIN SCFV FOR IMPROVED STABILITY AND EXPRESSION

7.1 Introduction

Protein-based therapeutics and especially antibodies are becoming increasingly important for the treatment of a broad range of human diseases [203]. Single-chain Fv (scFv) antibodies composed of the VL and VH domains connected by a short linker [204, 205] are attractive alternatives to a full-sized IgG antibody in situations that do not require the FC effector functions. Antigen-binding specificity of the shortened scFv is maintained while expression as a single continuous peptide sequence facilitates engineering and expression in a range of hosts. Some issues encountered by generation of scFvs from full IgGs or Fabs include reduced stability due to exposure of hydrophobic inter-domain faces that may promote aggregation and loss of the stabilizing function of constant-region immunoglobulin domains [206, 207] which can lead to decreased secretion yield [208]. A variety of methods including directed evolution have been used successfully to improve scFv stability and production [209-211].

Mammalian cells and especially Chinese hamster ovary (CHO) cells are currently the expression system of choice for the production of complex therapeutic proteins such as antibodies [212]. Due to significant advances in media and bioprocess optimization, the titers of complex proteins in CHO cells are significantly higher than microbial expression systems [213]. However, rapid engineering and screening of

proteins in CHO cells is unfeasible due to the time-consuming process of cell line development [214]. Since *Saccharomyces* have eukaryotic expression machinery and a fast doubling time relative to mammalian cells, yeast surface display is convenient for rapid screening of mutagenesis libraries by fluorescence-activated cell sorting (FACS) [215, 216]. Selected clones may then be transferred back into a mammalian expression system while maintaining improvements in stability [217-219].

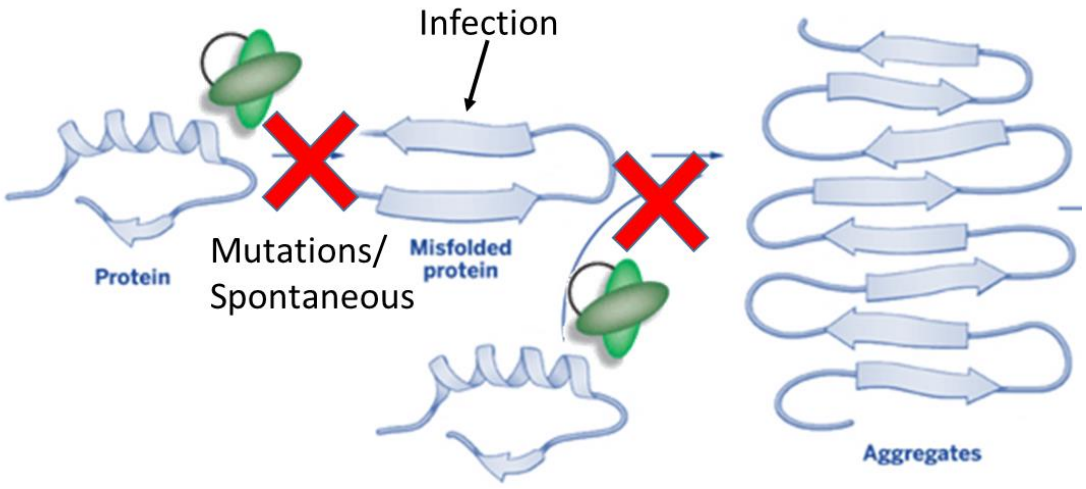


Figure 7.1 Native prion protein may be induced to misfold into a disease-conformation either spontaneously or by familial disease mutations. Misfolded protein may also be introduced exogenously by dietary intake. The native conformation is induced to misfold into the disease conformation by templated interaction with the misfolded protein. Addition of an antibody specific to the prion protein may block direct protein-protein interactions to halt the spread of the disease. (Adapted from Wolf LK. Alz Prion Connection. C&EN. 2012).

We chose to engineer the anti-prion antibody ICSM18 because of its therapeutic interest in prion disease [220], though the neurotoxicity of ICSM18 is a

subject of debate [221, 222] Prion disease is caused by misfolding of the native conformation into a disease-causing conformation by interaction with misfolded prion protein [223]. Antibodies targeting the prion protein may be able to inhibit this conversion process (**Figure 7.1**) and also be used to study PrP-PrP interactions [224-229]. In this study, the VL and VH domains from ICSM18 Fab [230] were fused by a linker to generate an ICSM18 scFv for expression in mammalian cells. The scFv was optimized by engineering on YSD which translated to improvements in thermal stability and specific productivity in mammalian cells.

7.2 Generation of an ICSM18 ScFv From ICSM18 Fab

To generate a smaller version of the ICSM18 Fab, we have generated the ICSM18 scFv by fusing the variable regions of the light and heavy chains of ICSM18 Fab with a 15 AA $(Gly_4Ser)_3$ linker [17]. The resulting protein sequence was converted to a DNA sequence optimized for mouse codon usage and cloned into a dual promoter lentiviral mammalian expression vector pHAGE. The pHAGE expression vector expresses scFv with an IgK light-chain secretion signal and a FLAG tag for detection and purification under the constitutive cytomegalovirus (CMV) promoter and an enhanced green fluorescent protein (EGFP) reporter under the constitutive human ubiquitin C (UBC) promoter (**Figure 7.2**). The EGFP reporter was used to confirm integration and relative degree of integration by median fluorescence.

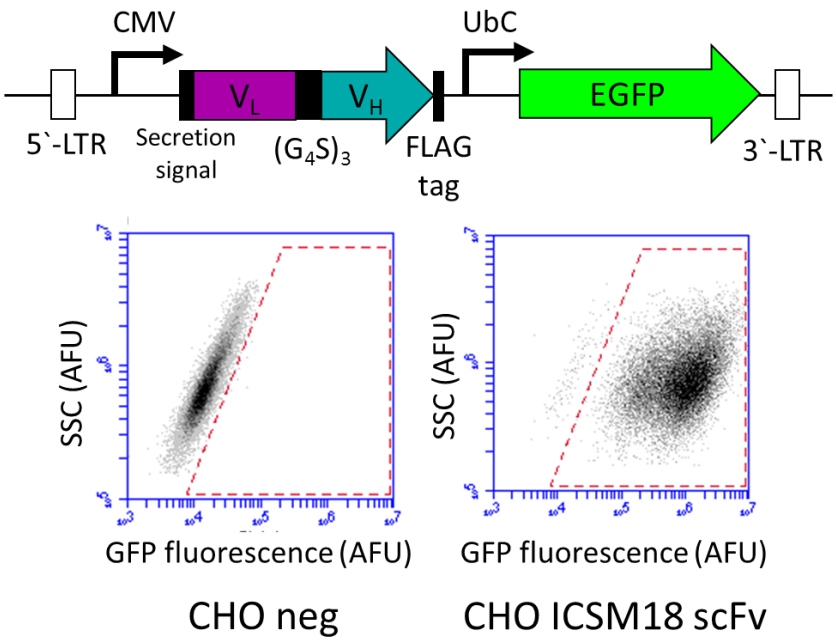


Figure 7.2 Diagram of the ICSM18 scFv construct. Light and Heavy variable domains were linked with a $(G_4S)_3$ linker and fused to an IgK secretion signal and a FLAG tag (DYKDDDDK) for detection and purification. The scFv construct was subcloned into a lentiviral expression vector under a constitutive CMV promoter and co-expressed with an EGFP reporter under a constitutive UbC promoter. Positive integration by lentivirus is detectable by flow cytometry for GFP fluorescence by comparing un-transduced CHO cells with CHO cells infected with the ICSM18 scFv cassette.

CHO cells adapted to serum-free suspension culture were transduced with pHAGE-ICSM18 scFv lentivirus to generate a stable CHO suspension cell line constitutively expressing and secreting ICSM18 scFv. Cultures were grown for 3-4 days in serum-free medium and the conditioned supernatant was probed by Western blot to assess scFv secretion capacity. We found that specific productivity (pg/cell/day; pcd) of secreted ICSM18 scFv was significantly lower than values reported in literature (**Table 7.1**). CHO ICSM18 cells were sorted by FACS for low,

intermediate, or high levels of EGFP reporter expression corresponding to different degrees of genomic integration. Specific productivity correlated linearly with median EGFP fluorescence (**Supp Figure B.1**), but the yield of the highest EGFP-expressing cells were still lower than reported values (~0.3 pcd). To improve the expression yield, we decided to take a protein engineering approach by using directed evolution with yeast surface display.

Product	Qp range (pg/cell/day)	Citation	
IgG-Fc	15-87	Zboray et al.	2015
mAb	0.2-24	Mason et al.	2013
mAb	2-3	Lee et al.	2005
mAb	3-50	Edros et al.	2014
mAb	3-75	Chusainow et al.	2009
mAb	15-40	Dorai et al.	2013
mAb	30-50	Reinhart et al.	2015
mAb	80-110	Fouser et al.	1992
mAb, Fc-fusion	16-49	Huang et al.	2010
scFv-Fc	0.8-3.3	Mayrhofer et al.	2014
scFv-Fc	20-40	Kawabe et al.	2012
scFv-fusion	5-10	Berger et al.	2013
ICSM18 scFv (WT)	<1	This study	
ICSM18 scFV (2.6.1)	7-8*	This study	

Table 7.1 Reported typical values of specific productivity (pg/cell/day) of different antibody variants and values of the wild type and engineered ICSM18 scFvs reported in this study.

7.3 Engineering of ICSM18 ScFv by Yeast Surface Display

The ICSM18 sequence was cloned into the yeast surface display vector pCTCON2. The resulting yeast vector expresses ICSM18 scFv with a c-myc tag fused to the yeast mating agglutinin protein Aga2p, which is exported to the cell surface

linked by two disulfides to the Aga1p protein linked covalently to the cell wall. The scFv fusion protein and Aga1p are expressed under the galactose-inducible GAL1 promoter for inducible overexpression of the yeast surface display construct.

Before engineering the scFv we decided to compare the binding affinity of the yeast surface displayed construct to the CHO-secreted WT ICSM18 scFv to see if affinity was maintained upon transfer into a different host and different fusion construct. WT ICSM18 scFv was scaled up in CHO culture, purified by FLAG affinity chromatography, and quantified by BCA assay after confirming purity by silver stain . Mouse neuroblastoma N2a cells were used because they express natively-folded mouse PrP on the cell surface. Purified scFv was titrated on the surface of non-enzymatically dissociated N2a cells and probed with antibodies against the scFv FLAG tag to estimate binding affinity (**Figure 7.3**). In a similar manner but with the scFv anchored on the yeast surface and titration of soluble recombinant labeled PrP, the yeast surface displayed ICSM18 scFv binding affinity was calculated (**Figure 7.3**). We found that the CHO-secreted scFv had a relatively high affinity (~2nM), however, the yeast displayed scFv had a significantly lower apparent affinity (**Figure 7.3**). We were unable to carry the titration to higher concentrations of PrP since PrP in its soluble form is highly unstable and will begin to precipitate out of solution. From the data we can at least conclude that the affinity is at least higher than 100nM which is a couple orders of magnitude worse that the CHO-secreted form.

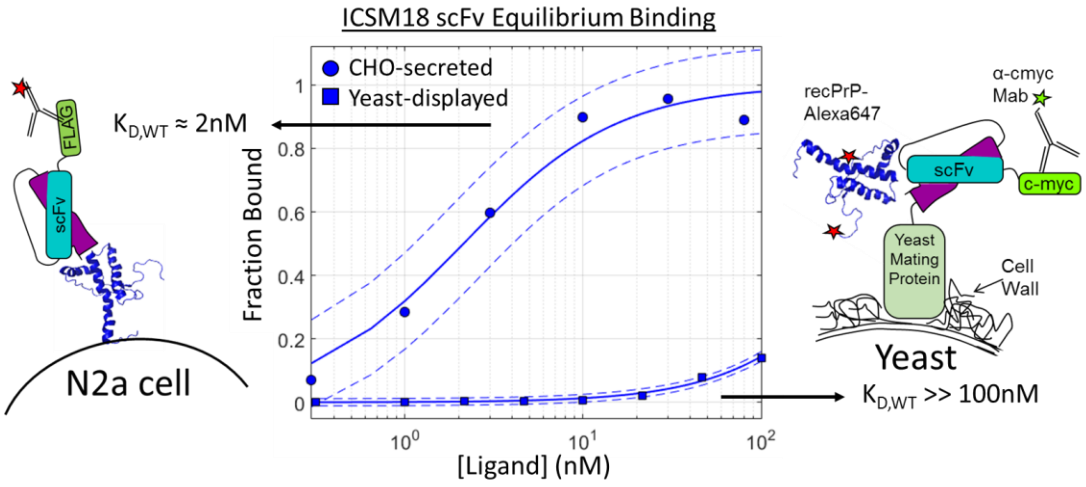


Figure 7.3 Equilibrium binding curves of WT ICSM18 scFv and PrP. For CHO-secreted scFv (blue circles), purified scFv was titrated on the surface of N2a mouse neuroblastoma cells expressing native PrP on the surface followed by labeling with an anti-FLAG antibody to detect binding by flow cytometry (schematic on the left). For yeast surface displayed scFv (blue squares), soluble labeled PrP was titrated on WT ICSM18 scFv expressed on the yeast surface (schematic on the right). Titrated ligand was always in excess and so a standard equilibrium binding curve may be applied. The measured affinity of CHO-secreted WT ICSM18 scFv was about 2nM and the affinity of yeast displayed ICSM18 scFv is greater than 100nM indicating at least a two order of magnitude loss in affinity.

