

UCLA

UCLA Electronic Theses and Dissertations

Title

Chemical Biology Approaches to Modeling and Treating Brain Disease

Permalink

<https://escholarship.org/uc/item/58x7h7qm>

Author

Jahanbakhsh, Shahriyar

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Chemical Biology Approaches to Modeling and Treating Brain Disease

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in Biochemistry, Molecular and Structural Biology

by

Shahriyar Jahanbakhsh

2019

© Copyright by
Shahriyar Jahanbakhsh

2019

ABSTRACT OF THE DISSERTATION

Chemical Biology Approaches to Modeling and Treating Brain Disease

by

Shahriyar Jahanbakhsh

Doctor of Philosophy in Biochemistry, Molecular and Structural Biology

University of California, Los Angeles, 2019

Professor Carla Marie Koehler, Chair

Neurodegenerative diseases such as Parkinson's and Alzheimer's affect millions of people and are uniformly untreatable. A significant unmet need exists for faster and higher throughput approaches to modeling neurodegenerative diseases and identifying candidate small molecule therapies. Herein, I present a new chemical biology approach to identifying small molecule candidates for modeling and treating neurodegenerative diseases, using pontocerebellar hypoplasia type 1B and amyotrophic lateral sclerosis as test cases. The approach combines *in silico* rational design, *in vitro* biophysical testing, and *in vivo* assays in zebrafish to yield a powerful discovery platform for small molecule models and therapies for neurodegenerative disease.

The dissertation of Shahriyar Jahanbakhsh is approved.

Jau-Nian Chen

Catherine F. Clarke

Alexander M. van der Blik

Carla Marie Koehler, Committee Chair

University of California, Los Angeles

2019

DEDICATION

To Ariya,
the most courageous kid I know

TABLE OF CONTENTS

	Page
List of Figures	vii
List of Tables	ix
Abbreviations	x
Acknowledgements	xiii
Vita	xvi
 Chapter 1: Neurodegenerative Diseases and a New Paradigm for Modeling and Treating Them	
Introduction to Neurodegenerative Diseases	2
Overview of pontocerebellar hypoplasias	2
Overview of amyotrophic lateral sclerosis	6
Zebrafish as a Model System	8
Overview of zebrafish as a model organism	8
References	14
 Chapter 2: A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B)	
Main Manuscript	24
Supporting Information	36
 Chapter 3: A Small Molecule Candidate for Prevention or Treatment of Amyotrophic Lateral Sclerosis and Other TDP-43 Proteinopathies	
Introduction	63
Paraquat as a model for ALS and Parkinson's disease	64
rTRD01: a candidate small molecule for treatment or prophylaxis	

of TDP-43 proteinopathies	65
Materials and Methods	65
Zebrafish culture	65
Treatment conditions	66
Imaging	66
Results	67
rTRD01 protects against paraquat-induced neurodegeneration in zebrafish.	68
rTRD01 is not a general antioxidant.	68
Discussion	74
References	75
Chapter 4: Toward a zebrafish model of pontocerebellar hypoplasia type 1D	
Introduction	80
MitoTimer: a fluorophore for assaying mitochondrial turnover	80
Materials and Methods	81
Zebrafish culture	81
Zebrafish mutagenesis	81
Results	81
Discussion	83
References	85

LIST OF FIGURES

Chapter 1: Neurodegenerative Diseases and a New Paradigm for Modeling and Treating Them	Page
Figure 1-1. <i>Tol2</i> transgenesis in zebrafish.	10
Figure 1-2. Zebrafish CRISPR/Cas9 mutagenesis workflow.	12
Chapter 2: A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B)	
Figure 1. RNA exosome structure highlighting PCH1B-associated mutations in human EXOSC3.	26
Figure 2. Effect of PCH1B mutations on EXOSC3 binding to adjacent subunits through a peptide array.	27
Figure 3. Profiling RNA binding landscape of EXOSC3.	27
Figure 4. PCH1B mutations on EXOSC3 altering the RNA binding landscape.	28
Figure 5. Docking of small molecules onto EXOSC3.	29
Figure 6. STD-NMR analysis of binding of small molecules to exosome subunits.	30
Figure 7. RNA/EXOSC3 disruption by small molecules.	31
Figure 8. ERD compounds induce a PCH1B-like phenotype in zebrafish.	31
Figure 9. Biophysical characterization of ERD03.	32
Figure 10. Transcriptomics in zebrafish after ERD03 incubation or EXOSC3-knockdown.	32
Figure 11. ERD03 effect on cerebellum development.	33
Supplementary Figure S1. Interaction of EXOSC3 with EXOSC1, EXOSC5 and EXOSC9.	44

Supplementary Figure S2. Stability of disease causing EXOSC3 mutants.	45
Supplementary Figure S3. EXOSC3 binding to adjacent subunits through peptide array.	46
Supplementary Figure S4. Z-score distribution of RNA 7-mers from RNAcompete on EXOSC3 proteins.	48
Supplementary Figure S5. Top hits from <i>in silico</i> docking poses for each domain.	49
Supplementary Figure S6. Comparison of human structures of EXOSC3 with homology model of zebrafish EXOSC3 .	50
Supplementary Figure S7. ERD03 disrupts EXOSC3-RNA interaction in a dose-dependent fashion.	51
Supplementary Figure S8. Effect of ERD03 and ERD18 on EXOSC3 binding to adjacent subunits assessed by peptide array.	52
 Chapter 3: A Small Molecule Candidate for Prevention or Treatment of Amyotrophic Lateral Sclerosis and Other TDP-43 Proteinopathies	
Figure 3-1. Domain architecture of TDP-43.	64
Figure 3-2. Toxicity titration and survival curve for rTRD01 exposure in zebrafish.	69
Figure 3-3. rTRD01 rescues paraquat-induced neurodegenerative phenotypes in zebrafish.	70
Figure 3-4. Neuroprotection by rTRD01 is not the result of chemical inactivation of paraquat by rTRD01.	72
Figure 3-5. rTRD01 does not rescue rotenone-induced neuronal death in zebrafish.	73
 Chapter 4: Toward a zebrafish model of pontocerebellar hypoplasia type 1D	
Figure 4-1. Heterozygous knockout of SLC25A46 in the zebrafish.	82

LIST OF TABLES

Chapter 2: A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B)	Page
Supplementary Table 2. Structures of ERD compounds.	54
Supplementary Table 4. Primers used to clone the cap proteins.	57
Supplementary Table 5. Primers used to generate EXOSC3 mutants.	58
Supplementary Table 6. qPCR primers used to assess mRNA levels in zebrafish embryos.	59

ABBREVIATIONS

Abbreviations

ALS	amyotrophic lateral sclerosis
ANOVA	analysis of variance
bp	base pair
CNS	central nervous system
CRISPR	clustered regularly interspaced short palindromic repeats
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dpf	days post-fertilization
ER	endoplasmic reticulum
FDA	Food and Drug Administration
GFP	green fluorescent protein
GTP	guanosine-5'-triphosphate
HIV	human immunodeficiency virus
hpf	hours post-fertilization
HSD	honestly significant difference
KH	K homology
mRNA	messenger RNA
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NTD	N-terminal domain
OMM	outer mitochondrial membrane
PCR	polymerase chain reaction

PE	phosphatidylethanolamine
PH	Pleckstrin homology
PQ	paraquat (<i>N,N'</i> -dimethyl-4,4'-bipyridinium dichloride)
PTU	phenylthiourea (phenylthiocarbamide)
RNA	ribonucleic acid
ROS	reactive oxygen species
RRM	RNA recognition motif
rTRD	<u>RRM</u> <u>TDP-43</u> <u>RNA</u> <u>disruptor</u>
SG	stress granule
sgRNA	single guide RNA
SLC	solute carrier
TDP-43	transactive response DNA binding protein of 43 kilodaltons
tRNA	transfer RNA

Metric Prefixes

k	kilo (1×10^3)
μ	micro (1×10^{-6})
m	mili (1×10^{-3})
n	nano (1×10^{-9})
p	pico (1×10^{-12})

Weights and Measures

Da	dalton
g	gram

L	liter
m	meter
M	molar
ppm	parts per million
v/v	volume in volume
w/v	weight in volume

ACKNOWLEDGEMENTS

Thank you Professor Carla Koehler for your generous mentorship, patient guidance, and steady financial support as my advisor and Committee Chair. Thank you Professors Jau-Nian Chen, Alex van der Bliet, and Cathy Clarke for graciously providing your time and expertise as members of my Committee.

Thank you Adam Langenbacher, Yali Zhao, and Alejandro Martorell for consultations and hands-on training. Alejandro, special thanks to you for all the time you spent discussing with me the culture of Catalunya and Spain—gracias, gràcies, y (aunque no habláis euskara en Catalunya) eskerrik asko!

Thank you to the brilliant and talented undergraduate researchers with whom I have had the privilege to work: Gary Shmorgon (Hellos Gary!), Kayla White, Jennifer Ly (Hellos Jenny!), Gloria Yeo, Parsia Vazirnia, Matthew Hennefarth, Lucas Vincenzi, and William Im. Special thanks to Gloria for serving as “the cartoonist” for our Hult Prize 2018 bid, and for reminding me that fast and crazy does not win the race (even though I still disagree). Matthew, I’m proud that your first published work is with me, and I’m happy you found Peter again after 20 years. Lucas, thank you for being the first to bring the glorious and glamorous culture of Italy to our floor. Molto grazie Luca! Ti auguro il meglio.

Thank you Adam Makhlef, Essam Assali, Nate Miller, Eva Koziolk, Chiara Montemurro, Jordan Whatley, Melania Abrahamian, Eriko Shimada, Michael Conti, Charles Wang, and Ivan Ramirez for your true friendship during my time in graduate school. Our discussions of the *Art of the Steal*, the Persian Gulf, Ryan Dellamaggiora, Chancellor Willy Brandt, DOCG (Denominazione di Origine Controllata e Garantita) wines, fiancées, languages,

self-cleaning toilets, Italian counts, K-pop, the Donald, *Art of the Deal*, and international “connections” to the Basque Country will always remain in my mind. Thank you Chiara, Alberto Tombolato and Guglielmo Graser for enriching the Italian theme of our floor.

Thank you to my friends Peter Shoemaker, Ben Wagstaff, Marcus Talke, Aaron Rubanowitz, Marc Roubé, and Albert Teplitsky for your excellent company during my time in graduate school. Exploring Los Angeles, San Francisco, San Diego, and Las Vegas with you guys was such a blast.

Thank you to the brilliant physicians and therapists who cared for me during my nearly seven years at UCLA: Mark Ackerman, Manuel Dizon, David McAllister, Sharon Hame, Casimir Dowd, Karen Yeung, Sean Wahlquist, David Sarraf, and Evan Brahce.

Thank you to my dear parents and brother for their love during my time in graduate school and to my mom for making me delicious food during my visits home.

Finally, thank you Sean Atamdede for your diligence in ensuring I had all the needed tools for my research. You steadfastly supported my work and made things easy for me, even when I did not return the favor.

Chapter 2 is a reprint of the published manuscript:

“A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B).” Liberty François-Moutal, Shahriyar Jahanbakhsh, Andrew D. L. Nelson, Debashish Ray, David D. Scott, Matthew R. Hennefarth, Aubin Moutal, Samantha Perez-Miller, Andrew J. Ambrose, Ahmed Al-Shamari, Philippe Coursodon, Bessie Meechoovet, Rebecca Reiman, Eric Lyons, Mark Beilstein, Eli Chapman, Quaid D. Morris, Kendall Van Keuren-Jensen, Timothy R. Hughes,

Rajesh Khanna, Carla Koehler, Joanna Jen, Vijay Gokhale, and May Khanna.
ACS Chemical Biology 2018 13(10), 3000–3010. DOI:
10.1021/acscchembio.8b00745.

Shahriyar Jahanbakhsh and Matthew R. Hennefarth performed the *in vivo* cerebellar size determination experiment and generated the accompanying figures and text. The remainder of the manuscript is the work of Liberty François-Moutal and the other authors. Copyright © 2018, American Chemical Society. Reprinted with permission.

Chapter 3 contains Shahriyar Jahanbakhsh's and Matthew R. Hennefarth's contributions to an unpublished manuscript. The other co-authors are Liberty François-Moutal, David D. Scott, Razaz Felemban, Victor Miranda, Melissa Sayegh, Samantha Perez-Miller, Vijay Gokhale, Daniela C. Zarnescu, and May Khanna, all at the University of Arizona, Tucson; these co-authors discovered the small molecule rTRD01 and performed the biophysical characterization of rTRD01.

Chapter 4 describes the joint work of Gary Shmorgon, Kayla White, and Shahriyar Jahanbakhsh. Gary Shmorgon and Kayla White generated the mutation described using CRISPR/Cas9 mutagenesis. Shahriyar Jahanbakhsh characterized the mutation and isolated the mutant line.

VITA

Education

B.S., Biochemistry, *summa cum laude*, June 2016, UCLA

M.S., Biochemistry, Molecular and Structural Biology, June 2016, UCLA

C.Phil., Biochemistry, Molecular and Structural Biology, June 2018, UCLA

Published Works

Jahanbakhsh, Shahriyar. “Novel regulators of social motility in the African trypanosome, *Trypanosoma brucei*.” [Print]. Ann Arbor, Michigan: ProQuest; 2016.

“A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B).” Liberty François-Moutal, **Shahriyar Jahanbakhsh**, Andrew D. L. Nelson, Debashish Ray, David D. Scott, Matthew R. Hennefarth, Aubin Moutal, Samantha Perez-Miller, Andrew J. Ambrose, Ahmed Al-Shamari, Philippe Coursodon, Bessie Meechoovet, Rebecca Reiman, Eric Lyons, Mark Beilstein, Eli Chapman, Quaid D. Morris, Kendall Van Keuren-Jensen, Timothy R. Hughes, Rajesh Khanna, Carla Koehler, Joanna Jen, Vijay Gokhale, and May Khanna. ACS Chemical Biology 2018 *13*(10), 3000–3010. DOI: 10.1021/acscchembio.8b00745.

Selected Awards and Honors

2013 Daniel Kivelson Research Fellowship, UCLA

2013 Hosmer Stone Prize for Excellence in Chemistry, UCLA

2014 Daniel Kivelson Research Fellowship, UCLA

2015 Elmer Carvey Scholarship, Aerospace Corporation

2015 Phi Beta Kappa, The Phi Beta Kappa Society

2015 Vice Provost’s Research Fellowship, UCLA

2016 Winner, UCLA/Perkins Coie Venture Competition, UCLA Business of Science Center

2017 3rd Prize, Knapp Venture Competition, Anderson School of Management

2018 Semifinalist, Hult Prize

2018 Finalist, Sandler Prize for New Entrepreneurs, Lowell Milken Institute

2019 Semifinalist, Hult Prize

Languages

Native (S-5): English, Persian

Fluent (S-4): Spanish, Portuguese

Conversational (S-2): Basque

Chapter 1:
Neurodegenerative Diseases and a New Paradigm for
Modeling and Treating Them

Introduction to Neurodegenerative Diseases

Neurodegenerative diseases afflict seven million people in the United States alone and 40 million people worldwide. At present, they are uniformly incurable and untreatable. The burden of neurodegenerative disease continues to increase as the world population ages, rendering the need for effective therapies ever more urgent.

Neurodegenerative disease is an umbrella term that encompasses well-known pathologies such as Parkinson's and Alzheimer's, as well as less common illnesses such as amyotrophic lateral sclerosis (ALS), Huntington's chorea, frontotemporal dementia, the spinocerebellar ataxias, and pontocerebellar hypoplasias, among many others. While displaying diverse pathophysiologies, neurodegenerative diseases exhibit two common themes: [1] their incidence and prevalence increase with age; and [2] most cases are sporadic, although familial forms, including monogenic forms with early age of onset, do exist.

The ensuing chapters describe the discovery of a small molecule model for a pontocerebellar hypoplasia and a small molecule candidate for prophylaxis or treatment of ALS, diseases to which we now turn our attention.

Overview of pontocerebellar hypoplasias

Pontocerebellar hypoplasias (PCH) are a group of very rare, heterogeneous conditions characterized by prenatal development of an abnormally small cerebellum and ventral pons. The exact prevalence of PCH is unknown, but is estimated to be less than 1/1,000,000. The main clinical feature is profound psychomotor retardation. In many cases, PCH is fatal early in life, although patients with less severe forms can survive into adolescence or young adulthood¹.

As of 2018, clinicians and researchers have defined 11 types of PCH on the basis of clinical and genetic criteria². These 11 types are further subcategorized according to the specific genetic mutation present. The various genes defective in PCH patients function in a broad range of processes including neuronal migration, mRNA degradation, mitochondrial fusion and fission, tRNA splicing, arginyl tRNA synthesis, selenocysteine synthesis, RNA processing, regulation of GTP synthesis, and intracellular vesicle transport, among others. The subtypes of PCH that disrupt mRNA degradation (PCH1B) and mitochondrial dynamics (PCH1D) are the focus of this dissertation.

Pontocerebellar hypoplasia type 1B

PCH1B is the second most common subtype of PCH after PCH2, and one of the best studied. As with all subtypes of PCH, PCH1B is characterized by profound cerebellar and pontine atrophy. PCH1B is additionally characterized by dysmorphic Purkinje cells, loss of granule cells, and degeneration of motor neurons in the anterior spinal horn reminiscent of spinal muscular atrophy. The lifespan of PCH1B patients ranges from neonatal death to survival into puberty³.

Pathogenesis of PCH1B⁴

PCH1B is caused by mutations in the EXOSC3 subunit of the RNA exosome, a complex of cytosolic proteins that processes and degrades RNA in eukaryotes^{5,6}. PCH1C is the only other subtype caused by mutations in an exosome subunit⁷. The RNA exosome is involved in the 3'—5' processing of small nucleolar RNAs, small nuclear RNAs, and ribosomal RNAs⁸. The overall structure of the RNA exosome core (EXOSC1–9) is conserved from prokaryotes to higher eukaryotes and contains a ring-like structure composed of six RNase Pleckstrin homology

domains that form a central channel able to accommodate single-stranded RNA, and three additional proteins that form a “cap-like” structure^{9,10,11}.

The “cap proteins”—the entry point of RNAs threading through the exosome—are EXOSC1, -2, and -3. EXOSC2 and -3 contain two putative RNA-binding domains, the KH (heterogeneous nuclear ribonucleoprotein K homology) and S1 (ribosomal protein S1) domains, while EXOSC1 only contains an S1 domain; all three proteins have a similar N-terminal domain (NTD). The S1 domains are positioned to bind RNA as it threads through the cap toward the core proteins and central channel of the RNA exosome¹², although it is not known if the KH domains also contribute to RNA binding.

Whole genome sequencing of patients with PCH1B has identified several mutations clustered in EXOSC3^{3,13,14,15,16}. Survival of the patients is correlated with the type of mutation and its association with a nonsense allele. Thus far, eight different point mutations have been reported, as either homozygous or heterozygous morphism associated with PCH1B (G31A, V80F, Y109N, D132A, G135E, A139P, and W238R)^{3,13,14,15,16}. The majority of these mutations cluster in the S1 domain, one mutation in the KH domain, and two in the NTD of EXOSC3.

Several studies have offered insights into the contribution of mutations in EXOSC3 to the pathogenesis of PCH1B, although a complete understanding remains elusive. Knockdown of EXOSC3 in zebrafish yields a PCH1B-like phenotype associated with common alterations in the expression of genes implicated in development, including *ataxin1b* and the homeobox gene *HOXC*¹⁷. Another study demonstrated the lack of stability of a yeast analog of EXOSC3-W238R and its failure to associate efficiently with the exosome in the presence of the wild-type yeast EXOSC3^{17,18}. A third study reported sequestration of EXOSC3-D132A in the cytosol of patient

fibroblasts and an accumulation of mRNA in the muscles of these patients; both events were associated with mitochondrial dysfunction¹⁹. The work I present herein, among other things, further advances our understanding of the pathogenesis of PCH1B.

Pontocerebellar hypoplasia type 1D

PCH1D is the most severe subtype within the PCH1 family. Clinically, PCH1D manifests as severe atrophy of the cerebellum and brainstem in addition to hypotonia, severe global developmental delay, and diffuse cerebral atrophy. Affected individuals succumb to their illness at latest a few months after birth².

Pathogenesis of PCH1D

PCH1D is caused by mutations in the solute carrier (SLC) protein SLC25A46. SLC25A46 is an outer mitochondrial membrane (OMM) protein with important functions in mitochondrial dynamics and cristae maintenance. SLC25A46 is an orphan member of the SLC25 family, meaning its endogenous ligand is unknown. SLC25A46 is distantly related to the yeast OMM protein Ugo1²⁰. SLC25A46 interacts, most likely transiently, with the mitofusins (required for mitochondrial fusion) and the mitochondrial contact site and cristae organizing system complex (required for cristae maintenance)²¹. In addition to causing PCH1D, mutations in SLC25A46 are known to cause a host of other neurodegenerative disorders, including optic atrophy²², the axonal peripheral neuropathy Charcot-Marie-Tooth type 2²², an optic atrophy spectrum disorder²³, Leigh syndrome²⁴, and progressive myoclonic ataxia with optic atrophy and neuropathy²⁵.

A loss of function mutation in SLC25A46 (L341P) that promotes SLC25A46's rapid degradation by the ubiquitin-proteasome system is among the mutations that cause PCH1D²¹.

The rapid degradation of SLC25A46-L341P is specific and independent of mitophagy and apoptosis, and results in oligomerization of mitofusin 1/2 and mitochondrial hyperfusion *in vitro*. The effect of knockout of SLC25A46 on mitochondrial morphology *in vivo* remains disputed, however, with some investigators reporting no mitochondrial hyperfusion²⁶ and others reporting large round mitochondria with swollen cristae in knockout mice²⁷. Clearly, additional studies are needed to clarify the role of SLC25A46 on mitochondrial morphology and the pathogenesis of PCH1D. In the closing chapter of this dissertation, I present progress toward establishing the first stable zebrafish genetic model of PCH1D.

Overview of amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease in the United States and motor neurone disease in Britain, is a progressive, degenerative condition of the motor neurons, the nerves that link the motor cortex to the skeletal muscles. From the time of onset of symptoms, ALS typically proves fatal in 2–4 years²⁸. ALS afflicts approximately 20,000 people in the United States with an annual incidence of two per 100,000 people²⁸. Most cases of ALS are sporadic, although a familial form does exist and accounts for 5–10 percent of ALS cases. As with PCH, a broad array of genetic defects can produce the degenerative phenotype observed in familial ALS²⁹.

ALS is characterized by degeneration of corticospinal motor neurons, the descending axons of which in the lateral spinal cord seem scarred (lateral sclerosis), and the death of spinal motor neurons, with secondary denervation and muscle wasting (amyotrophy). Corticospinal neurons make direct or indirect connections with spinal motor neurons, which innervate skeletal muscles and trigger muscle contraction³⁰. Clinically, patients suffer from progressive paralysis.

Over a third of patients also experience cognitive or behavioral impairment³¹. Although 145 years have passed since Jean-Martin Charcot first described ALS in 1874³², the disease remains incurable, and the sole FDA-approved therapeutic, riluzole, prolongs survival by a mere 2–3 months³³.

Pathogenesis of ALS

The pathogenesis of ALS is complex and multifactorial. Nonetheless, accumulation of the DNA- and RNA-binding protein transactive response DNA binding protein 43 (TDP-43) is the most significant pathological finding in approximately 95 percent of all ALS cases³⁴. Additionally, about 4.5 percent of ALS patients bear dominant mutations in TDP-43 itself³⁵.

TDP-43 is encoded by the *TARDBP* gene and was first identified as a binder of pyrimidine-rich DNA motifs in the long terminal repeat of HIV-1 (human immunodeficiency virus type 1)³⁶. Subsequent studies demonstrated TDP-43 as capable of binding mRNA and functioning in the regulation of mRNA splicing, translation, transport, and degradation^{37,38,39,40,41,42}. The discovery that TDP-43 is a major component of the ubiquitinated inclusions bodies found in the brains of patients suffering from ALS and frontotemporal lobar degeneration was a game-changer for the field^{43,44}. TDP-43-positive inclusions were subsequently found in patients with other neurodegenerative disease such as Parkinson's, Alzheimer's, frontotemporal dementia, primary progressive nonfluent aphasia, and semantic dementia, giving rise to "TDP-43 proteinopathy" as a new pathomechanism⁴⁵.

Under physiological conditions, TDP-43 is a predominantly nuclear protein. In patients with ALS and other neurodegenerative diseases (i.e. under pathological conditions), TDP-43 redistributes to the cytosol⁴⁶, where it is hyperphosphorylated at its C-terminus⁴⁷, ubiquitinated

by parkin⁴⁸ at its four C-terminal lysines (as well as N-terminal lysines)⁴⁹, and proteolytically cleaved by caspases and calpain proteases into 35 kDa and 25 kDa C-terminal fragments that form the major components of inclusion bodies⁵⁰. Phosphorylation appears to precede ubiquitination⁴⁷, and cleaved fragments are preferentially ubiquitinated and phosphorylated compared to full-length TDP-43^{47,51}. A plethora of dominantly-inherited point mutations in TDP-43 have also been characterized and are the cause of 5–10 percent of familial ALS cases⁵². These mutations may promote TDP-43 hyperphosphorylation and aggregation but are not necessary to induce TDP-43 proteinopathy⁵³.

While TDP-43 redistribution and aggregation is major pathological hallmark of ALS and other neurodegenerations, the pathogenesis of ALS, like that of other neurodegenerative diseases, is multifactorial. Other pathways such as oxidative stress, glutamate excitotoxicity, inflammation, accumulation of neurofilaments, and mitochondrial dysfunction all contribute to the pathogenesis of ALS⁵⁴.

Zebrafish as a Model System

Zebrafish (*Danio rerio*, formerly *Brachydanio rerio*) are freshwater fish native to South Asia. The use of zebrafish in laboratory research was pioneered in the 1970s by George Streisinger, whose laboratory at the University of Oregon was the first to clone a vertebrate animal⁵⁵. The ensuing decades saw widespread adoption of the zebrafish as a model organism.

Overview of zebrafish as a model organism

Zebrafish are excellent model organisms for the study of neuronal and mitochondrial biology given their high conservation of genetic information (including mitochondrial DNA) and physiological processes, inexpensive maintenance, and optical transparency facilitating direct

observation *in vivo*⁵⁶. The zebrafish genome shows approximately 70 percent homology with the human genome, and 84 percent of genes known to be associated with human disease have counterparts in the zebrafish⁵⁷. Moreover, embryonic development proceeds rapidly, with most major organs being present by two days post-fertilization (dpf). Zebrafish are also amenable to genetic manipulation *via* a variety of techniques, including *tol2*-mediated transgenesis⁵⁸ (**Figure 1-1**) and CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9-based mutagenesis (**Figure 1-2**)⁵⁹. Importantly for the research I present in this work, mounting evidence supports generally conserved genetics and physiology of major brain processes in mammals and zebrafish, and the duplicated central nervous system (CNS) genes in zebrafish mostly encode proteins with similar functions⁶⁰.