To determine whether the reduced apparent affinity is caused by the differences in yeast and mammalian expression systems or merely the act of tethering the scFv to a surface, we generated an integrated yeast line capable of secreting scFv by fusion to the alpha mating factor secretion signal. We found that scFv secreted by yeast is unstable and preferentially forms either SDS-insoluble dimers or degradation products which were cleaved at the flexible linker. By running a Western blot with non-reducing and reducing conditions, we found that the yeast secreted scFv forms higher order oligomers which dissociate to a stable dimer in the presence of reducer

(**Figure 7.4**). The presence of these disulfide-linked higher order oligomers indicates that the intra-domain disulfides within each immunoglobulin domain are interacting with other monomers showing that the monomers are not well-folded. Additionally, the yeast-secreted stable dimers were not able to dissociate in the presence of extra denaturant and boiling time which means that the dimer is highly stable.

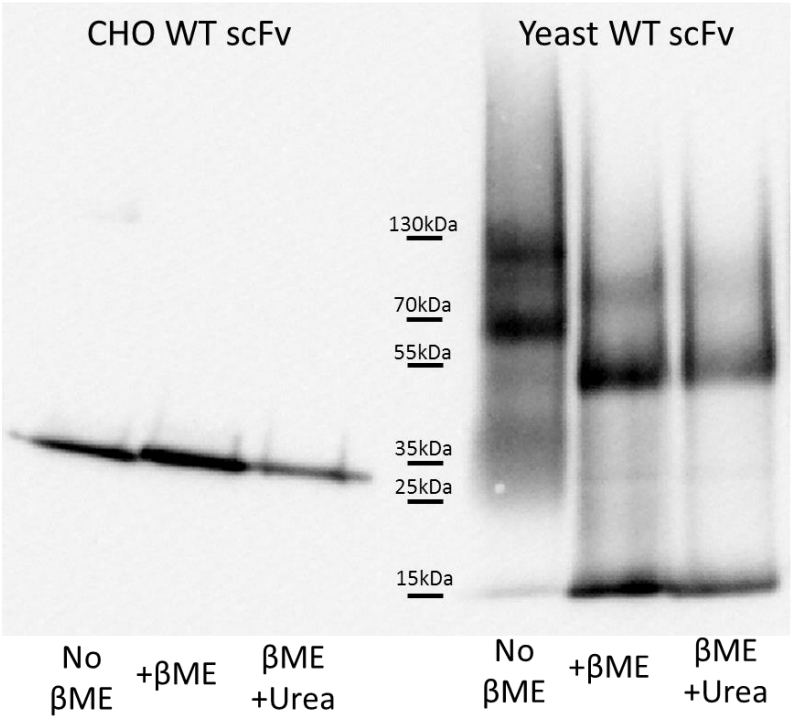


Figure 7.4 Comparison of CHO-secreted and yeast-secreted (pITy yeast-integrated secretion cassette) WT ICSM18 scFv shows that yeast expression and secretion machinery are unable to properly process the un-engineered WT ICSM18 scFv. Purified scFv samples were run on a standard SDS-PAGE gel either with no reducing agent (beta-mercaptoethanol) to maintain disulfide bonds, with reducing agent, or additionally with 4M urea denaturant and an additional 20 min of boiling time. The non-reduced samples show that yeast-secreted scFv forms disulfides with other monomers and the samples subjected to additional denaturation with urea show the formation of a highly stable irreversible dimer.

A directed evolution approach to engineering ICSM18 scFv was used by rounds of mutagenesis by error-prone PCR and sorting of the yeast libraries for improved clones by FACS (**Figure 7.5**). After two rounds of mutagenesis with 6-9 rounds of enrichment by FACS following each mutagenesis, a clone with the highest signal was selected for further analysis. All mutatgenesis and sorting was performed by Kyle Doolan. Clone 2.6.1 was sequenced by Sanger sequencing to identify which mutations were introduced. There were three amino acid substitutions in framework regions of the immunoglobulin domains, a single mutation in the flexible linker, and 4 silent mutations which accounted for improvements in yeast codon usage (**Figure 7.6**). A mutation near the binding pocket, R92G, may have reduced steric hindrance of the antigen to the binding pocket since arginine is a bulky reside. Two methionines in the light chain framework were mutated to an isoleucine (M21I) and a valine (M77V) respectively. Both of these new residues are slightly smaller and much higher on the hydrophobic index which may have helped them bury into the domain core better to improve stability. The last amino acid substitution was found in the linker region (G116D) which is unsurprising since it is known that mutating exposed linker residues to charged amino acids helps improve stability and solubility [231]. The improved codon frequency of the silent mutations was unsurprising since selection was performed in a yeast host.

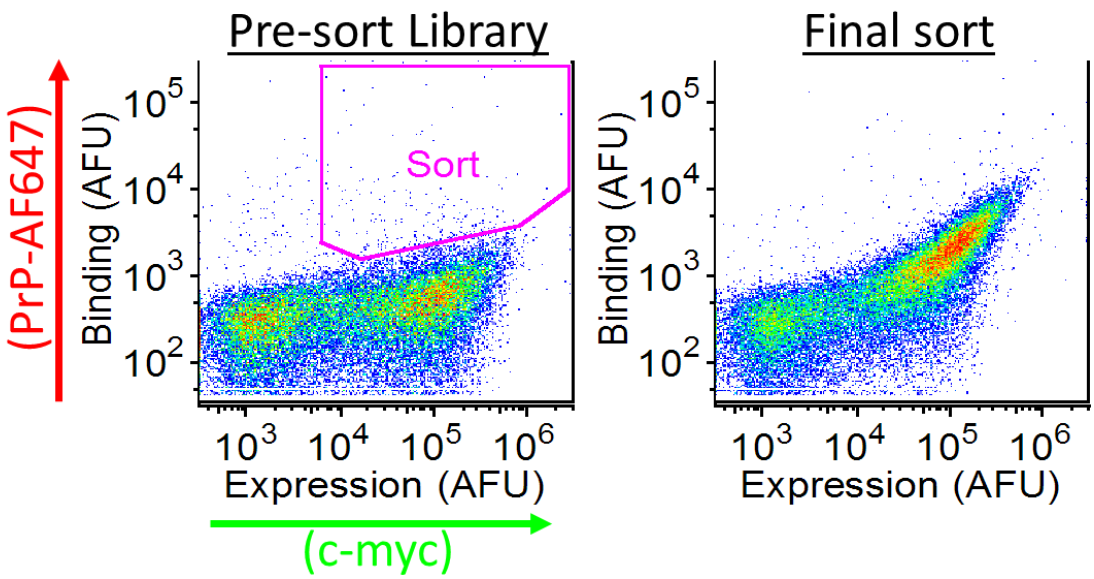


Figure 7.5 Yeast surface display system and directed evolution representative plots. The ICSM18 scFv is tethered to the yeast surface by expressing it as a fusion protein to a yeast mating protein and a c-myc tag (EQKLISEEDL) is added to the C-terminus for detection of expression. Labeled prion protein (recPrP-AF647) is added to test binding to the scFv. Sorting was conducted at a low PrP concentration (10nM) and after multiple rounds of mutagenesis and sorting, a population with higher binding was isolated. (Mutagenesis and sorting was performed by Kyle Doolan)

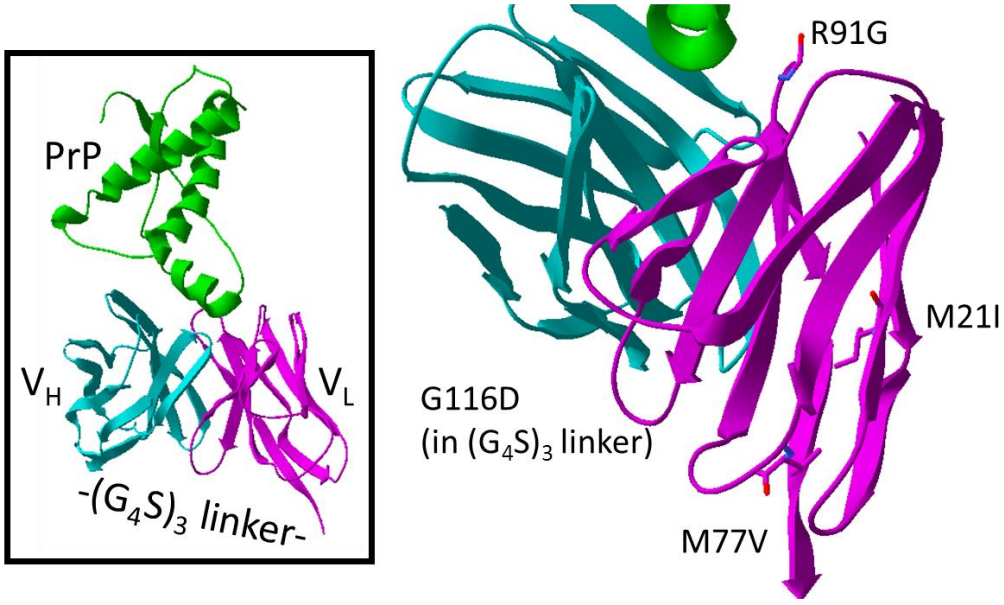


Figure 7.6 Ribbon model of the ICSM18 scFv and an enlarged view of the 2.6.1 mutant with four amino acid substitutions. Three of the mutations are in the light chain shown in front (magenta) and the fourth mutation is in the linker region (not rendered). The heavy chain is in the rear of the image (teal) and a segment of the prion protein (green) is shown to indicate the location of the binding pocket.

7.4 Characterization of Improved ICSM18 2.6.1 ScFv Mutant

Since the improved ICSM18 scFv mutant 2.6.1 had a higher PrP binding signal at the 10nM PrP sorting conditions, we hypothesized that the affinity on the yeast surface had improved. We titrated soluble labeled PrP on the 2.6.1 ICSM18 scFv yeast surface displayed construct and found that the apparent affinity had greatly improved to within an order of magnitude of the CHO-secreted WT ICSM18 scFv (**Figure 7.7**). The 2.6.1 yeast displayed mutant had an apparent affinity of ~18nM which is a significantly improvement relative to the un-mutated original construct. The improvement in affinity is likely due to improved folding stability allowing the

construct to fold into the proper conformation so that allosteric affects do not inhibit binding to a poorly folded scFv.

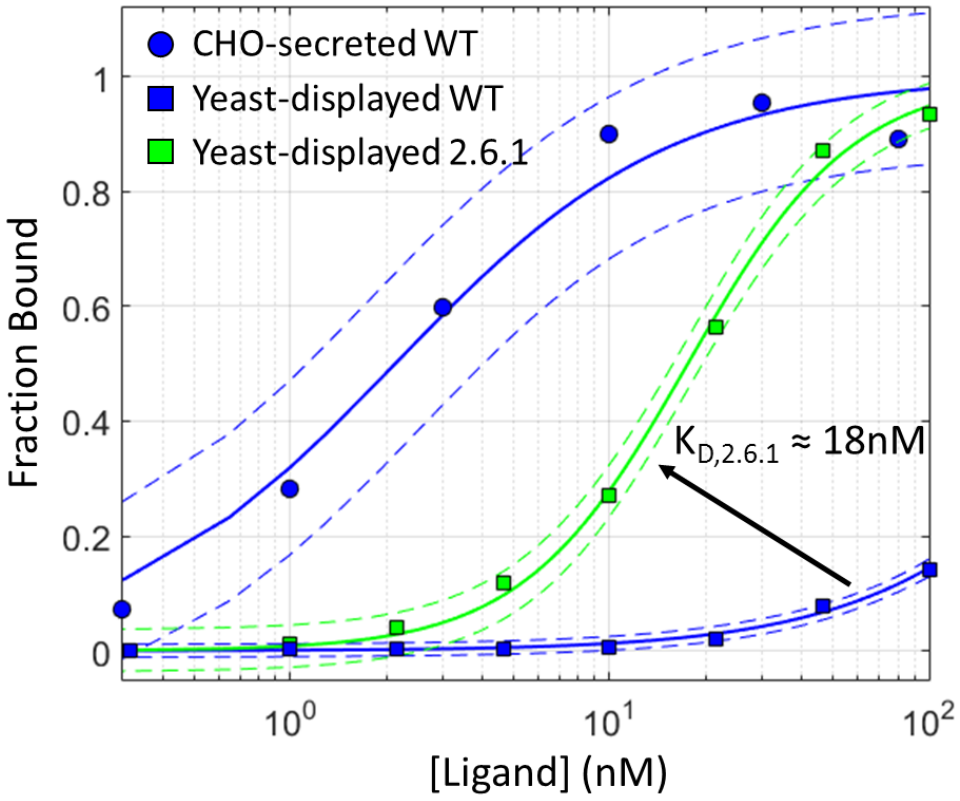


Figure 7.7 A binding titration of fluorescently-labeled soluble recombinant PrP on yeast expressing 2.6.1 (blue squares) or WT (green circles) ICSM18 scFv or a control anti-PrP scFv 6H4 (red triangles) to estimate scFv binding affinity. PrP concentration was limited by spontaneous precipitation above 100nM.