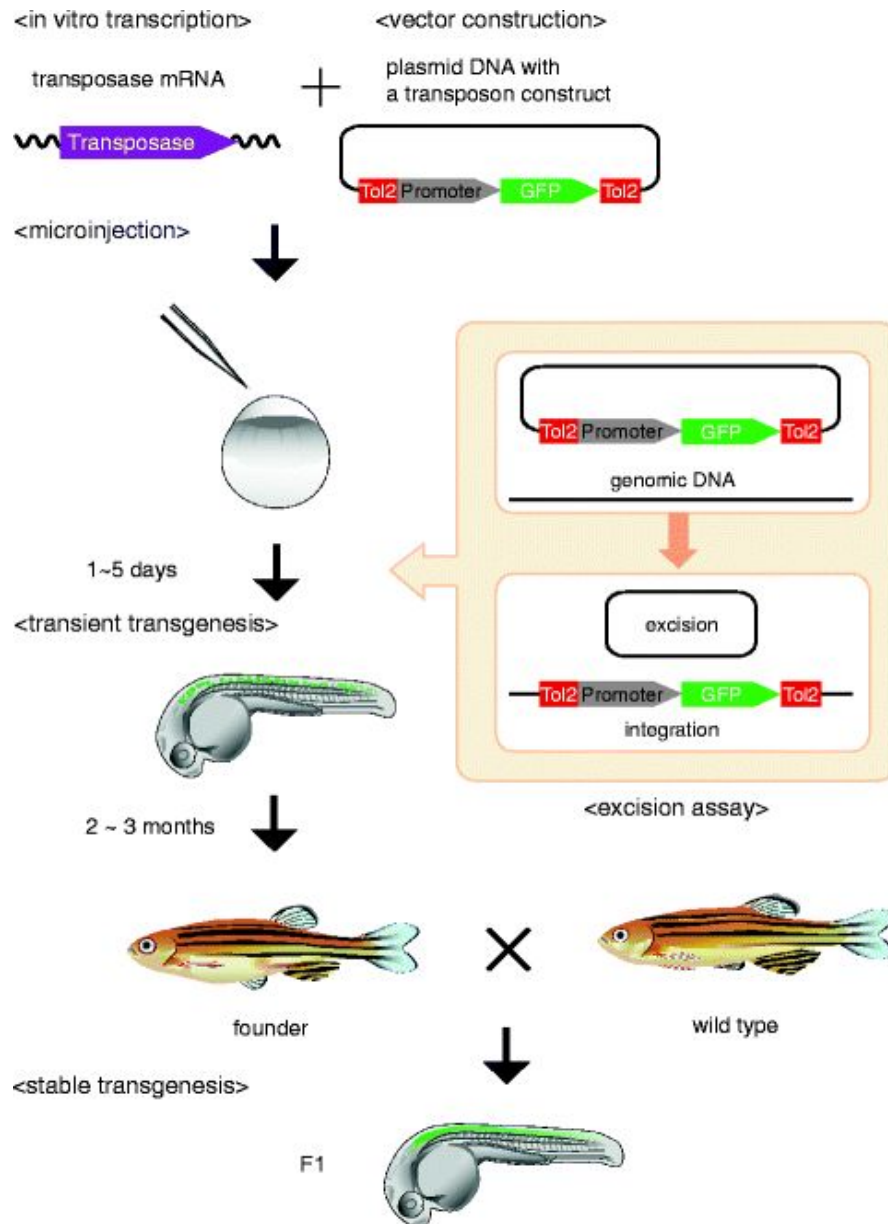


Figure 1-1⁵⁸. Tol2 transgenesis in zebrafish. Transposase mRNA and a plasmid DNA containing the gene of interest (here, GFP) flanked by minimal *tol2* element sequences are co-microinjected into nascently-fertilized embryos. The transposase is synthesized *in vivo* from mRNA and catalyzes excision of the Tol2 construct from the plasmid and its random transposition into the genome. Mosaic GFP expression is observed in the injected founder (F₀)

embryos. The injected embryos are raised to sexual maturity and crossed with wild-type fish. The integrated Tol2 construct is transmitted to the first filial (F_1) generation, which now show stable (non-mosaic) expression of the gene of interest. Figure reproduced with permission from Springer Nature.

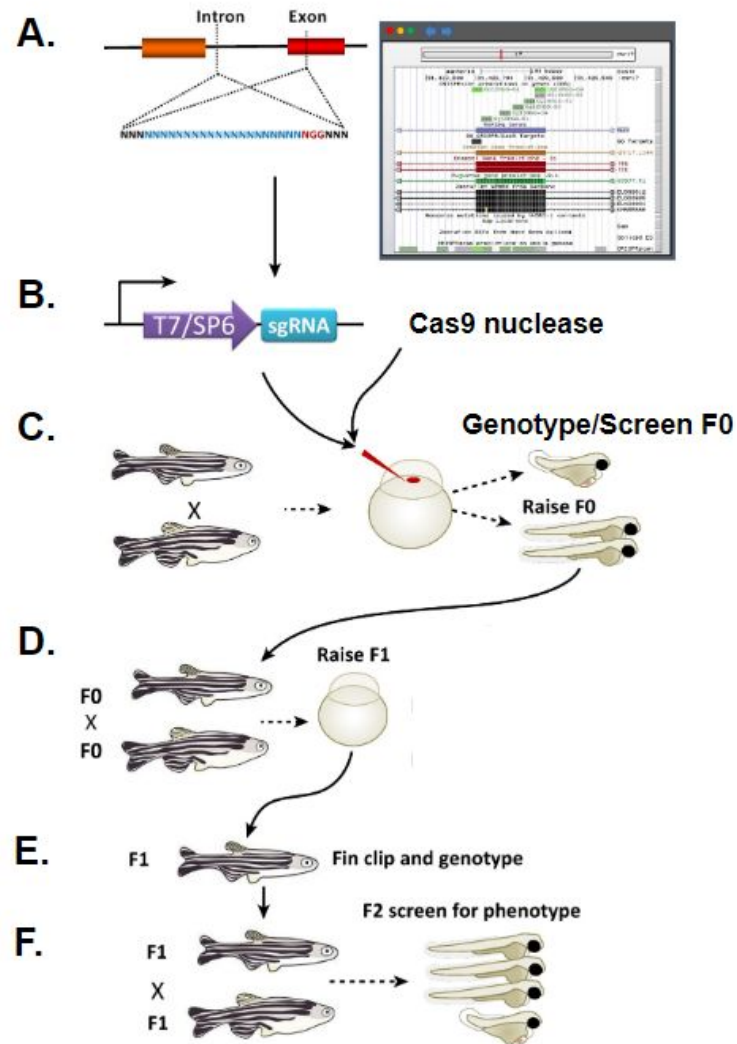


Figure 1-2⁶¹. Zebrafish CRISPR/Cas9 mutagenesis workflow. (A) The target DNA sequence is selected, (B) the single guide RNA (sgRNA) is synthesized and (C) co-injected together with Cas9 nuclease into the nascently-fertilized zebrafish embryo. The efficiency of mutagenesis is then evaluated *via* PCR-RFLP (restriction fragment length polymorphism) genotyping or phenotypic screen, if possible. (D) F₀ animals are raised to sexual maturity crossed to other F₀ animals. (E) The genotype and phenotype of the F₁ offspring are assessed to evaluate germline transmission. (F) Mutagenized F₁ animals are crossed to other mutagenized siblings to yield F₂ animals. Figure adapted with permission from Elsevier.

Zebrafish models of neurodegenerative disease

As a vertebrate animal, the zebrafish is not only an excellent model organism for human diseases in general, but is also well-suited as a model for neurodegenerative diseases in particular. The zebrafish brain shares extensive structural and functional homology with the human brain⁶²; genes that control the production of neurons and axonal guidance are shared among zebrafish and mammals⁶²; the zebrafish noradrenergic, serotonergic, and histaminergic systems are highly similar to those of mammals^{63,64,65,66,67,68,69,70,71,72} and, except for the lack of dopaminergic neurons in the zebrafish mesencephalon, the zebrafish dopaminergic system also highly resembles that of mammals⁷³.

Zebrafish have successfully been used to model a number of neurodegenerative diseases in recent years, including pontocerebellar hypoplasia, CNL2 disease (neuronal ceroid lipofuscinosis, a lysosomal storage disorder characterized by progressive intellectual and motor deterioration, seizures, and early death⁷⁴), Parkinson's, dementia with Lewy bodies, and ALS⁷⁵. Most of the models to date have been generated by injection of antisense morpholinos (synthetic oligomers consisting of DNA bases covalently bound to a phosphorodiamidate-linked methylenemorpholine backbone that knock down gene function) or mRNAs⁷⁵ into the zygote.

In the chapters that follow, I use small molecule- and CRISPR-based approaches, driven by chemical biology, to generate new zebrafish models for neurodegenerative disease.

References

1. Sánchez-Albisua, I., Frölich, S., Barth, P. G., Steinlin, M. & Krägeloh-Mann, I. Natural course of pontocerebellar hypoplasia type 2A. *Orphanet J. Rare Dis.* **9**, 70 (2014).
2. van Dijk, T., Baas, F., Barth, P. G. & Poll-The, B. T. What's new in pontocerebellar hypoplasia? An update on genes and subtypes. *Orphanet J. Rare Dis.* **13**, 92 (2018).
3. Wan, J. *et al.* Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.* **44**, 704–708 (2012).
4. François-Moutal, L. *et al.* A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B). *ACS Chemical Biology* **13**, 3000–3010 (2018).
5. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The Exosome: A Conserved Eukaryotic RNA Processing Complex Containing Multiple 3'→5' Exoribonucleases. *Cell* **91**, 457–466 (1997).
6. Anderson, J. S. J. The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *The EMBO Journal* **17**, 1497–1506 (1998).
7. Boczonadi, V. *et al.* EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia. *Nat. Commun.* **5**, 4287 (2014).
8. Allmang, C. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *The EMBO Journal* **18**, 5399–5410 (1999).

9. Makino, D. L., Baumgärtner, M. & Conti, E. Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature* **495**, 70–75 (2013).
10. Kowalinski, E. *et al.* Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. *Molecular Cell* **63**, 125–134 (2016).
11. Wasmuth, E. V., Januszyk, K. & Lima, C. D. Structure of an Rrp6–RNA exosome complex bound to poly(A) RNA. *Nature* **511**, 435–439 (2014).
12. Januszyk, K. & Lima, C. D. Structural components and architectures of RNA exosomes. *Adv. Exp. Med. Biol.* **702**, 9–28 (2011).
13. Halevy, A. *et al.* Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. *Journal of Neurology* **261**, 2165–2169 (2014).
14. Schwabova, J. *et al.* Homozygous EXOSC3 mutation c.92G→C, p.G31A is a founder mutation causing severe pontocerebellar hypoplasia type 1 among the Czech Roma. *J. Neurogenet.* **27**, 163–169 (2013).
15. Eggens, V. R. *et al.* EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations. *Orphanet J. Rare Dis.* **9**, 23 (2014).
16. Zanni, G. *et al.* Exome sequencing in a family with intellectual disability, early onset spasticity, and cerebellar atrophy detects a novel mutation in EXOSC3. *Neurogenetics* **14**, 247–250 (2013).
17. Giunta, M. *et al.* Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. *Hum. Mol. Genet.* **25**, 2985–2996 (2016).
18. Gillespie, A., Gabunilas, J., Jen, J. C. & Chanfreau, G. F. Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the

- exosome in. *RNA* **23**, 466–472 (2017).
19. Fasken, M. B. *et al.* Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast. *Genetics* **205**, 221–237 (2017).
 20. Wan, J. *et al.* Loss of function of SLC25A46 causes lethal congenital pontocerebellar hypoplasia. *Brain* **139**, 2877–2890 (2016).
 21. Steffen, J. *et al.* Rapid degradation of mutant SLC25A46 by the ubiquitin-proteasome system results in MFN1/2-mediated hyperfusion of mitochondria. *Mol. Biol. Cell* **28**, 600–612 (2017).
 22. Abrams, A. J. *et al.* Mutations in SLC25A46, encoding a UGO1-like protein, cause an optic atrophy spectrum disorder. *Nat. Genet.* **47**, 926–932 (2015).
 23. Nguyen, M. *et al.* Novel pathogenic SLC25A46 splice-site mutation causes an optic atrophy spectrum disorder. *Clin. Genet.* **91**, 121–125 (2017).
 24. Janer, A. *et al.* SLC25A46 is required for mitochondrial lipid homeostasis and cristae maintenance and is responsible for Leigh syndrome. *EMBO Mol. Med.* **8**, 1019–1038 (2016).
 25. Charlesworth, G. *et al.* SLC25A46 mutations underlie progressive myoclonic ataxia with optic atrophy and neuropathy. *Mov. Disord.* **31**, 1249–1251 (2016).
 26. Terzenidou, M. E. *et al.* Novel insights into SLC25A46-related pathologies in a genetic mouse model. *PLoS Genet.* **13**, e1006656 (2017).
 27. Li, Z. *et al.* Loss of SLC25A46 causes neurodegeneration by affecting mitochondrial dynamics and energy production in mice. *Hum. Mol. Genet.* **26**, 3776–3791 (2017).

28. Arthur, K. C. *et al.* Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat. Commun.* **7**, 12408 (2016).
29. Boylan, K. Familial Amyotrophic Lateral Sclerosis. *Neurol. Clin.* **33**, 807–830 (2015).
30. Taylor, J. P., Brown, R. H., Jr & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).
31. Crockford, C. *et al.* ALS-specific cognitive and behavior changes associated with advancing disease stage in ALS. *Neurology* **91**, e1370–e1380 (2018).
32. Rowland, L. P. & Shneider, N. A. Amyotrophic lateral sclerosis. *N. Engl. J. Med.* **344**, 1688–1700 (2001).
33. Miller, R. G., Mitchell, J. D. & Moore, D. H. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst. Rev.* CD001447 (2012).
34. Majumder, V., Gregory, J. M., Barria, M. A., Green, A. & Pal, S. TDP-43 as a potential biomarker for amyotrophic lateral sclerosis: a systematic review and meta-analysis. *BMC Neurol.* **18**, 90 (2018).
35. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: the FUS about TDP-43. *Cell* **136**, 1001–1004 (2009).
36. Ou, S. H., Wu, F., Harrich, D., García-Martínez, L. F. & Gaynor, R. B. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J. Virol.* **69**, 3584–3596 (1995).
37. Buratti, E. & Baralle, F. E. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator of CFTR Exon 9. *Journal of Biological Chemistry* **276**, 36337–36343 (2001).