To test the selected ICSM18 scFv clone 2.6.1 in a mammalian expression system, the sequence for 2.6.1 was cloned back into the pHAGE lentiviral vector. Lentivirus was prepared for 2.6.1 and CHO cells were transduced to generate a stable

CHO 2.6.1 cell line. The 2.6.1 mutant was then tested to ensure that it was still capable of binding its prion protein antigen when expressed by mammalian cells. Mutant 2.6.1 scFv was scaled up in CHO cell culture and purified by FLAG affinity chromatography similarly to the WT scFv and titrated on the surface of N2a cells for an equilibrium binding affinity curve. We found that the binding affinity of the improved mutant 2.6.1 (~2nM) was not significantly different from the non-mutated ICSM18 scFv (~2nM) (**Figure 7.8**). The mutations introduced into the 2.6.1 variant that were capable of improving apparent affinity by yeast surface display therefore did not affect affinity of the soluble CHO-secreted form of the antibody.

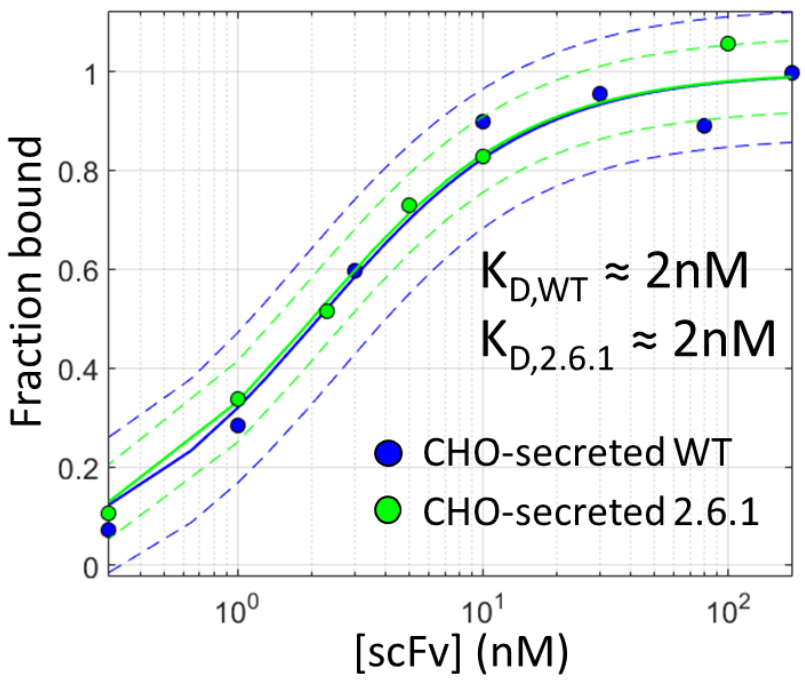


Figure 7.8 Binding affinity of purified CHO-secreted scFv was titrated on native PrP expressed on the surface of N2a mouse neuroblastoma cells. Bound scFv was detected by anti-FLAG antibody and the bound fraction was calculated by fitting to a standard binding equilibrium curve. The measured affinities of WT (blue circles) and 2.6.1 (green circles) ICSM18 scFv were around 2nM and not significantly different from each other.

Since the affinity appeared to be unchanged, we decided to test the purified scFvs for their biophysical properties to see what accounted for the significant difference by yeast display. Protein folding stability was measured by differential scanning fluorimetry [232] using the SYPRO Orange dye which undergoes a fluorescence shift upon interacting with hydrophobic regions of the protein. Samples of purified WT and 2.6.1 ICSM18 scFv were combined with SYPRO Orange and then

sample temperature was slowly ramped from 20°C to 90°C at a ramp rate of +1°C/min which fluorescence readings were taken after every Δ0.5°C interval. Both the WT and 2.6.1 scFv exhibited a single identifiable fluorescence transition indicating unfolding of the protein and suitability of DSF for measuring stability (**Figure 7.9**). The inflection point was found by taking the peak of the first derivative as a measure of melting temperature. The WT melting temperature of 44.5°C was significantly improved by Δ4.4°C in the mutant which had a melting temperature of 48.9°C ($p<10^{-3}$).

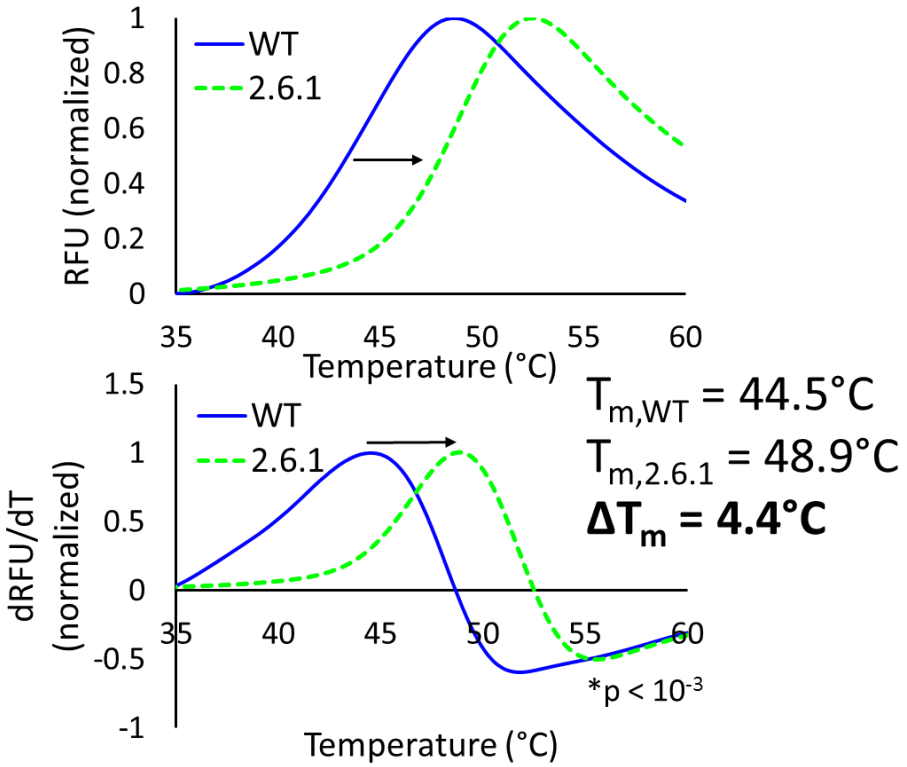


Figure 7.9 Thermal stability of purified WT and 2.6.1 ICSM18 scFv was measured by differential scanning fluorimetry (DSF) by measuring fluorescence of the hydrophobic-associating dye SYPRO Orange over a slow temperature ramp (0.5°C/30 sec) with a fluorescence reading at each Δ0.5°C interval. The scFvs had a well-defined transition from folded to unfolded states and by taking the first derivate and calculating the maximum, a meling temperature is found for each scFv, the melting temperature increased from $T_{m,WT}$ = 44.5°C to $T_{m,2.6.1}$ = 48.9°C (Δ 4.4°C, $p<10^{-3}$).

With an unchanged binding affinity for the prion protein antigen, we quantified the effect of stability improvement on specific productivity of secretion. We seeded CHO cultures of WT and 2.6.1 ICSM18 scFv and allowed them to constitutively secrete protein for equivalent amounts of time. The final concentration of scFv in solution was quantified by Western blot relative to a titration of purified and quantified control scFv. The final concentration of product was normalized by the

number of days as well as the initial and final cell densities to calculate the specific productivity. By comparison of the newly generated 2.6.1 CHO line and the highest sorted WT ICSM18 scFv CHO line for secretion of scFv, the CHO 2.6.1 cells had over 20-fold improvement in specific productivity. Since we found that productivity correlates linearly with median reporter fluorescence, we are able to calculate specific productivity relative to each integrated scFv expression cassette. Normalizing by median fluorescence of the EGFP reporter to account for differences in integration number, the improved 2.6.1 clone secretes over 150-fold better than the non-engineered ICSM18 scFv (**Figure 7.10**). The calculated specific productivity of 7.4 pcd of the mutant 2.6.1 CHO line therefore may be further improved by standard optimization methods such as removal of the reporter expression burden, increasing integrated cassette number by higher lentiviral transduction, and process optimization for seeding density, feeding schedule, and medium composition.

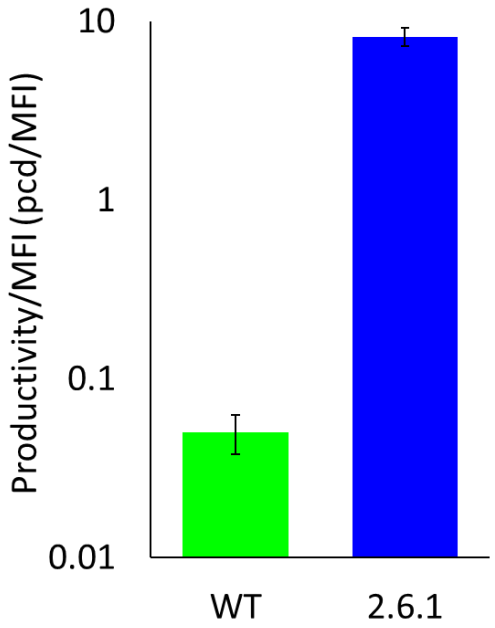


Figure 7.10 Specific productivity of CHO-secreted WT and 2.6.1 ICSM18 scFv. Comparing secretion on an absolute scale in pg/cell/day (pcd), mutant 2.6.1 performed 23-fold better than the highest-sorted WT cell line. By normalizing by reporter expression (in 10^{6}*pcd/MFI, mutant 2.6.1 performed 160-fold better per integrated cassette ($p<10^{-3}$).

7.5 Discussion

The choice of platform for engineering complex proteins has a significant effect on the efficiency of folding and secretion [217-219]. Here we show that improvement of folding stability in the eukaryotic yeast display system translates to significantly improved antibody yields in CHO cells. Previous work with yeast display demonstrated that an increase in total surface expression on the yeast correlates with an improvement in stability and secretion [233]. We found that our best-expressing mutant 2.6.1 had no detectable increase in yeast surface expression which underscores the point that increased surface expression may indicate improved stability/secretion but it is not a prerequisite for stability improvements.

While some reports show strong agreement between solubly-expressed and yeast displayed binding affinity measurements [21], we saw a significantly lower apparent affinity in the WT ICSM18 scFv by YSD while both the WT and mutant ICSM18 scFvs had similarly high affinity when secreted by mammalian cells. The effect of tethering the scFv by one of its termini or perhaps the choice of which terminus is tethered [234] may decrease apparent affinity. Therefore it is unsurprising that the mutant optimized and selected by YSD has an apparent affinity by YSD closer to that of the secreted form.

Most of the observed amino acid substitutions in the clone selected for further analysis were in the light chain and one substitution was found in the linker. Mutations in the light chain may have had a more significant effect since it was the chain tethered to the yeast surface and substitution for charged residues in the linker are known to improve stability [231]. None of the mutations were located in any of the CDRs of ICSM18 based on the crystal structure of the ICSM18 Fab [230] which explains the increase in folding stability without negatively affecting antigen-binding affinity. One of the substitutions, R92G, was located directly adjacent to the binding pocket of ICSM18, but did not have a significant effect on binding affinity when secreted by CHO cells. The R92G substitution may have instead improved potential steric hindrance of the bulky arginine group in tethered yeast display format, allowing the prion protein antigen to more easily access the ICSM18 binding site. All of the silent mutations found in the improved mutant ICSM18 scFv were for more common yeast codons which is expected because the clone was selected in yeast. Half of the silent mutations accounted for rarer codons in CHO cells, so improved expression of the mutant in CHO cells was likely due to improved stability rather than improved

translation codon usage bias. Substitution of the rarer mutated codons for more commonly used codons in CHO cells may further improve expression. It is important to note that binding affinity was not negatively affected since antigen binding was used as a selective pressure. Introducing known framework stability-enhancing mutations may have had an allosteric affect that could negatively influence antigen binding.

Aggregation propensity of the CHO-expressed ICSM18 scFvs was tested to determine whether misfolding of the WT ICSM18 scFv in yeast contributed to the decrease in apparent binding affinity (**Supp Figure B.2**). Despite the increase in stability of the mutant ICSM18 scFv, aggregation propensity was unchanged which is unsurprising since aggregation and folding stability are not necessarily correlated. Despite the de-stabilizing effect of directly truncating and linking the variable domains of an antibody into a scFv, both the WT and mutant ICSM18 scFvs were found to be stable monomers for at least several months at 4C in standard PBS.