38. Buratti, E. Nuclear factor TDP-43 and SR proteins promote *in vitro* and *in vivo* CFTR exon 9 skipping. *The EMBO Journal* **20**, 1774–1784 (2001).
39. Hefferon, T. W., Groman, J. D., Yurk, C. E. & Cutting, G. R. A variable dinucleotide repeat in the CFTR gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing. *Proceedings of the National Academy of Sciences* **101**, 3504–3509 (2004).
40. Buratti, E. *et al.* TDP-43 Binds Heterogeneous Nuclear Ribonucleoprotein A/B through Its C-terminal Tail. *Journal of Biological Chemistry* **280**, 37572–37584 (2005).
41. Ayala, Y. M. *et al.* Human, Drosophila, and C. elegans TDP43: Nucleic Acid Binding Properties and Splicing Regulatory Function. *Journal of Molecular Biology* **348**, 575–588 (2005).
42. Mercado, P. A., Ayala, Y. M., Romano, M., Buratti, E. & Baralle, F. E. Depletion of TDP 43 overrides the need for exonic and intronic splicing enhancers in the human apoA-II gene. *Nucleic Acids Res.* **33**, 6000–6010 (2005).
43. Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).
44. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
45. Cook, C., Zhang, Y.-J., Xu, Y.-F., Dickson, D. W. & Petrucelli, L. TDP-43 in neurodegenerative disorders. *Expert Opin. Biol. Ther.* **8**, 969–978 (2008).
46. Barmada, S. J. *et al.* Cytoplasmic mislocalization of TDP-43 is toxic to neurons and

- enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J. Neurosci.* **30**, 639–649 (2010).
47. Hasegawa, M. *et al.* Phosphorylated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Ann. Neurol.* **64**, 60–70 (2008).
 48. Hebron, M. L. *et al.* Parkin ubiquitinates Tar-DNA binding protein-43 (TDP-43) and promotes its cytosolic accumulation via interaction with histone deacetylase 6 (HDAC6). *J. Biol. Chem.* **288**, 4103–4115 (2013).
 49. Hans, F., Eckert, M., von Zweydford, F., Gloeckner, C. J. & Kahle, P. J. Identification and characterization of ubiquitylation sites in TAR DNA-binding protein of 43 kDa (TDP-43). *J. Biol. Chem.* **293**, 16083–16099 (2018).
 50. Zhang, Y.-J. *et al.* Progranulin mediates caspase-dependent cleavage of TAR DNA binding protein-43. *J. Neurosci.* **27**, 10530–10534 (2007).
 51. Neumann, M. *et al.* Phosphorylation of S409/410 of TDP-43 is a consistent feature in all sporadic and familial forms of TDP-43 proteinopathies. *Acta Neuropathol.* **117**, 137–149 (2009).
 52. Prasad, A., Bharathi, V., Sivalingam, V., Girdhar, A. & Patel, B. K. Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. *Front. Mol. Neurosci.* **12**, 25 (2019).
 53. Estes, P. S. *et al.* Wild-type and A315T mutant TDP-43 exert differential neurotoxicity in a *Drosophila* model of ALS. *Hum. Mol. Genet.* **20**, 2308–2321 (2011).
 54. Bonafede, R. & Mariotti, R. ALS Pathogenesis and Therapeutic Approaches: The Role of Mesenchymal Stem Cells and Extracellular Vesicles. *Front. Cell. Neurosci.* **11**, 80 (2017).

55. Streisinger, G., Walker, C., Dower, N., Knauber, D. & Singer, F. Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*). *Nature* **291**, 293–296 (1981).
56. Dooley, K. Zebrafish: a model system for the study of human disease. *Current Opinion in Genetics & Development* **10**, 252–256 (2000).
57. Howe, K. *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013).
58. Suster, M. L., Kikuta, H., Urasaki, A., Asakawa, K. & Kawakami, K. Transgenesis in Zebrafish with the Tol2 Transposon System. in *Transgenesis Techniques* (ed. Cartwright, E. J.) **561**, 41–63 (Humana Press, 2009).
59. Liu, J. *et al.* CRISPR/Cas9 in zebrafish: an efficient combination for human genetic diseases modeling. *Hum. Genet.* **136**, 1–12 (2017).
60. Kalueff, A. V., Stewart, A. M. & Gerlai, R. Zebrafish as an emerging model for studying complex brain disorders. *Trends Pharmacol. Sci.* **35**, 63–75 (2014).
61. Li, M., Zhao, L., Page-McCaw, P. S. & Chen, W. Zebrafish Genome Engineering Using the CRISPR-Cas9 System. *Trends Genet.* **32**, 815–827 (2016).
62. Tropepe, V. & Sive, H. L. Can zebrafish be used as a model to study the neurodevelopmental causes of autism? *Genes, Brain and Behavior* **2**, 268–281 (2003).
63. Eriksson, K. S., Peitsaro, N., Karlstedt, K., Kaslin, J. & Panula, P. Development of the histaminergic neurons and expression of histidine decarboxylase mRNA in the zebrafish brain in the absence of all peripheral histaminergic systems. *Eur. J. Neurosci.* **10**, 3799–3812 (1998).
64. Guo, S. *et al.* Mutations in the Zebrafish Unmask Shared Regulatory Pathways Controlling

- the Development of Catecholaminergic Neurons. *Developmental Biology* **208**, 473–487 (1999).
65. Guo, S. *et al.* Development of noradrenergic neurons in the zebrafish hindbrain requires BMP, FGF8, and the homeodomain protein soulless/Phox2a. *Neuron* **24**, 555–566 (1999).
 66. Guo, S. *et al.* A regulator of transcriptional elongation controls vertebrate neuronal development. *Nature* **408**, 366–369 (2000).
 67. Holzschuh, J., Ryu, S., Aberger, F. & Driever, W. Dopamine transporter expression distinguishes dopaminergic neurons from other catecholaminergic neurons in the developing zebrafish embryo. *Mech. Dev.* **101**, 237–243 (2001).
 68. McLean, D. L. & Fetcho, J. R. Ontogeny and innervation patterns of dopaminergic, noradrenergic, and serotonergic neurons in larval zebrafish. *J. Comp. Neurol.* **480**, 38–56 (2004).
 69. Ma, P. M. Catecholaminergic systems in the zebrafish. II. Projection pathways and pattern of termination of the locus coeruleus. *J. Comp. Neurol.* **344**, 256–269 (1994).
 70. Ma, P. M. Catecholaminergic systems in the zebrafish. III. Organization and projection pattern of medullary dopaminergic and noradrenergic neurons. *J. Comp. Neurol.* **381**, 411–427 (1997).
 71. Ma, P. M. Catecholaminergic systems in the zebrafish. IV. Organization and projection pattern of dopaminergic neurons in the diencephalon. *The Journal of Comparative Neurology* **460**, 13–37 (2003).
 72. Kaslin, J. & Panula, P. Comparative anatomy of the histaminergic and other aminergic systems in zebrafish (*Danio rerio*). *J. Comp. Neurol.* **440**, 342–377 (2001).

73. Panula, P. *et al.* Modulatory Neurotransmitter Systems and Behavior: Towards Zebrafish Models of Neurodegenerative Diseases. *Zebrafish* **3**, 235–247 (2006).
74. Jalanko, A. & Braulke, T. Neuronal ceroid lipofuscinoses. *Biochim. Biophys. Acta* **1793**, 697–709 (2009).
75. Martín-Jiménez, R., Campanella, M. & Russell, C. New zebrafish models of neurodegeneration. *Curr. Neurol. Neurosci. Rep.* **15**, 33 (2015).

Chapter 2:

A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B)

Chapter 2 is a reprint of the published manuscript:

“A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B).” Liberty François-Moutal, Shahriyar Jahanbakhsh, Andrew D. L. Nelson, Debashish Ray, David D. Scott, Matthew R. Hennefarth, Aubin Moutal, Samantha Perez-Miller, Andrew J. Ambrose, Ahmed Al-Shamari, Philippe Coursodon, Bessie Meechoovet, Rebecca Reiman, Eric Lyons, Mark Beilstein, Eli Chapman, Quaid D. Morris, Kendall Van Keuren-Jensen, Timothy R. Hughes, Rajesh Khanna, Carla Koehler, Joanna Jen, Vijay Gokhale, and May Khanna. ACS Chemical Biology 2018 13(10), 3000–3010. DOI: 10.1021/acscchembio.8b00745.

Shahriyar Jahanbakhsh and Matthew R. Hennefarth performed the *in vivo* cerebellar size determination experiment and generated the accompanying figures and text. The remainder of the manuscript is the work of Liberty François-Moutal and the other authors. Copyright © 2018, American Chemical Society. Reprinted with permission.

Main Manuscript

A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B)

Liberty François-Moutal,^{†,‡} Shahriyar Jahanbakhsh,[§] Andrew D. L. Nelson,^{||} Debashish Ray,[⊥] David D. Scott,^{†,‡} Matthew R. Hennefarth,[§] Aubin Moutal,[†] Samantha Perez-Miller,^{†,‡} Andrew J. Ambrose,[#] Ahmed Al-Shamari,[†] Philippe Coursodon,[†] Bessie Meechoovet,[∇] Rebecca Reiman,[∇] Eric Lyons,^{||} Mark Beilstein,^{||} Eli Chapman,[#] Quaid D. Morris,^{⊥,○,◆,¶} Kendall Van Keuren-Jensen,[∇] Timothy R. Hughes,^{⊥,○} Rajesh Khanna,^{†,‡} Carla Koehler,[§] Joanna Jen,⁺ Vijay Gokhale,[⊗] and May Khanna^{*,†,‡,⊗}

[†]Department of Pharmacology, College of Medicine, University of Arizona, Tucson, Arizona 85724, United States

[‡]Center for Innovation in Brain Science, Tucson, Arizona 85721, United States

[§]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States

^{||}School of Plant Sciences, University of Arizona, Tucson, Arizona 85721, United States

[⊥]Donnelly Centre, University of Toronto, Toronto, Canada M5S 3E1

[#]Pharmacology and Toxicology, School of Pharmacy, University of Arizona, Tucson, Arizona 85724, United States

[∇]Neurogenomics Division, TGen, Phoenix, Arizona 85004, United States

[○]Department of Molecular Genetics, University of Toronto, Toronto, Canada M5S 1A8

[◆]Department of Computer Science, University of Toronto, Toronto, Canada M5S 2E4

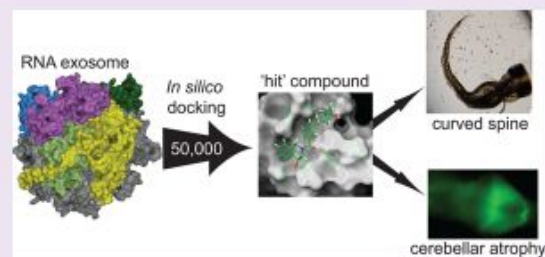
[¶]Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada M5S3G4

⁺Mount Sinai, New York, New York 10029, United States

[⊗]Bio5 Institute, University of Arizona, Tucson, Arizona, United States

Supporting Information

ABSTRACT: Mutations of EXOSC3 have been linked to the rare neurological disorder known as Pontocerebellar Hypoplasia type 1B (PCH1B). EXOSC3 is one of three putative RNA-binding structural cap proteins that guide RNA into the RNA exosome, the cellular machinery that degrades RNA. Using RNAcompete, we identified a G-rich RNA motif binding to EXOSC3. Surface plasmon resonance (SPR) and microscale thermophoresis (MST) indicated an affinity in the low micromolar range of EXOSC3 for long and short G-rich RNA sequences. Although several PCH1B-causing mutations in EXOSC3 did not engage a specific RNA motif as shown by RNAcompete, they exhibited lower binding affinity to G-rich RNA as demonstrated by MST. To test the hypothesis that modification of the RNA–protein interface in EXOSC3 mutants may be phenocopied by small molecules, we performed an *in-silico* screen of 50 000 small molecules and used enzyme-linked immunosorbent assays (ELISAs) and MST to assess the ability of the molecules to inhibit RNA-binding by EXOSC3. We identified a small molecule, EXOSC3-RNA disrupting (ERD) compound 3 (ERD03), which (i) bound specifically to EXOSC3 in saturation transfer difference nuclear magnetic resonance (STD-NMR), (ii) disrupted the EXOSC3–RNA interaction in a concentration-dependent manner, and (iii) produced a PCH1B-like phenotype with a 50% reduction in the cerebellum and an abnormally curved spine in zebrafish embryos. This compound also induced modification of zebrafish RNA expression levels similar to that observed with a morpholino against EXOSC3. To our knowledge, this is the first example of a small molecule obtained by rational design that models the abnormal developmental effects of a neurodegenerative disease in a whole organism.



The RNA exosome is ubiquitous cellular machinery that degrades and processes RNA.^{1–3} The RNA exosome is involved in the processing of small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), and rRNAs

Received: August 11, 2018

Accepted: August 24, 2018

Published: August 24, 2018

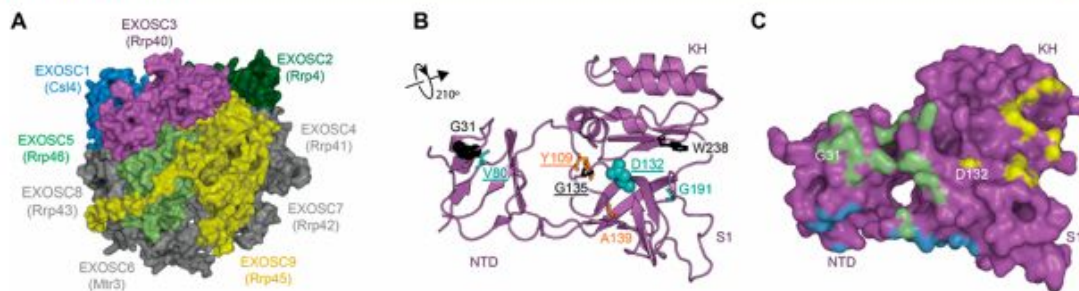


Figure 1. RNA exosome structure highlighting PCH1B-associated mutations in human EXOSC3. (A) The RNA exosome complex (PDB code: 2nn6²⁰) is composed of nine subunits. (B) Cartoon model of EXOSC3. PCH1B-associated amino acids shown as sticks (buried residues) and spheres (surface residues). Residues in black are associated with very poor survival,^{10,13} orange for those wherein survival is reduced when associated with D132A,¹⁰ and cyan for those associated with survival into adulthood.^{10,11,14} Mutations targeted in this study are underlined. NTD, N-terminal domain; S1, RNA binding S1; and KH, K homology domain. (C) Surface representation of EXOSC3 showing predicted contacts with other subunits. Colors refer to the EXOSC3 contacts with EXOSC5 (light green), EXOSC1 (blue), and EXOSC9 (yellow).