In summary, we have engineered an anti-prion scFv with higher stability and expression in mammalian cells for further testing of its therapeutic efficacy and probing the interactions between prion proteins. In support of previous work, we found that protein stability improvements engineered in heterologous hosts are conserved and further that stability improvements of a scFv in the eukaryotic YSD system may be a viable option to improve specific productivity of expression in mammalian cells. Our results show that while antibody properties determined by YSD may not correlate exactly with secreted monomeric antibodies, it is still a useful platform to rapidly engineer and select antibodies for improved expression in mammalian cells.

Chapter 8

CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

In the course of this work, we have demonstrated a convenient way to consider deconvoluting the complex nature of neurodegenerative diseases into intrinsic and extrinsic-focused models and demonstrated the engineering of a potential therapeutic antibody. The approaches and models developed here may be applied to a range of different neurodegenerative diseases to generate culture models in human cells. These improved models may be more representative of the *in vivo* disease state to yield more predictive results for identifying disease biomarkers and screening therapeutics.

The ability to reprogram neurons has significant implications for personalized medicine and high-throughput screens. In the first section, we demonstrated the proof-of-concept approach to high-throughput neuronal reprogramming using a screening platform of bi-transgenic mouse embryonic fibroblasts (MEFs). Transgenic reports allow for real-time monitoring of reprogramming progress and since MEFs are more amenable to reprogramming, this system may be used to screen for potential positive combinations of transcription factors before translating to human somatic cells. In the process of establishing this reprogramming platform, we were able to isolate and subclone over forty transcription factors which is a larger set than any published set of neuronal reprogramming library factors. In addition to demonstrating the feasibility of high-throughput reprogramming, in the next body of work, we generated a combinatorial model of cooperative library screening to guide the selection of

experimental parameters for the most efficient coverage of potential experimental phase space. This statistical model approach may also be applied to any libraries in which multiple elements are required for a positive result such as CRISPR guide RNA libraries, silencing RNA libraries, or random genetic circuit libraries.

In the next body of work, we thoroughly surveyed the tauopathy field by identifying the major accomplishments and state of the field in experimental models from the protein level, the cell culture level, and up to *in vivo* models in mice. We showed that other approaches to studying tau fibrillization typically rely upon non-physiological factors including the use of synthetic heparin-induced fibrils, fragments of expressed tau monomer, and pro-fibrillization mutations found in hereditary frontotemperal dementia patients. We then decided to establish an improved model of tau disease strain propagation in culture by investigating the appropriate choice of fibrils and expressed monomer to best capture the representative characteristics of fibril-propagation in the brain. By comparing synthetic fibrils and disease-associated strains in culture, we found that the conformational structure of tau in the fibrils is propagated and conserved in a cellular environment which is the first demonstration of conformational propagation of disease strains in a cellular environment. Additionally, we found that similar to fibrils propagated with recombinant purified tau, synthetic heparin-induced fibrils are significantly different from disease-associated strains. Finally, we demonstrated a stark contrast in fibrillization intracellularly of full length vs truncated 4R tau with synthetic or disease-associated fibrils from patients with AD, CBD, or PiD. The synthetic fibrils were capable of fibrillizing both full length and 4R tau, but the disease-associated strains could only elongate with full length tau

monomer suggesting that the minimum region of fibrillization for synthetic and disease-associated tau fibrils may be significantly different.

In the last part of this work, we compared a mutant scFv isolated from yeast surface display mutatgenesis and sorting by a previous lab member to the wild type variant. We found that the even though yeast and mammalian cells are both eukaryotic, their expression systems are different enough that the lower stability wild type scFv was not able to fold and express properly. By yeast surface display this manifested as a significant decrease in apparent affinity and by yeast secretion the expression difference was evident by the robust formation of an irreversible improperly-folded dimer. However, stability improvements identified in the yeast system directly translated to improvements in the mammalian system which were evidenced by the drastic improvement in expression efficiency by over 150-fold for scFv cassette. This demonstrates that heterologous-host engineering is a feasible approach to improve stability and expression on a protein engineering level without compromising antibody affinity. Another important implication of these results is that measured affinity of a scFv by yeast surface display it not necessarily representative of the affinity of the secreted form in mammalian cells. Typically, the final expressed construct capable of displaying on the yeast surface is likely folded well enough to bind its antigen if it is folded well enough to bypass proteolytic degradation inside the cell. Since the energy folding landscape of proteins is complex with many local minima, a displayed proteins in a pseudo-stable partially folded state may be capable of surface expression without being properly folded. The specific productivities of the mutant 2.6.1 ICSM18 scFv may be more readily scaled up for protein interaction experiments and *in vivo* studies.

8.2 Future work

8.2.1 Cellular Reprogramming

In the course of the work completed here, many new and exciting lines of research were uncovered that may be further investigated in the future. While the methods and models may be generally useful to apply to various diseases for different applications, the specific questions raised suggest particular experimental approaches. For the reprogramming work, the proof-of-concept groundwork was established and a large transcription factor library was assembled as a useful resource for screening reprogramming combinations. Further, a combinatorial model was implemented to assist in the selection of experimental parameters for maximizing the efficiency of library screening. The next steps are to validate the model with a select number of transcription factors including a known combination of factors such as Ascl1, Brn2, and Myt1l (ABM) capable of reprogramming MEFs into glutamatergic forebrain neurons as a positive control. Model validation may be carried out by transducing MEFs with several different combinations including the known three factors, three plus an inhibitory factor REST, ABM+REST plus 6 additional factors to simulate a mid-sized library, and ABM+REST with the entire 40-factor set to simulate a large library. The multiplicities of infectivity suggested by the combinatorial model may be used to compare the resulting percentage of positive neuronal cells to the value predicted by the model to ensure that the assumptions made by the model are valid. After model validation, the next step that follows naturally is implementing the high-throughput screening platform to derive MSNs and BFCNs for establishing culture models of HD and AD. This step will require labeling the resulting neurons with more specific markers of these neuronal subtypes before sorting to make sure that the

appropriate subtype of neuron is identified. An extension of this work that may increase the possibilities of neuronal subtype or increase the speed and efficiency of reprogramming is the use of small molecule inhibitor chromatin remodelers such as HDAC-inhibitors, methyltransferase inhibitors, or DNA methyltransferase inhibitors which should theoretically increase chromatin availability for transcription factors to exert their DNA transcription effects. This high-throughput platform may then be applied to the derivation of any neuronal subtype from any starting cell type by identifying sufficient neuronal markers to confirm positive reprogramming during the iterative process.

8.2.2 Cell Culture Model of Tauopathy

We have developed an improved model of tauopathy strain fibrillization in culture which has shown that disease-associated fibrils applied to cells expressing full length normal human tau may be more representative of the faithful propagation of these disease strains. Our results have yielded some interesting observations that raise new scientific questions that may lead to more insight into the mechanisms of tauopathy propagation. A surprising result in our work was the sharp difference in fibrillization of full length tau with disease-associated strains relative to the fragment 4R tau which was unable to elongate the disease fibrils as opposed to synthetic fibrils which were able to seed both full length and truncated tau. Since this result suggests that the minimum region required to interact and fibrillize may be different between synthetic and disease strains we generated a panel of tau constructs in which different regions of tau protein surrounding the 4R region were added to see which regions are necessary. This method may also be able to determine regions capable of discriminating between different disease strains which would be useful in both

diagnostics as well as better understanding of how these disease fibrils assemble differently for different tauopthies. In an attempt to generate stable inducible HEK cell lines expressing low but consistent levels of these different constructs to decrease microscopy background, we found that low but detectable intracellular concentrations of tau were insufficient to seed fibrillization upon addition of exogenous fibrils. Upon induction of expression to increase intracellular tau concentration, fibril seeds were again able to induce fibrillization. This result suggests that there is an intermediate minimum concentration necessary for fibrillization which may vary between tauopathies which may help explain why some cells are more or less prone to tauopathy spread possibly due to intracellular tau concentrations. Once we have fully characterized different disease strains of tau fibrils in HEK cells to normalize for cell-specific effects, the next step would be to apply these strains to primary neurons to investigate differences in uptake and toxicity between strains and further by applying these strains to different subtypes of neurons we can begin to understand how cell type-specific vulnerability varies between tauopathies because of conformational differences.

8.2.3 Anti-Prion ScFv Therapeutic Candidate

The engineering of ICSM18 scFv by mutagenesis and FACS yielded a significantly higher-expressing mutant that had the same affinity but improved stability as measured by differential scanning fluorimetry. Since the mutant identified has four different amino acid substitutions, the individual contribution of each mutation or whether these mutations work together cooperatively remains to be seen and so by looking at each individual mutation in isolation we can potentially draw some conclusions about antibody engineering in general to improve the process of

engineering other unrelated proteins. The arginine adjacent to the binding pocket that was mutated into a glycine would be one potential candidate to study to see if that mutation alone would be sufficient to improve stability. A next step for the improved 2.6.1 mutant would be to perform more rounds of mutagenesis and sorting specifically in the CDRs to isolate clones with improved affinity to the prion protein. Even before improving the affinity which is already physiologically decent at around 2nM, the scale up of purification of the mutant scFv is much easier since the expression levels were improved by over two orders of magnitude which makes *in vivo* studies of safety and efficacy with prion disease mouse models much more feasible.

Taken together, the work shown here demonstrates a significant contribution to the areas of neurodegenerative disease modeling and protein engineering to better establish more representative disease models for biomarker identification and therapeutic screening as well as the improved engineering of potential protein therapeutics. This work also raises some new interesting questions to look deeper into the underlying mechanisms of reprogramming, tau strain propagation, and heterologous-host protein engineering for future contributions to these fields.

REFERENCES

1. Ross, C.A. and M.A. Poirier, *Protein aggregation and neurodegenerative disease.* Nat Med, 2004. **10 Suppl**: p. S10-7.
2. Hsu, A.L., C.T. Murphy, and C. Kenyon, *Regulation of aging and age-related disease by DAF-16 and heat-shock factor.* Science, 2003. **300**: p. 1142-1145.
3. Cui, L., et al., *Transcriptional repression of PGC-1alpha by mutant huntingtin leads to mitochondrial dysfunction and neurodegeneration.* Cell, 2006. **127**(1): p. 59-69.
4. Cohen, E. and A. Dillin, *The insulin paradox: aging, proteotoxicity and neurodegeneration.* Nat Rev Neurosci, 2008. **9**(10): p. 759-67.
5. Gan, L. and L. Mucke, *Paths of convergence: sirtuins in aging and neurodegeneration.* Neuron, 2008. **58**(1): p. 10-4.
6. Cohen, E., et al., *Reduced IGF-1 signaling delays age-associated proteotoxicity in mice.* Cell, 2009. **139**(6): p. 1157-69.
7. Mattson, M.P. and T. Magnus, *Ageing and neuronal vulnerability.* Nat Rev Neurosci, 2006. **7**(4): p. 278-94.
8. Saxena, S. and P. Caroni, *Selective neuronal vulnerability in neurodegenerative diseases: from stressor thresholds to degeneration.* Neuron, 2011. **71**(1): p. 35-48.
9. Vonsattel, J.P., et al., *Neuropathological classification of Huntington's disease.* J. Neuropathol. Exp. Neurol., 1985. **44**: p. 559-577.
10. Selkoe, D.J., *Alzheimer's disease is a synaptic failure.* Science, 2002. **298**(5594): p. 789-91.
11. Hardy, J. and D.J. Selkoe, *The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.* Science, 2002. **297**: p. 353–356.
12. Hotamisligil, G.S., *Endoplasmic reticulum stress and the inflammatory basis of metabolic disease.* Cell, 2010. **140**(6): p. 900-17.
13. Koss, D.J., et al., *Soluble pre-fibrillar tau and beta-amyloid species emerge in early human Alzheimer's disease and track disease progression and cognitive decline.* Acta Neuropathol, 2016.
14. Morozova, O.A., et al., *Conformational features of tau fibrils from Alzheimer's disease brain are faithfully propagated by unmodified recombinant protein.* Biochemistry, 2013. **52**(40): p. 6960-7.
15. Poduslo, J.F., G.L. Curran, and C.T. Berg, *Macromolecular permeability across the blood-nerve and blood-brain barriers.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5705-9.