(rRNAs).⁴ The overall structure of the RNA exosome core (EXOSC1–9) is conserved from prokaryotes to higher eukaryotes and contains a ring-like structure composed of six RNase Pleckstrin homology (PH) domains that form a central channel able to accommodate single-stranded RNA and three additional proteins that form a “cap-like” structure^{5–8} (Figure 1A). The “cap proteins”—the entry point of RNAs threading through the exosome—are EXOSC1, -2, and -3. EXOSC2 and -3 contain two putative RNA-binding domains, the KH (heterogeneous nuclear ribonucleoprotein K homology) and S1 (ribosomal protein S1) domains, while EXOSC1 only contains an S1 domain; all three proteins have a similar N-terminal domain (NTD; Figures 1B and C). The S1 domains are positioned to bind RNA as it threads through the cap toward the core proteins and central channel of the RNA exosome,⁹ although it is not known if the KH domains also contribute to RNA binding.

Pontocerebellar Hypoplasia 1B (PCH1B) is a recessive disorder characterized by cerebellar hypoplasia, variable pontine atrophy, and progressive microcephaly with global developmental delay.¹⁰ Whole genome sequencing of patients with PCH1B reported several mutations clustered in EXOSC3.^{10–14} Survival of the patients is correlated with the type of mutation and its association with a nonsense allele. Thus far, eight different point mutations have been reported, as either homozygous or heterozygous morphism associated with PCH1B (G31A, V80F, Y109N, D132A, G135E, A139P, and W238R;^{10–14} Figure 1C). The majority of these mutations cluster in the S1 domain, one mutation in the KH domain, and two in the NTD of EXOSC3.

Knockdown of EXOSC3 in zebrafish resulted in a PCH1B-like phenotype associated with common alterations in the expression of genes implicated in development including *ataxin1b* and homeobox gene *HOXC*.¹⁶ Another recent study demonstrated the lack of stability of a yeast analog of EXOSC3 W238R, and its failure to associate efficiently with the exosome in the presence of the wild-type yeast EXOSC3, Rrp40.¹⁸ Yet another study reported sequestration of EXOSC3-D132A in the cytosol of patient fibroblasts and an accumulation of mRNA in muscles of these patients; both were linked to mitochondrial dysfunction.¹⁹ Thus, accumulating evidence points to a link between mutations in EXOSC3 and PCH1B development.

How mutations in EXOSC3 lead to PCH1B is unknown. In this study, we interrogated how EXOSC3 mutations perturb

interactions with other RNA exosome proteins through peptide array analysis and determined the RNA-binding preferences of EXOSC3 and several disease-causing mutations using RNAcompete and microscale thermophoresis (MST). We then used *in-silico* docking to target the RNA-binding interface of EXOSC3 and identified a small molecule, ERD03, able to induce a PCH1B-like phenotype in zebrafish. Using RNaseq and qPCR, we showed that RNA expression patterns were similar between EXOSC3-knockdown and ERD03-treated zebrafish. Finally, incubation of this compound with zebrafish embryos induced an atrophy of the cerebellum. To our knowledge, this is the first example of a small molecule obtained through targeted structure-based design that models a neurodegenerative disease.

RESULTS AND DISCUSSION

Assessing PCH1B Mutation’s Effect on Exosome Integrity. First, we compared the effect of the mutations on protein–protein contacts within the RNA exosome complex. Guided by the three-dimensional structure of the RNA exosome (PDB ID: 2nn6),²⁰ we identified proteins that directly contact EXOSC3 (Figure 1A, Supporting Information Figure S1): EXOSC1 (another cap protein), EXOSC5, and EXOSC9. Among the PCH1B mutations,^{10,12–14} only two out of eight mutations, G31A and D132A, are located near protein–protein interfaces in the structure (Figures 1B,C).

As a disease-causing mutant (EXOSC3-W238R) failed to associate efficiently with the exosome in the presence of the wild-type yeast EXOSC3, Rrp40,¹⁸ without being close to protein–protein interfaces, we sought to determine if other mutations of EXOSC3 (V80F, Y109N, D132A, and G135E) could alter its binding to EXOSC1, EXOSC5, and EXOSC9, either directly or through an allosteric affect. These mutations were selected based on (i) their location—V80F is in the NTD while Y109N, D132A, and G135E are clustered in the S1 domain (Figure 1B)—and (ii) their varying spectrum of patient survival^{10,13,14} (Supporting Information Figure S2A). While homozygous D132A mutation, the most common mutation, is linked to a mild phenotype, heterozygosity of D132A with V80F was found to cause a milder PCH1 phenotype.^{10,13,14} The homozygous G135E and heterozygous D132A/Y109N are found to cause severe phenotypes and are fatal within a few months of infancy.^{10,13}

For all subsequent experiments, we purified EXOSC3-wildtype (EXOSC3-WT) and EXOSC3-V80F, EXOSC3-

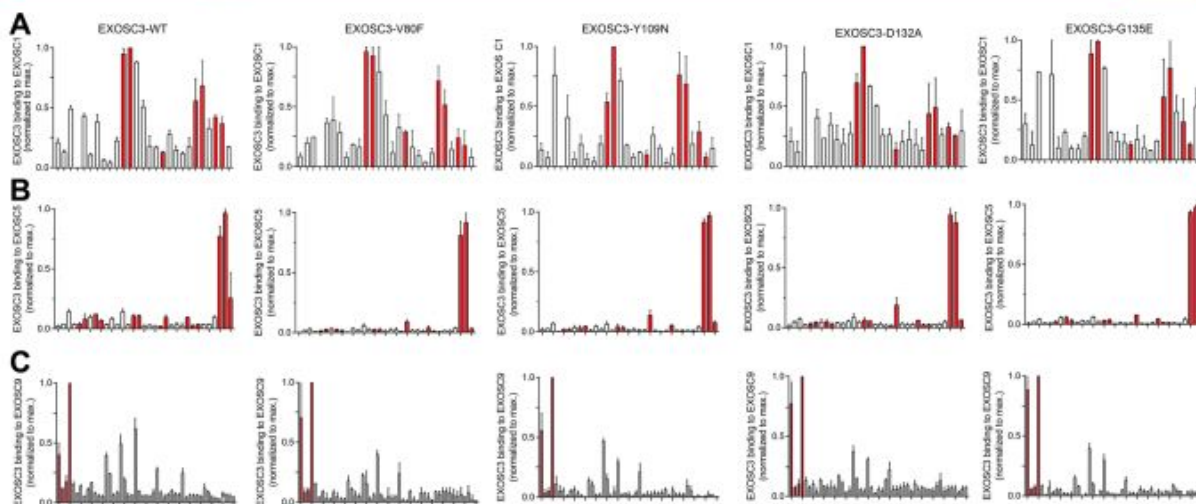


Figure 2. Effect of PCH1B mutations on EXOSC3 binding to adjacent subunits through a peptide array. Binding of EXOSC3-WT and four PCH1B mutants (V80F, Y109N, D132A, G135E) on immobilized 12-mer EXOSC1 (A), EXOSC5 (B), and EXOSC9 (C) peptides, in overlapping eight amino acid steps. The blots were scanned, and spot intensities were quantified and represented as a normalized signal. Predicted contacts with EXOSC3 are shown in red. Data: means \pm SD ($n = 2$).

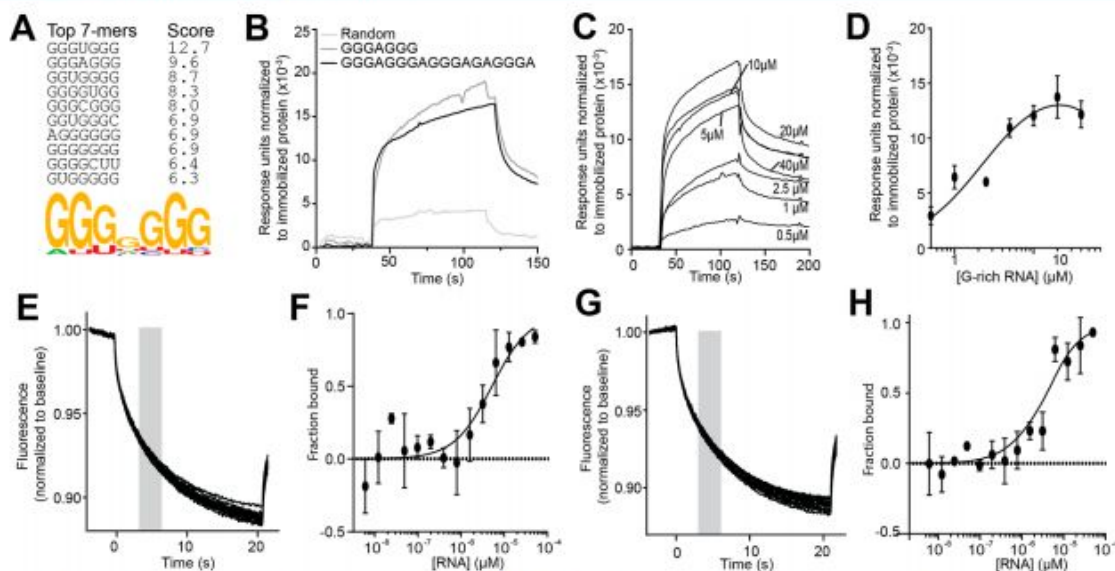


Figure 3. Profiling RNA binding landscape of EXOSC3. (A) RNA binding specificities of EXOSC3 were determined with RNacompete. The top 10 binding 7-mer sequences and their Z-scores are shown. Below, the consensus RNA motif is shown as a position weight matrix. (B) SPR sensorgrams showing binding of 200 μ M long G-rich RNA (GGGAGGGAGGGAGAGGGGA), short G-rich RNA (GGGAGGG), and random RNA (AGUCAUUC) on immobilized EXOSC3. The sensorgrams were normalized to the molecular weight of the respective RNAs. (C) SPR sensorgrams showing binding of increasing concentrations of long G-rich RNA to immobilized EXOSC3 (~ 0.5 ng). (D) SPR data were fit to a one-site specific binding equation ($r^2 = 0.91$) yielding an apparent K_d of $2.4 \pm 1 \mu$ M. Data: means \pm SD ($n = 3$). Some error bars are smaller than the symbols. (E) Thermographs of NT-647-GST-EXOSC3 with 0.0015–50 μ M of long G-rich RNA. The gray bar represents the time point at which the MST measurements were used in F. (F) MST values yielded an apparent K_d of $5.8 \pm 1.3 \mu$ M (see Methods for details). (G) Thermographs of NT-647-GST-EXOSC3 with 0.0015–50 μ M of short G-rich RNA. The gray bar represents the time point at which the MST measurements were used in (H). (H) MST values yielded an apparent K_d of $3.58 \pm 2.5 \mu$ M (see Methods for details).

Y109N, EXOSC3-D132A, and EXOSC3-G135E (Supporting Information Figure S2B). To assess potential effects on protein–protein contacts, we constructed peptide arrays that encompass the three proteins interfacing with EXOSC3: EXOSC1, EXOSC5, and EXOSC9. The data were normalized to a maximum binding peptide that was consistently found in

each array (Figure 2); the points of contacts based on the crystal structure are highlighted in red (Figure 2). There appears to be partial overlap of predictions and array data for EXOSC1 and -9 and very little overlap for EXOSC5 (Figure 2, Supporting Information Figure S3); the predicted contacts may be affected by crystal packing. On the basis of the

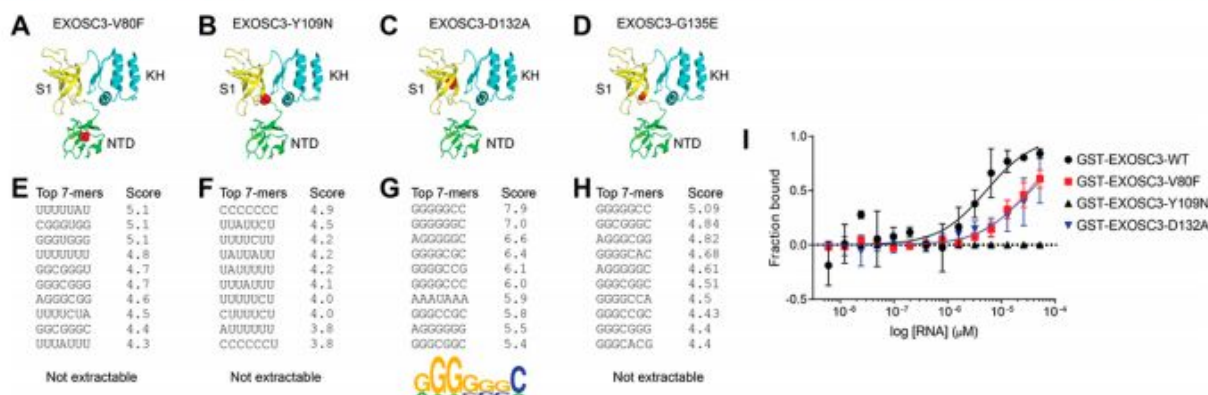


Figure 4. PCH1B mutations on EXOSC3 altering the RNA binding landscape. Cartoon models of EXOSC3 mutants (in red ball–stick representation): V80F (A), Y109N (B), D132A (C), and G135E (D). The top 10 7-mers with their corresponding Z-scores show the RNA binding landscape for EXOSC3 mutants V80F (E), Y109N (F), D132A (G), and G135E (H); an RNA motif as a position weight matrix was extractable for D132A. (I) MST experiments yielded apparent K_d values of $31.8 \pm 11 \mu\text{M}$ (V80F) and $46.8 \pm 19 \mu\text{M}$ (D132A), which were higher compared to $5.8 \pm 1.3 \mu\text{M}$ (EXOSC3-WT). No association could be detected between long G-rich RNA (GGGAGGGAGGGAGAGGGGA) and GST-EXOSC3-Y109N; no binding curve was extractable. Means \pm SD ($n = 3$). Some error bars are smaller than the symbols.

normalized data, there are no differences in the ability of EXOSC3 mutants when compared to wild-type to bind to points of contacts of EXOSC1, -5, and -9 (Figure 2). These results suggest that mutations are unlikely to impact exosome integrity through loss of EXOSC3 contact.

RNA-Binding by EXOSC3 Is Modified by PCH1B Mutations. As the mutations did not affect EXOSC3's ability to interact with RNA exosome subunits and are clustered in the RNA-binding domains, we analyzed their impact on RNA binding using RNAcompete^{21,22} (see Supporting Information Methods) and microscale thermophoresis (MST).

The top 10 7-mers that emerged from EXOSC3-WT converged on a G-rich sequence (top sequences: GGGUGGG and GGGAGGG; Figure 3A, Supporting Information Table 1, Supporting Information Figure S4A). We thus tested the ability of EXOSC3 to bind a short G-rich RNA (GGGAGGG), a random RNA (AGUCAUUC), and a long G-rich sequence (GGGAGGGAGGGAGAGGGGA) using surface plasmon resonance (SPR). After normalizing the signal to the molecular weight of the corresponding RNA, we noted no difference in binding of EXOSC3 to the short and long G-rich sequences, while we observed little binding of the random RNA to EXOSC3 (Figure 3B). Further SPR experiments revealed binding of the long G-rich RNA to EXOSC3 with a dissociation constant (K_d) of $2.4 \pm 1 \mu\text{M}$ (Figure 3C,D). MST showed an apparent K_d of $5.8 \pm 1.3 \mu\text{M}$ of NT647-GST-EXOSC3 with long G-rich RNA (Figure 3E,F) and an apparent K_d of $3.6 \pm 2.5 \mu\text{M}$ of NT647-GST-EXOSC3 with short G-rich RNA (Figure 3G,H).

Together, our results demonstrate that EXOSC3 prefers G-rich RNA sequences. However, it is not clear whether the aforementioned EXOSC3 mutations associated with PCH1B also impact RNA binding. We thus used RNAcompete to analyze the RNA-binding preferences of EXOSC3-V80F, Y109N, D132A, and G135E (Figure 4B–D). Except EXOSC3-D132A, all mutants yielded low confidence RNA-compete data (Figure 4E–H, Supporting Information Table 1, Supporting Information Figure S4B–E). For all EXOSC3 mutants, the relative RNA-bindings, approximated by Z-score values, are reduced compared to EXOSC3-WT (Figure 4E–H and Supporting Information Figure S4). EXOSC3-D132A was

the only mutant that yielded a motif—GGGGGG—similar, but not identical, to the GGGAGGG motif of EXOSC3-WT (Figure 4G).

MST was next used to assess the effect of disease-causing mutations on EXOSC3's binding to long G-rich RNA (Figure 4I). Robust thermomodification of labeled EXOSC3-WT, EXOSC3-V80F, EXOSC3-Y109N, and EXOSC3-D132A was observed; NT647-EXOSC3-G135E had a high propensity to adsorb to the capillaries and therefore was not used further. Compared to a K_d of $5.8 \pm 1.3 \mu\text{M}$ of EXOSC3-WT (Figure 3F), EXOSC3-V80F and D132A decreased the K_d values for RNA to $31.8 \pm 11 \mu\text{M}$ and $46.8 \pm 19 \mu\text{M}$, respectively (Figure 4I). No binding was detected for EXOSC3-Y109N under our conditions (Figure 4I).

Combined, our results demonstrate that EXOSC3 recognizes G-rich sequences and that mutations V80F, Y109N, D132A, and G135E have impaired and/or modified RNA binding. The defects in RNA binding and enrichment of RNA motifs of mutant EXOSC3 proteins shows some correlation with PCH1B disease severity. For example, mutant EXOSC3-V80F has reduced interaction with EXOSC1 and EXOSC9 (peptide array), only slightly decreases RNA binding (MST), and is linked to a good survival rate; mutant EXOSC3-Y109N, linked with a poor survival,¹³ showed no binding to G-rich RNA (MST). Mutant EXOSC3-G135E, also associated with poor survival,¹³ has no enriched binding motif (RNAcompete), and mutant EXOSC3-D132A, which has a similar motif as EXOSC3-WT, is linked to the mildest form of PCH1B.^{10,13}

Rationale for Targeting EXOSC3/RNA Interface. Modeling neurodegenerative diseases is challenging given the complexity arising from development, identification of a vulnerable window of onset, and the multifactorial etiology of these diseases.^{26,27} Compounding the problem is the fact that often neurodegenerative diseases arise in patients due to diverse mutations within protein families, thus making it difficult to create models that capture the full spectrum of the disease. A differentiated approach is to understand the structural and mechanistic basis of diseases. We demonstrated that RNA binding is altered by mutations in the S1 domain of EXOSC3. Next, we asked if targeting RNA/EXOSC3 interfaces with small molecules could mimic the altered

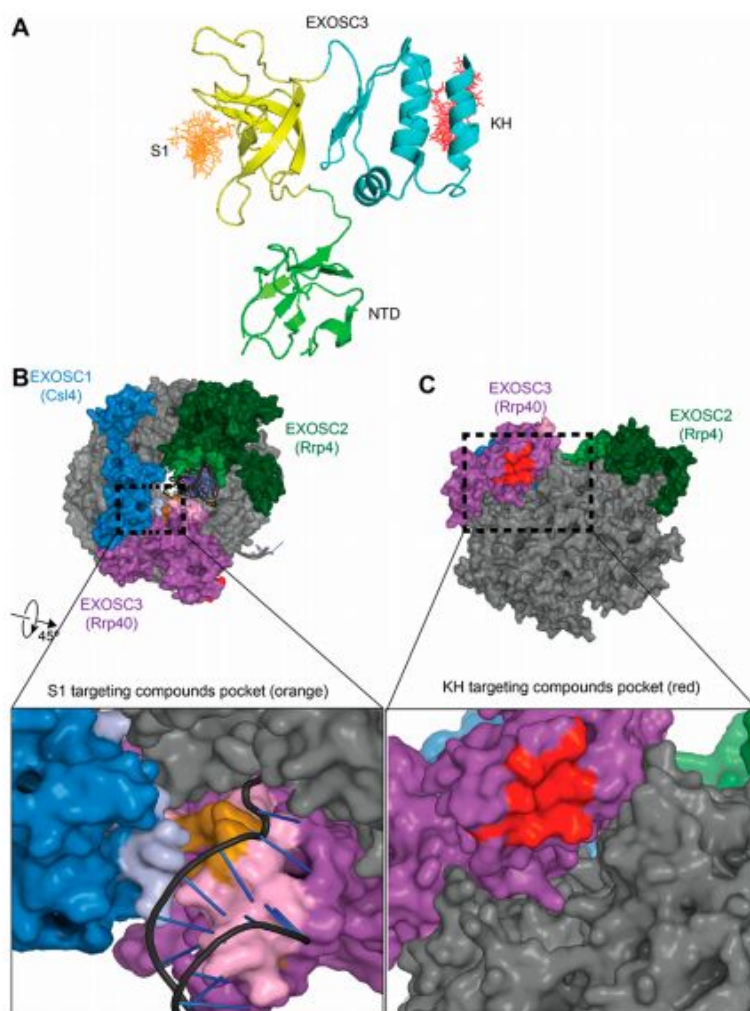


Figure 5. Docking of small molecules onto EXOSC3. (A) The top 20 molecules (stick representation) were docked on the S1 (yellow) or the KH (blue) domain of EXOSC3 (carton representation). (B) RNA exosome with core components colored in gray. Cap proteins EXOSC1, EXOSC2, and EXOSC3 are shown in blue, dark green, and purple, respectively. Close up of S1 domain binding site (orange) showing significant overlap with a predicted RNA interface on EXOSC3 (pink). Cartoon representation of RNA from yeast structure 5jea (for reference). EXOSC2 is omitted in the inset. (C) Rotated view of the RNA exosome showing the docking pocket in the KH domain (red). Close-up of KH targeting compounds docking site showing pocket accessibility.

RNA-binding landscape of EXOSC3 to recapitulate cerebellar atrophy, the hallmark of PCH1B.

Our rationale was that successful identification of such molecules could enrich our understanding on the importance of the RNA-binding domain for EXOSC3's function and lead to induction of a PCH1B-like phenotype in zebrafish embryos, thus creating a new model of this disease. Zebrafish have emerged as a useful model organism for developmental biology, because of their physiologic similarities to humans,²⁸ their fast-developing rate, and their translucent skin during larval stages permitting brain imaging. Previous studies demonstrated that morpholino knockdown of EXOSC3 in zebrafish embryos caused the phenotype of a short, curved spine and underdeveloped brain, associated with poor motility and death by 3 days post-fertilization (dpf).^{10,16} Imaging of the Purkinje cells in EXOSC3-knockdown zebrafish further demonstrated cerebellar defects. Those observations modeled

critical elements of the clinical disease of PCH1B and established EXOSC3-knockdown in the zebrafish as a facile model of this disease.

***In-Silico* Screen of EXOSC3 RNA–Protein Interfaces.**

Using the Schrödinger Suite, we performed an *in-silico* docking experiment targeting residues in the two RNA binding domains—Gln244, Asn251, Ile252, Leu271, and Ser274 in the KH domain and Phe146, Glu147, Arg152, Asn188, and Met190 in the S1 domain. A 50 000-compound library was docked on a 10 Å grid surrounding these residues (Figure 5) using Glide. On the basis of this docking, two pockets emerged in the KH and S1 domains of EXOSC3 (Figure 5A). The S1-targeting compounds appeared to cluster in a region where RNA threads into the exosome (Figure 5B). The predicted RNA binding interface and the molecule binding site show a modest overlap (Figure 5B, inset). While the KH domain compounds are not positioned directly in the path of RNA

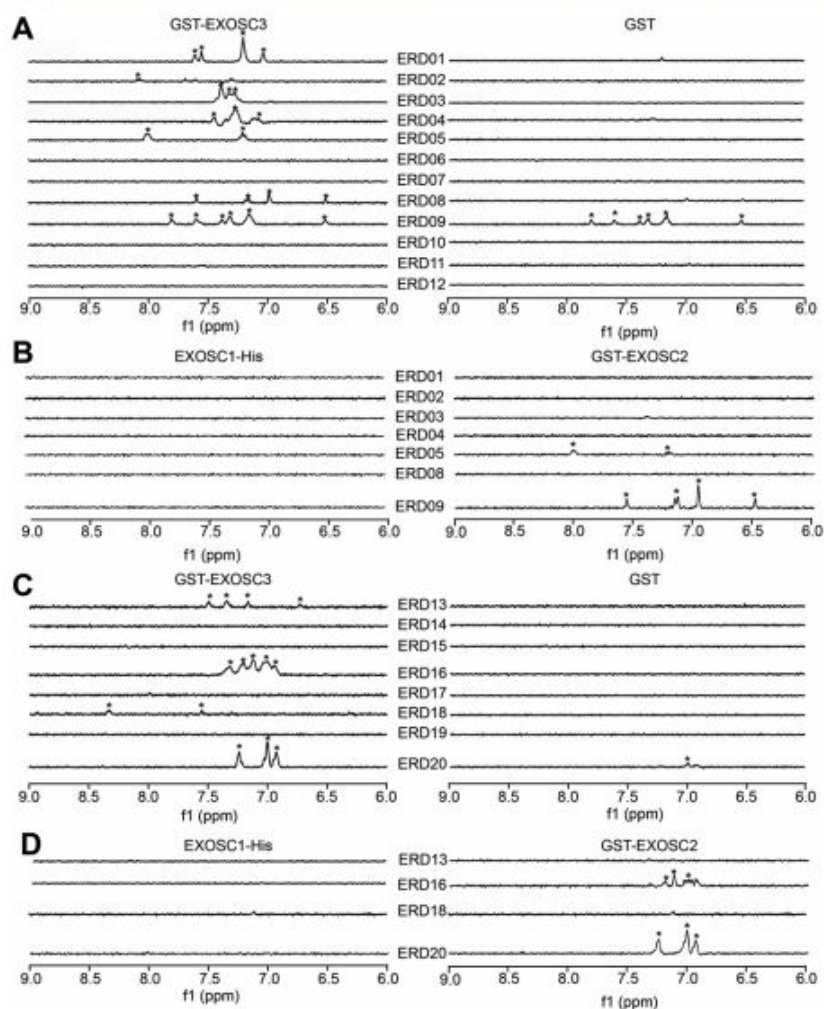


Figure 6. STD-NMR analysis of binding of small molecules to exosome subunits. 1D ^1H STD-NMR showing on-resonance difference spectrum of 500 μM (EXOSC3/RNA disrupting (ERD)) compounds with 5 μM proteins. The aromatic region of the spectrum (6–9 ppm) is shown. (A) 1D ^1H STD-NMR spectra of KH domain-targeting ERD01–ERD12 compounds binding to GST-EXOSC3 (left) or GST (right). (B) 1D ^1H STD-NMR of ERD compounds that bound to GST-EXOSC3, but not GST, were tested against EXOSC1-His (left) and GST-EXOSC2 (right). (C, D) Same as A and B except with compounds ERD13–ERD20 (targeting S1 domain). (*protons reproducibly found in independent experiments ($n = 3$)).

threading into the exosome, the KH domain RNA-binding interface is accessible and may be necessary for RNA binding (Figure 5C).

Biophysical Triage of Putative EXOSC3-RNA-Disrupting (ERD) Compounds. The top 20 selected compounds from *in-silico* docking (Supporting Figure S5, Supporting Table 2) were divided into molecules that target the KH domain (EXOSC3 RNA Disruptors ERD01–ERD12) and molecules that target the S1 interface (ERD13–ERD20).