16. Felgenhauer, K., *Protein size and cerebrospinal fluid composition.* Klin Wochenschr, 1974. **52**(24): p. 1158-64.
17. Doolan, K.M. and D.W. Colby, *Conformation-dependent epitopes recognized by prion protein antibodies probed using mutational scanning and deep sequencing.* J Mol Biol, 2015. **427**(2): p. 328-40.
18. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method.* Methods Enzymol, 2002. **350**: p. 87-96.
19. Parekh, R.N., M.R. Shaw, and K.D. Wittrup, *An integrating vector for tunable, high copy, stable integration into the dispersed Ty delta sites of Saccharomyces cerevisiae.* Biotechnol Prog, 1996. **12**(1): p. 16-21.
20. Valente, K.N., et al., *Recovery of Chinese hamster ovary host cell proteins for proteomic analysis.* Biotechnol J, 2014. **9**(1): p. 87-99.
21. Colby, D.W., et al., *Engineering antibody affinity by yeast surface display.* Methods Enzymol, 2004. **388**: p. 348-58.
22. Chao, G., et al., *Isolating and engineering human antibodies using yeast surface display.* Nat Protoc, 2006. **1**(2): p. 755-68.
23. Ma, J. and S. Lindquist, *Wild-type PrP and a mutant associated with prion disease are subject to retrograde transport and proteasome degradation.* Proc. Natl. Acad. Sci. U.S.A., 2001. **98**: p. 14955-14960.
24. Mehlhorn, I., et al., *High-level expression and characterization of a purified 142-residue polypeptide of the prion protein.* Biochemistry, 1996. **35**(17): p. 5528-37.
25. Lu, B.Y., P.J. Beck, and J.Y. Chang, *Oxidative folding of murine prion mPrP(23-231).* Eur. J. Biochem., 2001. **268**: p. 3767-3773.
26. Boivin, S., S. Kozak, and R. Meijers, *Optimization of protein purification and characterization using Thermofluor screens.* Protein Expr Purif, 2013. **91**(2): p. 192-206.
27. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.
28. Park, I.H., et al., *Disease-specific induced pluripotent stem cells.* Cell, 2008. **134**(5): p. 877-86.
29. Zhang, N., et al., *Characterization of Human Huntington's Disease Cell Model from Induced Pluripotent Stem Cells.* PLoS Curr, 2010. **2**: p. RRN1193.
30. Karumbayaram, S., et al., *Directed Differentiation of Human-Induced Pluripotent Stem Cells Generates Active Motor Neurons.* Stem Cells, 2009. **27**(4): p. 806-811.
31. Wang, A., et al., *Induced pluripotent stem cells for neural tissue engineering.* Biomaterials, 2011. **32**(22): p. 5023-32.
32. Vierbuchen, T., et al., *Direct conversion of fibroblasts to functional neurons by defined factors.* Nature, 2010. **463**(7284): p. 1035-41.

33. Pang, Z.P., et al., *Induction of human neuronal cells by defined transcription factors.* Nature, 2011.
34. Soldner, F., et al., *Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors.* Cell, 2009. **136**(5): p. 964-77.
35. Gaspard, N. and P. Vanderhaeghen, *From stem cells to neural networks: recent advances and perspectives for neurodevelopmental disorders.* Dev Med Child Neurol, 2011. **53**(1): p. 13-7.
36. Caiazzo, M., et al., *Direct generation of functional dopaminergic neurons from mouse and human fibroblasts.* Nature, 2011.
37. Masserdotti, G., S. Gascon, and M. Gotz, *Direct neuronal reprogramming: learning from and for development.* Development, 2016. **143**(14): p. 2494-510.
38. Ma, L., et al., *Human embryonic stem cell-derived GABA neurons correct locomotion deficits in quinolinic acid-lesioned mice.* Cell Stem Cell, 2012. **10**(4): p. 455-64.
39. Duan, L., et al., *Stem cell derived basal forebrain cholinergic neurons from Alzheimer's disease patients are more susceptible to cell death.* Mol Neurodegener, 2014. **9**: p. 3.
40. Yagi, T., et al., *Modeling familial Alzheimer's disease with induced pluripotent stem cells.* Hum Mol Genet, 2011. **20**(23): p. 4530-9.
41. Yahata, N., et al., *Anti-Abeta drug screening platform using human iPS cell-derived neurons for the treatment of Alzheimer's disease.* PLoS One, 2011. **6**(9): p. e25788.
42. Qiang, L., et al., *Directed conversion of Alzheimer's disease patient skin fibroblasts into functional neurons.* Cell, 2011. **146**(3): p. 359-71.
43. Israel, M.A., et al., *Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells.* Nature, 2012.
44. Bissonnette, C.J., et al., *The controlled generation of functional basal forebrain cholinergic neurons from human embryonic stem cells.* Stem Cells, 2011. **29**(5): p. 802-11.
45. Son, E.Y., et al., *Conversion of mouse and human fibroblasts into functional spinal motor neurons.* Cell Stem Cell, 2011. **9**(3): p. 205-18.
46. Pfisterer, U., et al., *Direct conversion of human fibroblasts to dopaminergic neurons.* Proc Natl Acad Sci U S A, 2011. **108**(25): p. 10343-8.
47. Yamanaka, S. and K. Takahashi, *[Induction of pluripotent stem cells from mouse fibroblast cultures]*. Tanpakushitsu Kakusan Koso, 2006. **51**(15): p. 2346-51.
48. Zhang, Y., et al., *Rapid single-step induction of functional neurons from human pluripotent stem cells.* Neuron, 2013. **78**(5): p. 785-98.
49. Davis, R.L., H. Weintraub, and A.B. Lassar, *Expression of a single transfected cDNA converts fibroblasts to myoblasts.* Cell, 1987. **51**(6): p. 987-1000.
50. Heins, N., et al., *Glial cells generate neurons: the role of the transcription factor Pax6.* Nat Neurosci, 2002. **5**(4): p. 308-15.

51. Bussmann, L.H., et al., *A robust and highly efficient immune cell reprogramming system.* Cell Stem Cell, 2009. **5**(5): p. 554-66.
52. Chanda, S., et al., *Generation of induced neuronal cells by the single reprogramming factor ASCL1.* Stem Cell Reports, 2014. **3**(2): p. 282-96.
53. Yoo, A.S., et al., *MicroRNA-mediated conversion of human fibroblasts to neurons.* Nature, 2011. **476**(7359): p. 228-31.
54. Stellmann, C. and P. Bornarel, *Titrations of viruses on cell cultures. Pharmacology or single particle hypothesis.* Arch Gesamte Virusforsch, 1972. **36**(3): p. 205-17.
55. Fehse, B., et al., *Pois(s)on--it's a question of dose.* Gene Ther, 2004. **11**(11): p. 879-81.
56. Spillantini, M.G. and M. Goedert, *Tau pathology and neurodegeneration.* The Lancet Neurology, 2013. **12**(6): p. 609-622.
57. Murray, M.E., et al., *Clinicopathologic assessment and imaging of tauopathies in neurodegenerative dementias.* Alzheimers Res Ther, 2014. **6**(1): p. 1.
58. Williams, D.R. and A.J. Lees, *How do patients with parkinsonism present? A clinicopathological study.* Intern Med J, 2009. **39**(1): p. 7-12.
59. Goedert, M., et al., *Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease.* Neuron, 1989. **3**(4): p. 519-26.
60. Weingarten, M.D., et al., *A protein factor essential for microtubule assembly.* Proc Natl Acad Sci U S A, 1975. **72**(5): p. 1858-62.
61. Cleveland, D.W., S.Y. Hwo, and M.W. Kirschner, *Physical and chemical properties of purified tau factor and the role of tau in microtubule assembly.* J Mol Biol, 1977. **116**(2): p. 227-47.
62. Morris, M., et al., *The many faces of tau.* Neuron, 2011. **70**(3): p. 410-26.
63. Hanger, D.P., et al., *Intracellular and extracellular roles for tau in neurodegenerative disease.* J Alzheimers Dis, 2014. **40 Suppl 1**: p. S37-45.
64. Schwalbe, M., et al., *Predictive atomic resolution descriptions of intrinsically disordered hTau40 and alpha-synuclein in solution from NMR and small angle scattering.* Structure, 2014. **22**(2): p. 238-49.
65. von Bergen, M., et al., *Tau aggregation is driven by a transition from random coil to beta sheet structure.* Biochim Biophys Acta, 2005. **1739**(2-3): p. 158-66.
66. Fischer, D., et al., *Conformational changes specific for pseudophosphorylation at serine 262 selectively impair binding of tau to microtubules.* Biochemistry, 2009. **48**(42): p. 10047-55.
67. Goedert, M. and R. Jakes, *Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization.* EMBO J, 1990. **9**(13): p. 4225-30.
68. Goedert, M., et al., *Tau proteins of Alzheimer paired helical filaments: abnormal phosphorylation of all six brain isoforms.* Neuron, 1992. **8**(1): p. 159-68.

69. Schmidt, M.L., et al., *Tau isoform profile and phosphorylation state in dementia pugilistica recapitulate Alzheimer's disease.* Acta Neuropathol, 2001. **101**(5): p. 518-24.
70. Noda, K., et al., *Quantitative analysis of neurofibrillary pathology in a general population to reappraise neuropathological criteria for senile dementia of the neurofibrillary tangle type (tangle-only dementia): the Hisayama Study.* Neuropathology, 2006. **26**(6): p. 508-18.
71. Flament, S., et al., *Abnormal Tau proteins in progressive supranuclear palsy. Similarities and differences with the neurofibrillary degeneration of the Alzheimer type.* Acta Neuropathol, 1991. **81**(6): p. 591-6.
72. Ksiezak-Reding, H., et al., *Ultrastructure and biochemical composition of paired helical filaments in corticobasal degeneration.* Am J Pathol, 1994. **145**(6): p. 1496-508.
73. Togo, T., N. Cookson, and D.W. Dickson, *Argyrophilic grain disease: neuropathology, frequency in a dementia brain bank and lack of relationship with apolipoprotein E.* Brain Pathol, 2002. **12**(1): p. 45-52.
74. Tolnay, M., et al., *Argyrophilic grain disease and Alzheimer's disease are distinguished by their different distribution of tau protein isoforms.* Acta Neuropathol, 2002. **104**(4): p. 425-34.
75. Delacourte, A., et al., *Specific pathological Tau protein variants characterize Pick's disease.* J Neuropathol Exp Neurol, 1996. **55**(2): p. 159-68.
76. Prusiner, S.B., *Novel proteinaceous infectious particles cause scrapie.* Science, 1982. **216**(4542): p. 136-44.
77. Safar, J., F.E. Cohen, and S.B. Prusiner, *Quantitative traits of prion strains are enciphered in the conformation of the prion protein.* Arch Virol Suppl, 2000(16): p. 227-35.
78. Legname, G., et al., *Continuum of prion protein structures enciphers a multitude of prion isolate-specified phenotypes.* Proc Natl Acad Sci U S A, 2006. **103**(50): p. 19105-10.
79. Peden, A.H. and J.W. Ironside, *Molecular pathology in neurodegenerative diseases.* Curr Drug Targets, 2012. **13**(12): p. 1548-59.
80. Goedert, M., et al., *Prion-like mechanisms in the pathogenesis of tauopathies and synucleinopathies.* Curr Neurol Neurosci Rep, 2014. **14**(11): p. 495.
81. Holmes, B.B. and M.I. Diamond, *Prion-like properties of Tau protein: the importance of extracellular Tau as a therapeutic target.* J Biol Chem, 2014. **289**(29): p. 19855-61.
82. Clavaguera, F., et al., *Prion-like transmission and spreading of tau pathology.* Neuropathol Appl Neurobiol, 2014.
83. Seeley, W.W., et al., *Neurodegenerative diseases target large-scale human brain networks.* Neuron, 2009. **62**(1): p. 42-52.
84. Zhou, J., et al., *Predicting regional neurodegeneration from the healthy brain functional connectome.* Neuron, 2012. **73**(6): p. 1216-27.

85. Raj, A., A. Kuceyeski, and M. Weiner, *A network diffusion model of disease progression in dementia.* Neuron, 2012. **73**(6): p. 1204-15.
86. Prusiner, S.B., *Some speculations about prions, amyloid, and Alzheimer's disease.* N Engl J Med, 1984. **310**(10): p. 661-3.
87. Lee, V.M., M. Goedert, and J.Q. Trojanowski, *Neurodegenerative tauopathies.* Annu Rev Neurosci, 2001. **24**: p. 1121-59.
88. Hutton, M., et al., *Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17.* Nature, 1998. **393**(6686): p. 702-5.
89. Wolfe, M.S., *Tau mutations in neurodegenerative diseases.* J Biol Chem, 2009. **284**(10): p. 6021-5.
90. Novak, P., M. Prcina, and E. Kontsekova, *Tauons and prions: infamous cousins?* J Alzheimers Dis, 2011. **26**(3): p. 413-30.
91. Braak, H. and E. Braak, *Neuropathological stageing of Alzheimer-related changes.* Acta Neuropathol, 1991. **82**(4): p. 239-59.
92. Feany, M.B. and D.W. Dickson, *Neurodegenerative disorders with extensive tau pathology: a comparative study and review.* Ann Neurol, 1996. **40**(2): p. 139-48.
93. Frost, B., et al., *Conformational diversity of wild-type Tau fibrils specified by templated conformation change.* J Biol Chem, 2009. **284**(6): p. 3546-51.
94. Liu, L., et al., *Trans-synaptic spread of tau pathology in vivo.* PLoS One, 2012. **7**(2): p. e31302.
95. de Calignon, A., et al., *Propagation of tau pathology in a model of early Alzheimer's disease.* Neuron, 2012. **73**(4): p. 685-97.
96. Clavaguera, F., et al., *Transmission and spreading of tauopathy in transgenic mouse brain.* Nat Cell Biol, 2009. **11**(7): p. 909-13.
97. Iba, M., et al., *Synthetic tau fibrils mediate transmission of neurofibrillary tangles in a transgenic mouse model of Alzheimer's-like tauopathy.* J Neurosci, 2013. **33**(3): p. 1024-37.
98. Clavaguera, F., et al., *Brain homogenates from human tauopathies induce tau inclusions in mouse brain.* Proc Natl Acad Sci U S A, 2013. **110**(23): p. 9535-40.
99. Sanders, D.W., et al., *Distinct tau prion strains propagate in cells and mice and define different tauopathies.* Neuron, 2014. **82**(6): p. 1271-88.
100. Collinge, J. and A.R. Clarke, *A general model of prion strains and their pathogenicity.* Science, 2007. **318**(5852): p. 930-936.
101. Irwin, D.J., et al., *Acetylated tau neuropathology in sporadic and hereditary tauopathies.* Am J Pathol, 2013. **183**(2): p. 344-51.
102. Clavaguera, F., et al., *Peripheral administration of tau aggregates triggers intracerebral tauopathy in transgenic mice.* Acta Neuropathol, 2014. **127**(2): p. 299-301.
103. Falcon, B., et al., *Conformation determines the seeding potencies of native and recombinant tau aggregates.* J Biol Chem, 2015. **290**(2): p. 1049-65.