Following docking, saturation transfer difference NMR (STD-NMR) was used to test the ability of compounds to bind to EXOSC3. GST was used as a control for off-target binding. STD-NMR demonstrated that seven compounds targeting KH (ERD 01, 02, 03, 04, 05, 08, 09) and four compounds targeting S1 (ERD13, 16, 18, 20) bound EXOSC3 (Figure 6A,C). Compounds ERD09 and ERD20 also bound GST, typifying pan assay interference compounds (PAINS)^{30,31} (Figure 6A and C, right panels). Compounds

ERD16, ERD20, and ERD05 also bound to GST-EXOSC2, but not to EXOSC1-His (Figure 6B,D).

With the above biophysical triage steps, we identified several molecules that met key criteria: (i) binding to EXOSC3 but not EXOSC1 or EXOSC2; (ii) no binding to GST; and (iii) did not fit the structure of PAINS compounds (“false positive” in high throughput screen). Selected compounds were ERD01, ERD02, ERD03, ERD04, and ERD08 (for the KH domain) and ERD13 and ERD18 (for the S1 domain).

Disruption of EXOSC3/RNA Interactions with ELISA. We next asked if these compounds, used at 50 μM , could disrupt EXOSC3–RNA interactions using an enzyme-linked immunosorbent assay (ELISA). As expected, binding was observed between G-rich RNA and EXOSC3 (Figure 7A). Among compounds that target the KH domain, ERD01 was the most effective with ~45% inhibition, followed by ERD04 and ERD02 (~38% and 35% inhibition, respectively), and ERD03 with ~18% inhibition. Among compounds that target

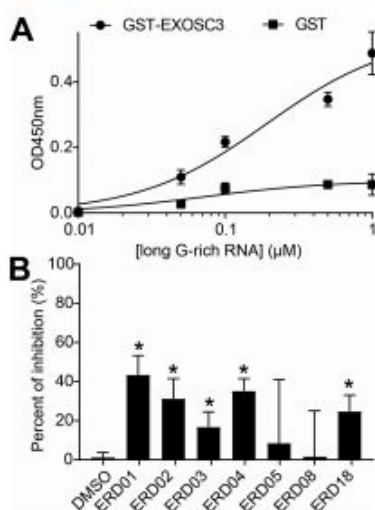


Figure 7. RNA/EXOSC3 disruption by small molecules. (A) 300 nM GST-EXOSC3 was incubated with increasing concentrations of long G-rich RNA. The data were fit to a one-site specific binding equation ($r^2 = 0.94$). Long G-rich RNA bound to EXOSC3 with a half-saturation concentration of ~ 100 nM. Binding of GST alone demonstrated minimal background binding to RNA. Data: means \pm SD ($n = 6$). Some error bars are smaller than the symbols. (B) ERD compounds specifically binding to EXOSC3 were tested for their ability to disrupt EXOSC3/RNA interaction. 50 μ M ERDs were incubated with 0.2 μ M RNA and GST-EXOSC3. Data: mean \pm SD (*one-way ANOVA, $p < 0.05$, $n = 6-12$).

the S1 domain, ERD18 inhibited binding by $\sim 30\%$ (Figure 7B).

An ERD Compound Recapitulates PCH1B Phenotype in Development in Zebrafish Embryos. To examine possible functional consequences of ERD compounds, we incubated zebrafish embryos with selected ERDs and monitored their phenotype. No major differences are expected between the structures of human and zebrafish EXOSC3 proteins in terms of the sites that were targeted by the small molecules (Supporting Information Figure S6). A morpholino that targets the AUG translational initiation site (AUG morpholino) resulted in a pronounced phenotype with most zebrafish having a curved spine and a high incidence of death, similar to the phenotype previously reported;¹⁰ control morpholinos did not affect zebrafish development (Figure 8A,B). Of five ERD compounds tested in zebrafish, ERD18 and ERD04 led to $\sim 75\%$ and 90% death, respectively, by 3–5 days postfertilization (dpf), whereas ERD01, ERD02, and ERD03 induced death equivalent to DMSO. Only ERD03 caused significant spinal curvature (Figure 8A–C) and was thus selected as the lead compound that might phenocopy the PCH1B disease in zebrafish.

Characterization of Binding Affinity and Inhibition of RNA-EXOSC3 Interaction by ERD03. We showed that ERD03 binds specifically EXOSC3 (Figure 6), disrupts the EXOSC3/RNA interface (Figure 7), and causes a spinal curvature in zebrafish (Figure 8). To further characterize ERD03, we measured ERD03's binding affinity for EXOSC3 using MST. Photobleaching experiments revealed that ERD03 bound specifically to EXOSC3 with an apparent K_d of 17 ± 7 μ M (Figure 9A and B); no binding to GST was observed.

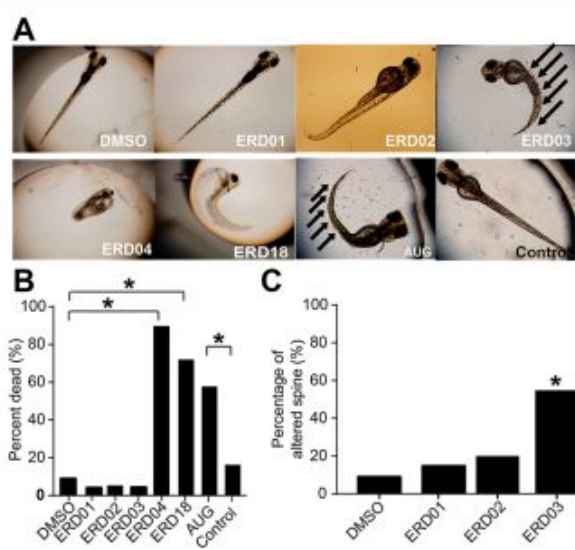


Figure 8. ERD compounds induce a PCH1B-like phenotype in zebrafish. (A) Zebrafish embryos were incubated with 50 μ M ERDs, DMSO (0.5%), or injected with morpholinos (AUG vs control) for 5 days before imaging. (B) The percent of dead zebrafish, assessed as having no heartbeat, is shown for each condition. (C) Bar graphs of percentage of zebrafish with short, curved spine (* χ^2 test, $p < 0.001$, $n \geq 50$).

Next, we measured inhibition of RNA/EXOSC3 by ERD03 using MST. We first incubated EXOSC3 with several concentrations of long G-rich RNA with DMSO or 50 μ M ERD03 (Figure 9C). ERD03 decreased EXOSC3 binding to long G-rich RNA at several concentrations of RNA. At a concentration of 3.1 μ M of RNA, ERD03 inhibited binding by 54% (Figure 9C, arrow). This 3.1 μ M concentration of RNA was then used against several concentrations of ERD03. Data indicated a concentration-dependent inhibition for ERD03 (Figure 9D). Inhibition was also measured using ELISA but yielded less inhibition, albeit still with a concentration dependency (Supporting Information Figure S7A). We also tested the effect of ERD03 on EXOSC3 interaction with short G-rich RNA using MST, and we observed an inhibition similar to long G-rich RNA at 50 μ M of ERD03 (Supporting Information Figure S7B).

To examine whether the compounds interfered with protein–protein interactions in the RNA exosome proteins, we tested EXOSC3's interactions with EXOSC1, EXOSC5, and EXOSC9 using peptide arrays in the presence of ERD03 and ERD01: no major differences were found between the peptide–protein interactions of EXOSC3 with the compounds (Supporting Information Figure S8).

Optimization of ERD03 to gain better potency would be advantageous in order to more finely tune EXOSC3/RNA inhibition. The docking model of ERD03 (Supporting Information Figure S5) with EXOSC3 provides us a starting point for lead optimization. The structure of ERD03 consists of an isoquinoline ring with a butylbenzoyl group attached to a nitrogen atom (Supporting Information Table 2), and an initial lead optimization program could focus on the butylbenzoyl group with an aim to gain additional hydrogen bonding and other interactions in the active site.

ERD03 Recapitulates EXOSC3 Knockdown Effects on RNA Transcriptomic Profiles in Zebrafish. Knockout of

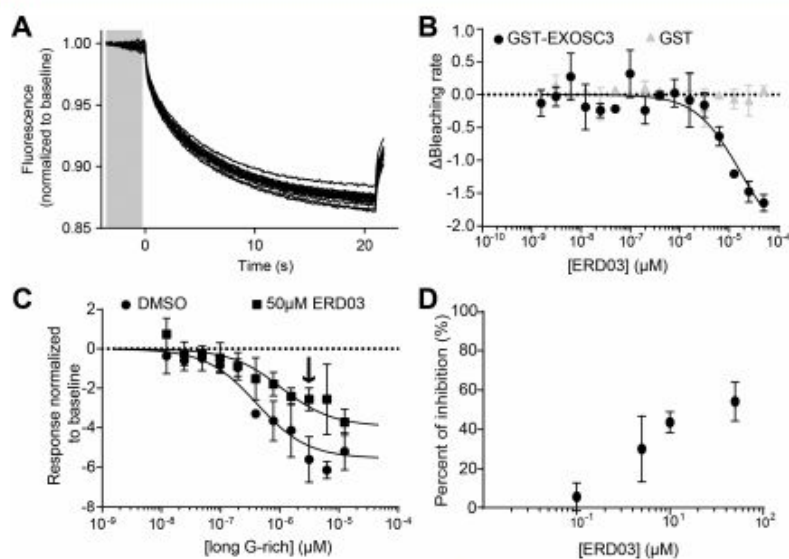


Figure 9. Biophysical characterization of ERD03. (A) Thermographs of NT-647-GST-EXOSC3 with ranging concentrations of ERD03 (0.00153–50 μM). The gray bar represents the time point at which the MST measurements were used in B. (B) Photobleaching rates of EXOSC3's initial fluorescence in the presence of 0.00153–50 μM ERD03 were used to determine ERD03 affinity for EXOSC3. The data were fitted (see Methods) with an apparent $K_d = 17 \pm 7 \mu\text{M}$ ($r^2 = 0.96$). GST protein was used as a negative control. Data: mean \pm SD ($n = 3$). (C) MST values of EXOSC3 with 0.012–12.5 μM of long G-rich RNA and 1% (v/v) DMSO or 50 μM ERD03. Arrow denotes 3.12 μM RNA, which caused $54 \pm 10\%$ inhibition of EXOSC3-RNA by ERD03. (D) Concentration-dependent curves were obtained for ERD03's disruption of RNA-EXOSC3 interaction at 3.12 μM of RNA. Data are represented as mean \pm SD ($n = 3$).

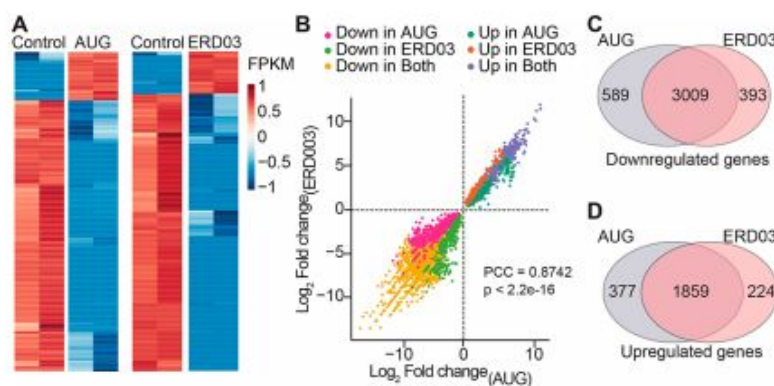


Figure 10. Transcriptomics in zebrafish after ERD03 incubation or EXOSC3-knockdown. (A) Heatmap of \log_2 FPKM (fragments per kilobase of transcript per million mapped reads) values for the top 1000 differentially expressed genes in morpholino knockdown samples (AUG) or samples treated with ERD03 compared to control. FPKM represents the absolute value of a transcript expression. (B) Distribution of \log_2 fold-change values for all differentially expressed genes (relative to control with an adjusted $p < 0.01$) in the ERD03-treated and AUG-knockdown samples. Values are color-coded based on the treatment in which they were significantly differentially expressed (relative to controls). The Pearson correlation coefficient (PCC) was used to assess similarity between AUG and ERD03 treatments ($p < 2.2e-16$). Venn diagram depicting distribution of overlapping downregulated (C) or upregulated (D) RNAs in the treatments. Full data: <https://www.ncbi.nlm.nih.gov/sra-bioproject>, ID PRJNA470927.

the EXOSC3 in mouse and in zebrafish has been previously shown to induce accumulation of certain types of RNA including long noncoding RNA (lncRNA).^{16,32,35} To address if ERD03 does the same, RNaseq analysis was performed on 3 dpf zebrafish incubated with ERD03 or injected with AUG morpholino (Figure 10, Supporting Information Table 3). Most transcripts appear to be affected in a similar fashion by AUG injection and ERD03 incubation compared to the control (Figure 10A). Change in relative expression of each transcript, expressed as \log_2 fold change, correlated well between the two conditions (Figure 10B) as demonstrated by

a Pearson correlation factor of 0.87 ($p < 2.2e-16$). A total of 3009 RNAs were downregulated and 1859 RNAs were upregulated in both AUG and ERD03 treatments (Figure 10C,D), leading to an $\sim 80\%$ overlap of differentially expressed RNAs (75%, $p < 0.01$; 83%, $p < 0.05$). As previously described, lncRNAs constituted a greater percentage of the upregulated vs downregulated genes (13.5% vs 6%, $p < 0.001$ by Fisher's exact test). There is substantial similarity and overlap in the effect on lncRNAs between EXOSC3 knockdown and ERD03 treatment. The observed strong positive correlation in gene expression suggests that pathways downstream of EXOSC3