104. Boluda, S., et al., *Differential induction and spread of tau pathology in young PS19 tau transgenic mice following intracerebral injections of pathological tau from Alzheimer's disease or corticobasal degeneration brains.* Acta Neuropathol, 2015. **129**(2): p. 221-37.
105. Tanaka, M., et al., *Conformational variations in an infectious protein determine prion strain differences.* Nature, 2004. **428**(6980): p. 323-8.
106. Klimova, N., N. Makarava, and I.V. Baskakov, *The diversity and relationship of prion protein self-replicating states.* Virus Res, 2014.
107. Kretzschmar, H. and J. Tatzelt, *Prion disease: a tale of folds and strains.* Brain Pathol, 2013. **23**(3): p. 321-32.
108. Sano, K., et al., *Conformational properties of prion strains can be transmitted to recombinant prion protein fibrils in real-time quaking-induced conversion.* J Virol, 2014. **88**(20): p. 11791-801.
109. Kidd, M., *Paired helical filaments in electron microscopy of Alzheimer's disease.* Nature, 1963. **197**: p. 192-3.
110. Ksiezak-Reding, H., K. Morgan, and D.W. Dickson, *Tau immunoreactivity and SDS solubility of two populations of paired helical filaments that differ in morphology.* Brain Res, 1994. **649**(1-2): p. 185-96.
111. Crowther, R.A. and M. Goedert, *Abnormal tau-containing filaments in neurodegenerative diseases.* J Struct Biol, 2000. **130**(2-3): p. 271-9.
112. Zhukareva, V., et al., *Biochemical analysis of tau proteins in argyrophilic grain disease, Alzheimer's disease, and Pick's disease : a comparative study.* Am J Pathol, 2002. **161**(4): p. 1135-41.
113. Schweers, O., et al., *Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure.* J Biol Chem, 1994. **269**(39): p. 24290-7.
114. Wolynes, P.G., J.N. Onuchic, and D. Thirumalai, *Navigating the folding routes.* Science, 1995. **267**(5204): p. 1619-20.
115. Dobson, C.M., *Protein folding and misfolding.* Nature, 2003. **426**(6968): p. 884-90.
116. Selkoe, D.J., *Folding proteins in fatal ways.* Nature, 2003. **426**(6968): p. 900-4.
117. Wille, H., et al., *Alzheimer-like paired helical filaments and antiparallel dimers formed from microtubule-associated protein tau in vitro.* J Cell Biol, 1992. **118**(3): p. 573-84.
118. Crowther, R.A., et al., *Assembly of Alzheimer-like filaments from full-length tau protein.* FEBS Lett, 1994. **337**(2): p. 135-8.
119. Goedert, M., et al., *Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans.* Nature, 1996. **383**(6600): p. 550-3.
120. Perez, M., et al., *Polymerization of tau into filaments in the presence of heparin: the minimal sequence required for tau-tau interaction.* J Neurochem, 1996. **67**(3): p. 1183-90.

121. Kampers, T., et al., *RNA stimulates aggregation of microtubule-associated protein tau into Alzheimer-like paired helical filaments.* FEBS Lett, 1996. **399**(3): p. 344-9.
122. Wilson, D.M. and L.I. Binder, *Free fatty acids stimulate the polymerization of tau and amyloid beta peptides. In vitro evidence for a common effector of pathogenesis in Alzheimer's disease.* Am J Pathol, 1997. **150**(6): p. 2181-95.
123. von Bergen, M., et al., *Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5129-34.
124. Friedhoff, P., et al., *Rapid assembly of Alzheimer-like paired helical filaments from microtubule-associated protein tau monitored by fluorescence in solution.* Biochemistry, 1998. **37**(28): p. 10223-30.
125. Barghorn, S. and E. Mandelkow, *Toward a unified scheme for the aggregation of tau into Alzheimer paired helical filaments.* Biochemistry, 2002. **41**(50): p. 14885-96.
126. Chirita, C.N., et al., *Triggers of full-length tau aggregation: a role for partially folded intermediates.* Biochemistry, 2005. **44**(15): p. 5862-72.
127. Ramachandran, G. and J.B. Udgaonkar, *Understanding the kinetic roles of the inducer heparin and of rod-like protofibrils during amyloid fibril formation by Tau protein.* J Biol Chem, 2011. **286**(45): p. 38948-59.
128. Siddiqua, A., et al., *Conformational basis for asymmetric seeding barrier in filaments of three- and four-repeat tau.* J Am Chem Soc, 2012. **134**(24): p. 10271-8.
129. Dinkel, P.D., et al., *Variations in filament conformation dictate seeding barrier between three- and four-repeat tau.* Biochemistry, 2011. **50**(20): p. 4330-6.
130. Furukawa, Y., K. Kaneko, and N. Nukina, *Tau protein assembles into isoform- and disulfide-dependent polymorphic fibrils with distinct structural properties.* J Biol Chem, 2011. **286**(31): p. 27236-46.
131. Carlson, S.W., et al., *A complex mechanism for inducer mediated tau polymerization.* Biochemistry, 2007. **46**(30): p. 8838-49.
132. Wischik, C.M., C.R. Harrington, and J.M. Storey, *Tau-aggregation inhibitor therapy for Alzheimer's disease.* Biochem Pharmacol, 2014. **88**(4): p. 529-39.
133. Tenreiro, S., K. Eckermann, and T.F. Outeiro, *Protein phosphorylation in neurodegeneration: friend or foe?* Front Mol Neurosci, 2014. **7**: p. 42.
134. Noble, W., et al., *The importance of tau phosphorylation for neurodegenerative diseases.* Front Neurol, 2013. **4**: p. 83.
135. Frost, B., R.L. Jacks, and M.I. Diamond, *Propagation of tau misfolding from the outside to the inside of a cell.* J Biol Chem, 2009. **284**(19): p. 12845-52.
136. Nonaka, T., et al., *Seeded aggregation and toxicity of {alpha}-synuclein and tau: cellular models of neurodegenerative diseases.* J Biol Chem, 2010. **285**(45): p. 34885-98.

137. Guo, J.L. and V.M. Lee, *Seeding of normal Tau by pathological Tau conformers drives pathogenesis of Alzheimer-like tangles.* J Biol Chem, 2011. **286**(17): p. 15317-31.
138. Santa-Maria, I., et al., *Paired helical filaments from Alzheimer disease brain induce intracellular accumulation of Tau protein in aggresomes.* J Biol Chem, 2012. **287**(24): p. 20522-33.
139. Guo, J.L. and V.M. Lee, *Neurofibrillary tangle-like tau pathology induced by synthetic tau fibrils in primary neurons over-expressing mutant tau.* FEBS Lett, 2013. **587**(6): p. 717-23.
140. Holmes, B.B., et al., *Heparan sulfate proteoglycans mediate internalization and propagation of specific proteopathic seeds.* Proc Natl Acad Sci U S A, 2013. **110**(33): p. E3138-47.
141. Wu, J.W., et al., *Small misfolded Tau species are internalized via bulk endocytosis and anterogradely and retrogradely transported in neurons.* J Biol Chem, 2013. **288**(3): p. 1856-70.
142. Allen, B., et al., *Abundant tau filaments and nonapoptotic neurodegeneration in transgenic mice expressing human P301S tau protein.* J Neurosci, 2002. **22**(21): p. 9340-51.
143. Probst, A., et al., *Axonopathy and amyotrophy in mice transgenic for human four-repeat tau protein.* Acta Neuropathol, 2000. **99**(5): p. 469-81.
144. Ahmed, Z., et al., *A novel in vivo model of tau propagation with rapid and progressive neurofibrillary tangle pathology: the pattern of spread is determined by connectivity, not proximity.* Acta Neuropathol, 2014. **127**(5): p. 667-83.
145. Peeraer, E., et al., *Intracerebral injection of preformed synthetic tau fibrils initiates widespread tauopathy and neuronal loss in the brains of tau transgenic mice.* Neurobiol Dis, 2015. **73**: p. 83-95.
146. Lasagna-Reeves, C.A., et al., *Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau.* Sci Rep, 2012. **2**: p. 700.
147. De Genst, E., A. Messer, and C.M. Dobson, *Antibodies and protein misfolding: From structural research tools to therapeutic strategies.* Biochim Biophys Acta, 2014. **1844**(11): p. 1907-1919.
148. Chai, X., et al., *Passive immunization with anti-Tau antibodies in two transgenic models: reduction of Tau pathology and delay of disease progression.* J Biol Chem, 2011. **286**(39): p. 34457-67.
149. Yanamandra, K., et al., *Anti-tau antibodies that block tau aggregate seeding in vitro markedly decrease pathology and improve cognition in vivo.* Neuron, 2013. **80**(2): p. 402-14.
150. Yanamandra, K., et al., *Anti-tau antibody reduces insoluble tau and decreases brain atrophy.* Annals of Clinical and Translational Neurology, 2015: p. n/a-n/a.

151. Kosik, K.S., et al., *Tau in situ hybridization in normal and Alzheimer brain: localization in the somatodendritic compartment.* Ann Neurol, 1989. **26**(3): p. 352-61.
152. Pooler, A.M., et al., *Physiological release of endogenous tau is stimulated by neuronal activity.* EMBO Rep, 2013. **14**(4): p. 389-94.
153. Yamada, K., et al., *Neuronal activity regulates extracellular tau in vivo.* J Exp Med, 2014. **211**(3): p. 387-93.
154. Tapiola, T., et al., *Cerebrospinal Fluid beta-Amyloid 42 and Tau Proteins as Biomarkers of Alzheimer-Type Pathologic Changes in the Brain.* Archives of Neurology, 2009. **66**(3): p. 382-389.
155. Mohamed, N.V., et al., *Spreading of tau pathology in Alzheimer's disease by cell-to-cell transmission.* Eur J Neurosci, 2013. **37**(12): p. 1939-48.
156. Nickel, W. and C. Rabouille, *Mechanisms of regulated unconventional protein secretion.* Nat Rev Mol Cell Biol, 2009. **10**(2): p. 148-55.
157. Kfoury, N., et al., *Trans-cellular propagation of Tau aggregation by fibrillar species.* J Biol Chem, 2012. **287**(23): p. 19440-51.
158. Dujardin, S., et al., *Neuron-to-neuron wild-type Tau protein transfer through a trans-synaptic mechanism: relevance to sporadic tauopathies.* Acta Neuropathol Commun, 2014. **2**: p. 14.
159. Kim, W., et al., *Interneuronal transfer of human tau between Lamprey central neurons in situ.* J Alzheimers Dis, 2010. **19**(2): p. 647-64.
160. Plouffe, V., et al., *Hyperphosphorylation and cleavage at D421 enhance tau secretion.* PLoS One, 2012. **7**(5): p. e36873.
161. Chai, X., J.L. Dage, and M. Citron, *Constitutive secretion of tau protein by an unconventional mechanism.* Neurobiol Dis, 2012. **48**(3): p. 356-66.
162. Saman, S., et al., *Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease.* J Biol Chem, 2012. **287**(6): p. 3842-9.
163. Holmes, B.B., et al., *Proteopathic tau seeding predicts tauopathy in vivo.* Proc Natl Acad Sci U S A, 2014. **111**(41): p. E4376-85.
164. Lim, J.P. and P.A. Gleeson, *Macropinocytosis: an endocytic pathway for internalising large gulps.* Immunol Cell Biol, 2011. **89**(8): p. 836-43.
165. Snow, A.D., et al., *The presence of heparan sulfate proteoglycans in the neuritic plaques and congophilic angiopathy in Alzheimer's disease.* Am J Pathol, 1988. **133**(3): p. 456-63.
166. Mercer, J. and A. Helenius, *Virus entry by macropinocytosis.* Nat Cell Biol, 2009. **11**(5): p. 510-20.
167. Guo, J.L. and V.M. Lee, *Cell-to-cell transmission of pathogenic proteins in neurodegenerative diseases.* Nat Med, 2014. **20**(2): p. 130-8.
168. Bretteville, A. and E. Planel, *Tau aggregates: toxic, inert, or protective species?* J Alzheimers Dis, 2008. **14**(4): p. 431-6.
169. Price, J.L. and J.C. Morris, *Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease.* Ann Neurol, 1999. **45**(3): p. 358-68.

170. Gomez-Isla, T., et al., *Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease.* Ann Neurol, 1997. **41**(1): p. 17-24.
171. Santacruz, K., et al., *Tau suppression in a neurodegenerative mouse model improves memory function.* Science, 2005. **309**(5733): p. 476-81.
172. Spires, T.L., et al., *Region-specific dissociation of neuronal loss and neurofibrillary pathology in a mouse model of tauopathy.* Am J Pathol, 2006. **168**(5): p. 1598-607.
173. Sydow, A., et al., *Tau-induced defects in synaptic plasticity, learning, and memory are reversible in transgenic mice after switching off the toxic Tau mutant.* J Neurosci, 2011. **31**(7): p. 2511-25.
174. Kimura, T., et al., *Hyperphosphorylated tau in parahippocampal cortex impairs place learning in aged mice expressing wild-type human tau.* EMBO J, 2007. **26**(24): p. 5143-52.
175. Taniguchi, T., et al., *Transgenic mice expressing mutant (N279K) human tau show mutation dependent cognitive deficits without neurofibrillary tangle formation.* FEBS Lett, 2005. **579**(25): p. 5704-12.
176. Wittmann, C.W., et al., *Tauopathy in Drosophila: neurodegeneration without neurofibrillary tangles.* Science, 2001. **293**(5530): p. 711-4.
177. Yoshiyama, Y., et al., *Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model.* Neuron, 2007. **53**(3): p. 337-51.
178. Chafekar, S.M., et al., *Pharmacological tuning of heat shock protein 70 modulates polyglutamine toxicity and aggregation.* ACS Chem Biol, 2012. **7**(9): p. 1556-64.
179. Stein, K.C. and H.L. True, *Structural variants of yeast prions show conformer-specific requirements for chaperone activity.* Mol Microbiol, 2014. **93**(6): p. 1156-71.
180. Sweeny, E.A., et al., *The Hsp104 N-Terminal Domain Enables Disaggregase Plasticity and Potentiation.* Mol Cell, 2015.
181. Jaworski, T., S. Kugler, and F. Van Leuven, *Modeling of tau-mediated synaptic and neuronal degeneration in Alzheimer's disease.* Int J Alzheimers Dis, 2010. **2010**.
182. Trojanowski, J.Q. and V.M. Lee, *Pathological tau: a loss of normal function or a gain in toxicity?* Nat Neurosci, 2005. **8**(9): p. 1136-7.
183. Morales, I., et al., *Tau oligomers and fibrils induce activation of microglial cells.* J Alzheimers Dis, 2013. **37**(4): p. 849-56.
184. Amadoro, G., et al., *NMDA receptor mediates tau-induced neurotoxicity by calpain and ERK/MAPK activation.* Proc Natl Acad Sci U S A, 2006. **103**(8): p. 2892-7.
185. Gomez-Ramos, A., et al., *Extracellular tau promotes intracellular calcium increase through M1 and M3 muscarinic receptors in neuronal cells.* Mol Cell Neurosci, 2008. **37**(4): p. 673-81.

186. Stohr, J., et al., *Distinct synthetic Abeta prion strains producing different amyloid deposits in bigenic mice.* Proc Natl Acad Sci U S A, 2014. **111**(28): p. 10329-34.
187. Arosio, P., et al., *Quantification of the concentration of Abeta42 propagons during the lag phase by an amyloid chain reaction assay.* J Am Chem Soc, 2014. **136**(1): p. 219-25.
188. Bulic, B., et al., *Tau protein and tau aggregation inhibitors.* Neuropharmacology, 2010. **59**(4-5): p. 276-89.
189. Crowe, A., et al., *Identification of aminothienopyridazine inhibitors of tau assembly by quantitative high-throughput screening.* Biochemistry, 2009. **48**(32): p. 7732-45.
190. Taniguchi, S., et al., *Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins.* J Biol Chem, 2005. **280**(9): p. 7614-23.
191. Brody, D.L. and D.M. Holtzman, *Active and passive immunotherapy for neurodegenerative disorders.* Annu Rev Neurosci, 2008. **31**: p. 175-93.
192. Boutajangout, A., D. Quartermain, and E.M. Sigurdsson, *Immunotherapy targeting pathological tau prevents cognitive decline in a new tangle mouse model.* J Neurosci, 2010. **30**(49): p. 16559-66.
193. Bi, M., et al., *Tau-targeted immunization impedes progression of neurofibrillary histopathology in aged P301L tau transgenic mice.* PLoS One, 2011. **6**(12): p. e26860.
194. Sigurdsson, E.M., *Immunotherapy targeting pathological tau protein in Alzheimer's disease and related tauopathies.* J Alzheimers Dis, 2008. **15**(2): p. 157-68.
195. Kontsekova, E., et al., *First-in-man tau vaccine targeting structural determinants essential for pathological tau-tau interaction reduces tau oligomerisation and neurofibrillary degeneration in an Alzheimer's disease model.* Alzheimers Res Ther, 2014. **6**(4): p. 44.
196. Clarke, G., et al., *A one-hit model of cell death in inherited neuronal degenerations.* Nature, 2000. **406**(6792): p. 195-9.
197. Dobson, C.M., *Principles of protein folding, misfolding and aggregation.* Semin Cell Dev Biol, 2004. **15**(1): p. 3-16.
198. Lansbury, P.T., Jr., *Structural neurology: are seeds at the root of neuronal degeneration?* Neuron, 1997. **19**(6): p. 1151-4.
199. Kidd, M., *Paired helical filaments in electron microscopy in Alzheimer's disease.* Nature, 1963. **197**: p. 192-193.
200. Takeda, S., et al., *Neuronal uptake and propagation of a rare phosphorylated high-molecular-weight tau derived from Alzheimer's disease brain.* Nat Commun, 2015. **6**: p. 8490.
201. Takahashi, M., et al., *Extracellular association of APP and tau fibrils induces intracellular aggregate formation of tau.* Acta Neuropathol, 2015. **129**(6): p. 895-907.

202. Stancu, I.C., et al., *Templated misfolding of Tau by prion-like seeding along neuronal connections impairs neuronal network function and associated behavioral outcomes in Tau transgenic mice.* Acta Neuropathol, 2015. **129**(6): p. 875-94.
203. Chiu, M.L. and G.L. Gilliland, *Engineering antibody therapeutics.* Curr Opin Struct Biol, 2016. **38**: p. 163-73.
204. Huston, J.S., et al., *Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli.* Proc Natl Acad Sci U S A, 1988. **85**(16): p. 5879-83.
205. Glockshuber, R., et al., *A comparison of strategies to stabilize immunoglobulin Fv-fragments.* Biochemistry, 1990. **29**(6): p. 1362-7.
206. Worn, A. and A. Pluckthun, *Stability engineering of antibody single-chain Fv fragments.* Journal of Molecular Biology, 2001. **305**(5): p. 989-1010.
207. Rothlisberger, D., A. Honegger, and A. Pluckthun, *Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability.* J Mol Biol, 2005. **347**(4): p. 773-89.
208. Jager, V., et al., *High level transient production of recombinant antibodies and antibody fusion proteins in HEK293 cells.* BMC Biotechnol, 2013. **13**: p. 52.
209. Lim, A.W., et al., *Directed evolution of human scFvs in DT40 cells.* Protein Eng Des Sel, 2016. **29**(2): p. 39-48.
210. Graff, C.P., et al., *Directed evolution of an anti-carcinoembryonic antigen scFv with a 4-day monovalent dissociation half-time at 37{degrees}C.* Protein Eng Des Sel, 2004. **17**(4): p. 293-304.
211. Luginbuhl, B., et al., *Directed evolution of an anti-prion protein scFv fragment to an affinity of 1 pM and its structural interpretation.* J. Mol. Biol., 2006. **363**: p. 75-97.
212. Ferrer-Miralles, N., et al., *Microbial factories for recombinant pharmaceuticals.* Microb Cell Fact, 2009. **8**: p. 17.
213. Maccani, A., et al., *Pichia pastoris secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins.* Biotechnol J, 2014. **9**(4): p. 526-37.
214. Lai, T., Y. Yang, and S.K. Ng, *Advances in Mammalian cell line development technologies for recombinant protein production.* Pharmaceuticals (Basel), 2013. **6**(5): p. 579-603.
215. Boder, E.T. and K.D. Wittrup, *Yeast surface display for screening combinatorial polypeptide libraries.* Nat Biotechnol, 1997. **15**(6): p. 553-7.
216. Gai, S.A. and K.D. Wittrup, *Yeast surface display for protein engineering and characterization.* Curr Opin Struct Biol, 2007. **17**(4): p. 467-73.
217. Schaefer, J.V. and A. Pluckthun, *Transfer of engineered biophysical properties between different antibody formats and expression systems.* Protein Eng Des Sel, 2012. **25**(10): p. 485-506.

218. Demarest, S.J., et al., *Engineering stability into Escherichia coli secreted Fabs leads to increased functional expression.* Protein Eng Des Sel, 2006. **19**(7): p. 325-36.
219. Lombana, T.N., et al., *Optimizing antibody expression by using the naturally occurring framework diversity in a live bacterial antibody display system.* Sci Rep, 2015. **5**: p. 17488.
220. White, A.R., et al., *Monoclonal antibodies inhibit prion replication and delay the development of prion disease.* Nature, 2003. **422**: p. 80-83.
221. Klohn, P.C., et al., *PrP antibodies do not trigger mouse hippocampal neuron apoptosis.* Science, 2012. **335**(6064): p. 52.
222. Reimann, R.R., et al., *Differential Toxicity of Antibodies to the Prion Protein.* PLoS Pathog, 2016. **12**(1): p. e1005401.
223. Colby, D.W. and S.B. Prusiner, *Prions.* Cold Spring Harb Perspect Biol, 2011. **3**(1): p. a006833.
224. Enari, M., E. Flechsig, and C. Weissmann, *Scrapie prion protein accumulation by scrapie-infected neuroblastoma cells abrogated by exposure to a prion protein antibody.* Proc Natl Acad Sci U S A, 2001. **98**(16): p. 9295-9.
225. Peretz, D., et al., *Antibodies inhibit prion propagation and clear cell cultures of prion infectivity.* Nature, 2001. **412**: p. 739-743.
226. Donofrio, G., et al., *Paracrine inhibition of prion propagation by anti-PrP single-chain Fv miniantibodies.* Journal of virology, 2005. **79**(13): p. 8330-8338.
227. Campana, V., et al., *Development of antibody fragments for immunotherapy of prion diseases.* Biochem J, 2009. **418**(3): p. 507-15.
228. Song, C.H., et al., *Effect of intraventricular infusion of anti-prion protein monoclonal antibodies on disease progression in prion-infected mice.* J Gen Virol, 2008. **89**(Pt 6): p. 1533-44.
229. Fujita, K., et al., *Effects of a brain-engraftable microglial cell line expressing anti-prion scFv antibodies on survival times of mice infected with scrapie prions.* Cell Mol Neurobiol, 2011. **31**(7): p. 999-1008.
230. Antonyuk, S.V., et al., *Crystal structure of human prion protein bound to a therapeutic antibody.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(8): p. 2554-2558.
231. Chen, X., J.L. Zaro, and W.C. Shen, *Fusion protein linkers: property, design and functionality.* Adv Drug Deliv Rev, 2013. **65**(10): p. 1357-69.
232. Malik, K., et al., *Differential scanning fluorimetry: rapid screening of formulations that promote the stability of reference preparations.* J Pharm Biomed Anal, 2013. **77**: p. 163-6.
233. Shusta, E.V., et al., *Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency.* J Mol Biol, 1999. **292**(5): p. 949-56.
234. Pepper, L.R., et al., *A decade of yeast surface display technology: where are we now?* Comb Chem High Throughput Screen, 2008. **11**(2): p. 127-34.

Appendix A

DERIVATION OF LIBRARY SCREENING LIKELIHOOD EQUATION

Lemma 0.1. *Suppose that X is a set with cardinality x and $Y, Z \subseteq X$ have cardinalities y and z, respectively. If the set of strings (ordered lists with repitition) of length k whose entries come from X is uniformly distributed then the probability of selecting a string that contains every element of Y (at least once) and no elements of Z is*

$$P^Y(X-Z) = \begin{cases} \frac{\sum_{i=0}^{y}(-1)^i\binom{y}{i}(x-z-i)^k}{x^k} & \text{if } Y \cap Z = \emptyset, \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Proof. If $Y \cap Z \neq \emptyset$ then it is immediate that $P^Y(X-Z) = 0$, so assume that $Y \cap Z = \emptyset$ and suppose that we wish to construct a string of length k from the set of letters $X = \{x_i\}_{i=1}^{x}$, with repitition allowed. It is not difficult to convince ourselves that there are x^k such strings possible by considering that there are x selections for the first entry, x selections for the second, and so on. Since the choice of one letter doesn't affect any of the other choices we can use the multiplication rule to achieve our desired result.

Let us denote the set of all possible strings whose entries are from $X - Z$ by E. Now suppose we set $E_i = \{d \in E | x_i \notin d\}$, the set of strings not containing x_i. By simply pretending that there is one fewer letter to choose from, we can conclude that there must be $(x-z-1)^k$ such strings possible. By relabeling X let us think of $Y = \{x_i\}_{i=1}^{y}$.

The set of all strings in the letters of $X - Z$ where any of the letters of Y could be missing is the union

$$\bigcup_{i=1}^{y} E_i.$$

The generalized principle of inclusion-exclusion states that

$$\left|\bigcup_{i=1}^{y} E_i\right| = \sum_{i=1}^{y}(-1)^{i+1}\left(\sum_{1\leq j_1<j_2<\ldots<j_i\leq y} |E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_i}|\right),$$

where the vertical bars denote cardinality. So the first step is to compute the cardinality of $|E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_i}|$ for $1 \leq j_1 < j_2 < \ldots < j_i \leq y$. Since we can select any element of $X - Z$ except $x_{j_1}, x_{j_2}, \ldots, x_{j_i}$ then the number of selections for each entry is $(x - z - i)$. Since there are k selections and the selections are independent of one another, there are $(x-z-i)^k$ total strings in $|E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_i}|$.

Now if we note that the actual choices of $j_1, j_2, \ldots, j_i$ didn't affect the number of strings, only the number of $j's$ did then we can conclude that

$$|E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_i}| = |E_{t_1} \cap E_{t_2} \cap \cdots \cap E_{t_i}|$$

for any other choice of subscripts. For example, if $i = 3$ then $|E_1 \cap E_3 \cap E_4| = |E_1 \cap E_2 \cap E_4|$. So

$$\begin{aligned}\sum_{1 \le j_1 < j_2 < \ldots < j_i \le y} |E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_i}| &= \sum_{1 \le j_1 < j_2 < \ldots < j_i \le y} |E_1 \cap E_2 \cap \cdots E_i| \\ &= (x - z - i)^k \sum_{1 \le j_1 < j_2 < \ldots < j_i \le y} 1 \\ &= \binom{y}{i}(x - z - i)^k,\end{aligned}$$

since there are $\binom{y}{i}$ such choices of numbers $1 \le j_1 < j_2 < \ldots < j_i \le y$.

Given this, then the formula becomes

$$\left|\bigcup_{i=1}^{y} E_i\right| = \sum_{i=1}^{y} (-1)^{i+1} \binom{y}{i} (x - z - i)^k.$$

However, if we want to count the number of strings in E that *must* contain the letters of Y then we simply count *all* of the possible strings and subtract the number that are missing one or more letters from Y:

$$|E| - \left|\bigcup_{i=1}^{y} E_i\right| = (x - z)^k - \sum_{i=1}^{y} (-1)^{i+1} \binom{y}{i} (x - z - i)^k = \sum_{i=0}^{y} (-1)^i \binom{y}{i} (x - z - i)^k.$$

If we were to select a random string and all strings were equally likely to be selected then the probability of selecting a string that must contain all of the letters of $Y = \{x_i\}_{i=1}^{y}$ (at least once) is

$$\frac{\sum_{i=0}^{y} (-1)^i \binom{y}{i} (x - z - i)^k}{x^k}, \tag{2}$$

which concludes our proof. □

Appendix B

SUPPLEMENTAL INFO FOR HETEROLOGOUS-HOST ENGINEREERING OF AN ANTI-PRION PROTEIN SCFV FOR IMPROVED STABILITY AND EXPRESSION

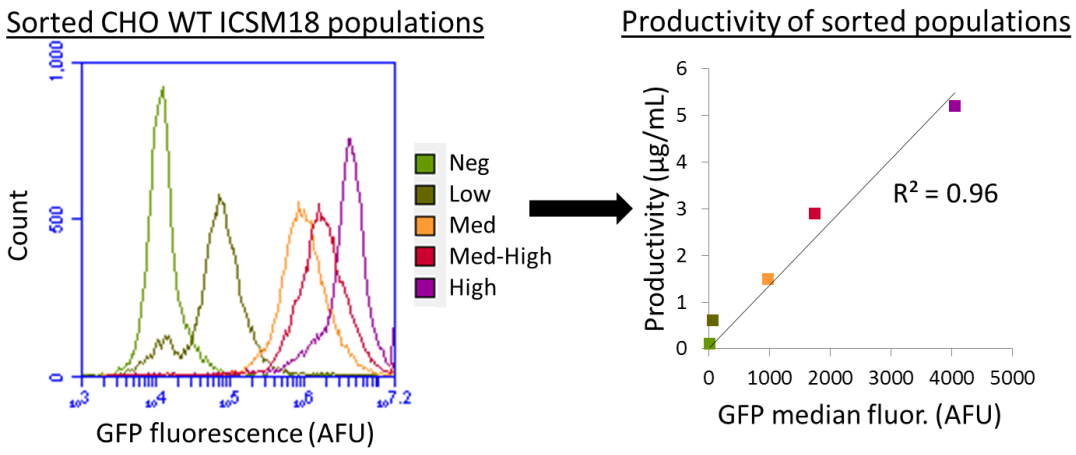


Figure B.1 CHO cells transduced with WT ICSM18 scFv lentivirus were sorted by fluorescence-activated cell sorting (FACS) for different levels of EGFP reporter fluorescence. The sorted sub-populations were tested for productivity by Western blot and the productivity was found with be linear with respect to reporter fluorescence.

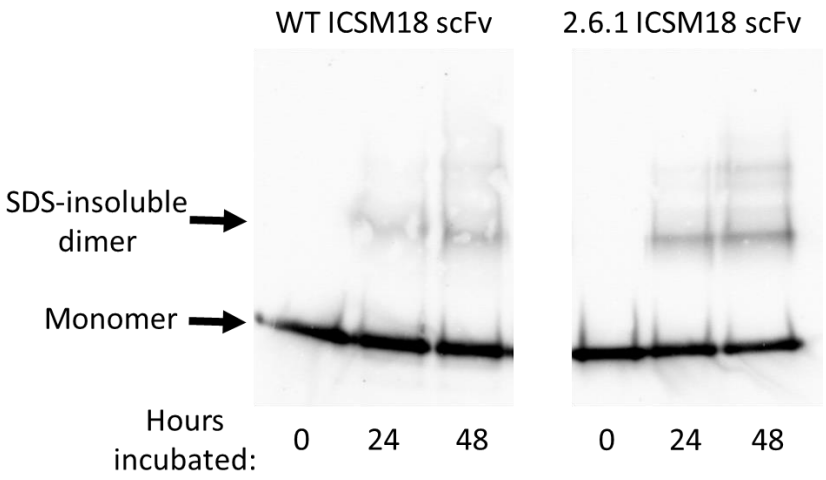
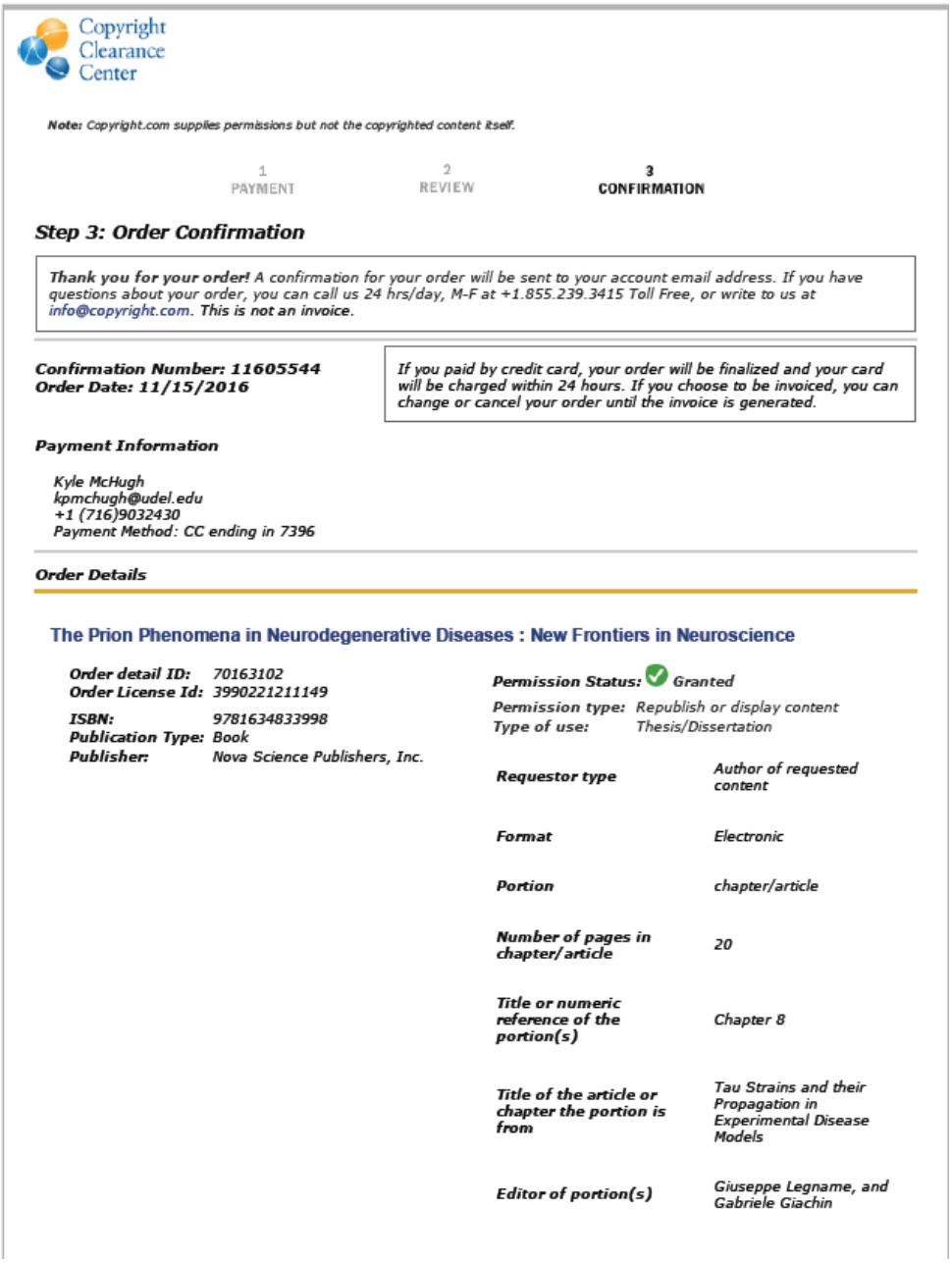


Figure B.2 Aggregation propensity was tested by incubation of scFv in PBS at 37C for 0, 24, or 48 hours followed by reducing SDS-PAGE and Western blot. Small quantities of SDS-insoluble dimer were detected at 24 and 48 hours, however the mutant 2.6.1 did not have significantly less aggregation that the WT ICSM18 scFv.

Appendix C

PERMISSION LETTERS



Copyright Clearance Center

Note: Copyright.com supplies permissions but not the copyrighted content itself.

1 PAYMENT 2 REVIEW 3 CONFIRMATION

Step 3: Order Confirmation

Thank you for your order! A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us 24 hrs/day, M-F at +1.855.239.3415 Toll Free, or write to us at info@copyright.com. This is not an invoice.

Confirmation Number: 11605544
Order Date: 11/15/2016

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.

Payment Information

Kyle McHugh
kpmchugh@udel.edu
+1 (716)9032430
Payment Method: CC ending in 7396

Order Details

The Prion Phenomena in Neurodegenerative Diseases : New Frontiers in Neuroscience

Order detail ID:	70163102
Order License Id:	3990221211149
ISBN:	9781634833998
Publication Type:	Book
Publisher:	Nova Science Publishers, Inc.

Permission Status:	Granted
Permission type:	Republish or display content
Type of use:	Thesis/Dissertation
Requestor type	Author of requested content
Format	Electronic
Portion	chapter/article
Number of pages in chapter/article	20
Title or numeric reference of the portion(s)	Chapter 8
Title of the article or chapter the portion is from	Tau Strains and their Propagation in Experimental Disease Models
Editor of portion(s)	Giuseppe Legname, and Gabriele Giachin

Author of portion(s)	Kyle P. McHugh, Olga A. Morozova and David W. Colby
Volume of serial or monograph	N/A
Page range of portion	157-176
Publication date of portion	2015 - 3rd quarter
Rights for	Main product
Duration of use	Life of current edition
Creation of copies for the disabled	no
With minor editing privileges	no
For distribution to	United States
In the following language(s)	Original language of publication
With incidental promotional use	no
Lifetime unit quantity of new product	Up to 499
Made available in the following markets	professional
The requesting person/ organization	Kyle McHugh
Order reference number	
Author/ Editor	Kyle McHugh
The standard identifier of New Work	N/A
Title of New Work	Engineering Antibody Therapeutics and Cell Culture Models for Neurodegenerative Diseases
Publisher of New Work	University of Delaware
Expected publication	Nov 2016

date

Estimated size (pages) 180

Note: This item will be invoiced or charged separately through CCC's **RightsLink** service. More info **$ 47.50**

This is not an invoice.

Total order items: 1 **Order Total: 47.50 USD**