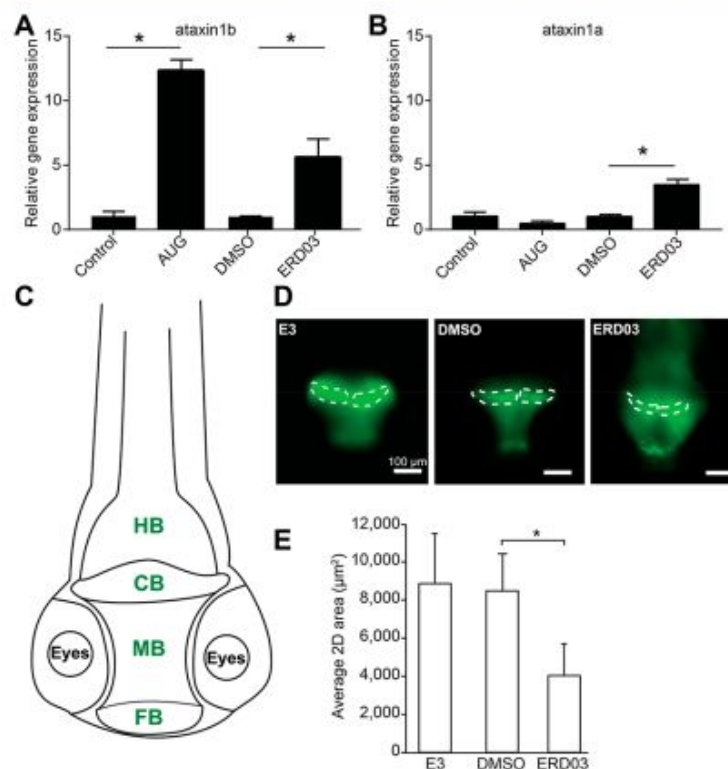


Figure 11. ERD03 effect on cerebellum development. Gene expression analysis of *atxn1b* (A) and *atxn1a* (B) in zebrafish following ERD03 incubation or EXOSC3-knockdown. Data: mean \pm SD (*one-way ANOVA, $p < 0.001$, $n \geq 50$). (C) Schematic depicting the localization of the zebrafish brain. The TDL6 zebrafish line expresses GFP in neurons of the central nervous system. Green text represents tissues expressing GFP. HB, hindbrain; CB, cerebellum; MB, midbrain; FB, forebrain. (D) Fluorescent images of 3dpf zebrafish embryos' brains incubated with E3 buffer, DMSO (0.5%, v/v), or 50 μ M ERD03. The area used to calculate cerebellum average size is represented in white within each image. (E) Average cerebellum size per 2D cross-section is shown. Data: mean \pm SD (*one-way ANOVA with Tukey HSD posthoc test, $p < 0.05$, $n = 3$).

are being impacted in a similar manner and degree between the ERD03 and EXOSC3 knockdown backgrounds. These data further support our assertion that ERD03 can be used as a pharmacological tool to modulate EXOSC3 activity and mimic a PCH1B phenotype.

ERD03 Induces an Atrophy of the Zebrafish Cerebellum, a Clinical Feature of PCH1B. PCH1B patients exhibit cerebellar atrophy with dysmorphic Purkinje cells,¹⁰ and EXOSC3 knockdown in zebrafish previously showed significant cerebellum defects.¹⁶ Ataxin, a DNA-binding protein encoded by the *ATXN1* gene, is important for differentiation of Purkinje cells in the mouse,³⁴ and defects in *ATXN1* were linked to spinocerebellar ataxia type 1,³⁵ a disorder similar to PCH1B, characterized by progressive movement problems due to Purkinje cell degeneration and subsequent cerebellar atrophy.³⁶ Furthermore, *atxn1b* mRNA was previously reported to be significantly upregulated with EXOSC3-knockdown in zebrafish embryos and was linked to a toxic effect on the cerebellum development.¹⁶ We performed quantitative RT-PCR (qPCR) on RNA extracted from zebrafish incubated with ERD03, DMSO control, or injected with AUG morpholino (Figure 11A). ERD03 induced a ~6-fold upregulation of *atxn1b* mRNA (Figure 11A). Although *atxn1a* mRNA was not modified by AUG morpholino, there is a minor accumulation of *atxn1a* mRNA in embryos with ERD03 (Figure 11B).

Because of the ability of ERD03 to induce *atxn1b* mRNA accumulation in zebrafish, linked to an underdeveloped cerebellum, we imaged zebrafish brains in the presence of ERD03. We used the Tübingen driver line (TDL6) zebrafish line where galactose (Gal4)-driven green-fluorescent protein (GFP) expression marks neurons of the central nervous system (Figure 11C).³⁷ We incubated TDL6 zebrafish embryos with either buffer E3, DMSO, or 50 μ M ERD03 and imaged the brains at 3 dpf (Figure 11D). Incubation with ERD03 resulted in zebrafish cerebella half the size of control brains in DMSO (Figure 11D and E). Taken together, these results demonstrate that disrupting the EXOSC3–RNA interaction in zebrafish embryos recapitulates several hallmarks of PCH1B: (i) modification of RNA transcription profiles, (ii) spine curvature, and (iii) accumulation of *atxn1b* mRNA combined with an atrophy of the cerebellum. On the basis of these observations, we propose that ERD03 could be used as a chemical tool to model the progression of PCH1B neurological disease through disruption of the EXOSC3/RNA interactions. Further studies with ERD03 will shed light on mechanisms of RNA processing and development of the cerebellum. We may uncover additional mechanisms to the RNA displacement since ERD03 binding is 3-fold weaker than RNA binding to EXOSC3; alternatively, such a dramatic phenotype in the zebrafish may be due to cumulative RNA defects.

The translational significance of this work is the possibility of discovering small molecules to manage progression of

PCH1B-phenotype in model organisms and could be applicable to other neurological disorders. This chemical approach to mimic neurological disorders caused by mutations may be an instructive example of how to model rare diseases and could be utilized widely in the drug discovery field to develop testable mechanistic hypotheses and drive therapeutic development. To our knowledge, this is the *first* example of a small molecule obtained by rational design that models the abnormal developmental effects of a neurodegenerative disease in a whole organism and could become an invaluable tool when genetic modifications are difficult to obtain in conventional model species.

EXPERIMENTAL PROCEDURES

See the Supporting Information Methods.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.8b00745.

RNAcompete 7-mer sequences and Z-scores (XLSX)
RNaseq results (XLSX)

experimental procedures; detailed map of interactions between EXOSC3, EXOSC1, EXOSC5, and EXOSC9; stability of disease causing mutations on EXOSC3; raw data of peptide arrays between EXOSC3 and EXOSC1, EXOSC5, and EXOSC9; Z-score distribution of RNA 7-mers from RNAcompete on EXOSC3 proteins; top hits from *in silico* docking poses for each domain; comparison of human structures of EXOSC3 with homology model of zebrafish EXOSC3; ERD03 disruption of EXOSC3-RNA interaction by MST; effect of ERD03 and ERD18 binding to adjacent subunits assayed by peptide array; compound structures of top hits from *in silico* docking; primers used for protein cloning; primers used to generate EXOSC3 mutants; qPCR primers used to assess mRNA levels in zebrafish embryos (PDF)

AUTHOR INFORMATION

Corresponding Author

*Address: Department of Pharmacology, College of Medicine, University of Arizona, 1501 North Campbell Drive, P.O. Box 245050, Tucson, AZ 85724, USA. Office phone: (520) 626-2147. Fax: (520) 626-2204. E-mail: maykhanna@email.arizona.edu.

ORCID

Andrew D. L. Nelson: 0000-0001-9896-1739

Andrew J. Ambrose: 0000-0002-2932-4514

May Khanna: 0000-0002-9989-6374

Author Contributions

L.F.-M., S.J., A.D.L.N., D.R., D.D.S., M.H., P.C., and V.G. performed the experiments; L.F.-M. and M.K. wrote the manuscript; and all of the authors participated in revising the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by grants from the 2014 AHSC-CDA Arizona Health Sciences Career Development Award and AZ ABRC ADHS16-162407 Arizona grant (to M.K.); NSF grant IOS 1444490 (to E.L. and M.B.); NIH R01HG008613 (to T.R.H. and Q.D.M.) and NIH R01GM61721; and USAMRAA W81XWH-17-1-0333 (to C.K.).

REFERENCES

- (1) Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3' to 5' exoribonucleases. *Cell* 91, 457–466.
- (2) Decker, C. J. (1998) The exosome: a versatile RNA processing machine. *Curr. Biol.* 8, R238–R240.
- (3) Anderson, J. S. J., and Parker, R. (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SK12 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.* 17, 1497–1506.
- (4) Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* 18, 5399–5410.
- (5) Makino, D. L., Baumgärtner, M., and Conti, E. (2013) Crystal structure of an rna-bound 11-subunit eukaryotic exosome complex. *Nature* 495, 70–75.
- (6) Kowalinski, E., Kögel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., and Conti, E. (2016) Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. *Mol. Cell* 63, 125–134.
- (7) Makino, D. L., and Conti, E. (2013) Structure determination of an 11-subunit exosome in complex with RNA by molecular replacement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 69, 2226–2235.
- (8) Wasmuth, E. V., Januszky, K., and Lima, C. D. (2014) Structure of an Rps6-RNA exosome complex bound to poly(A) RNA. *Nature* 511, 435–439.
- (9) Januszky, K., and Lima, C. D. (2010) Structural components and architectures of RNA exosomes. *Adv. Exp. Med. Biol.* 702, 9–28.
- (10) Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schöneborn, S., Menezes, M. P., Hong, J. E., Leong, D. W., Senderek, J., Salman, M. S., Chitayat, D., Seeman, P., Von Moers, A., Graul-Neumann, L., Kornberg, A. J., Castro-Gago, M., Sobrido, M. J., Sanefuji, M., Shieh, P. B., Salamon, N., Kim, R. C., Vinters, H. V., Chen, Z., Zerres, K., Ryan, M. M., Nelson, S. F., and Jen, J. C. (2012) Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.* 44, 704–708.
- (11) Halevy, A., Lerer, I., Cohen, R., Kornreich, L., Shuper, A., Gamliel, M., Zimmerman, B. El, Korabi, I., Meiner, V., Straussberg, R., and Lossos, A. (2014) Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. *J. Neurol.* 261, 2165–2169.
- (12) Schwabova, J., Brozkova, D. S., Petrak, B., Mojziso, M., Pavlickova, K., Haberlova, J., Mrazkova, L., Hedvicakova, P., Hornofova, L., Kaluzova, M., Fenc, F., Krutova, M., Zamecnik, J., and Seeman, P. (2013) Homozygous EXOSC3 mutation c.92G→C, p.G31A is a founder mutation causing severe pontocerebellar hypoplasia type 1 among the Czech Roma. *J. Neurogenet.* 27, 163–169.
- (13) Eggens, V. R. C., Barth, P. G., Niermeijer, J. M. F., Berg, J. N., Darin, N., Dixit, A., Fluss, J., Foulds, N., Fowler, D., Hortobágyi, T., Jacques, T., King, M. D., Makrythanasis, P., Máté, A., Nicoll, J. A. R., O'Rourke, D., Price, S., Williams, A. N., Wilson, L., Suri, M., Sztroiha, L., Dijns-De Wissel, M. B., Van Meegen, M. T., Van Ruisven, F., Aronica, E., Troost, D., Majoie, C. B. L. M., Marquering, H. A., Poll-Thé, B. T., and Baas, F. (2014) EXOSC3 mutations in pontocerebellar hypoplasia type 1: Novel mutations and genotype-phenotype correlations. *Orphanet J. Rare Dis.* 9, 23.
- (14) Zanni, G., Scotton, C., Passarelli, C., Fang, M., Barresi, S., Dallapiccola, B., Wu, B., Gualandi, F., Ferlini, A., Bertini, E., and Wei,

- W. (2013) Exome sequencing in a family with intellectual disability, early onset spasticity, and cerebellar atrophy detects a novel mutation in EXOSC3. *Neurogenetics* 14, 247–250.
- (15) Boczonadi, V., Müller, J. S., Pyle, A., Munkley, J., Dor, T., Quartararo, J., Ferrero, I., Karcagi, V., Giunta, M., Polvikoski, T., Birchall, D., Princzinger, A., Cinnamon, Y., Lützkendorf, S., Piko, H., Reza, M., Florez, L., Santibanez-Koref, M., Griffin, H., Schuelke, M., Elpeleg, O., Kalaydjieva, L., Lochmüller, H., Elliott, D. J., Chinnery, P. F., Edvardson, S., and Horvath, R. (2014) EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia. *Nat. Commun.*, 5, DOI: 10.1038/ncomms5287.
- (16) Giunta, M., Edvardson, S., Xu, Y., Schuelke, M., Gomez-Duran, A., Boczonadi, V., Elpeleg, O., Müller, J. S., and Horvath, R. (2016) Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. *Hum. Mol. Genet.* 25, 2985–2996.
- (17) Gillespie, A., Gabunilas, J., Jen, J. C., and Chanfreau, G. F. (2017) Mutations of EXOSC3/Rp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome. *RNA* 23, 466–472.
- (18) Fasken, M. B., Losh, J. S., Leung, S. W., Brutus, S., Avin, B., Vaught, J. C., Potter-Birriel, J., Craig, T., Conn, G. L., Mills-Lujan, K., Corbett, A. H., and van Hoof, A. (2017) Insight into the RNA exosome complex through modeling pontocerebellar hypoplasia type 1b disease mutations in yeast. *Genetics* 205, 221–237.
- (19) Schottmann, G., Picker-Minh, S., Schwarz, J. M., Gill, E., Rodenburg, R. J. T., Stenzel, W., Kaindl, A. M., and Schuelke, M. (2017) Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia. *Mitochondrion* 37, 46–54.
- (20) Liu, Q., Greimann, J. C., and Lima, C. D. (2006) Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. *Cell* 127, 1223–1237.
- (21) Ray, D., Ha, K. C. H., Nie, K., Zheng, H., Hughes, T. R., and Morris, Q. D. (2017) RNAcompete methodology and application to determine sequence preferences of unconventional RNA-binding proteins. *Methods* 118–119, 3–15.
- (22) Ray, D., Kazan, H., Chan, E. T., Castillo, L. P., Chaudhry, S., Talukder, S., Blencowe, B. J., Morris, Q., and Hughes, T. R. (2009) Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nat. Biotechnol.* 27, 667–670.
- (23) Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L. H., Dale, R. K., Smith, S. A., Yarosh, C. A., Kelly, S. M., Nabet, B., Mecnas, D., Li, W., Laishram, R. S., Qiao, M., Lipshitz, H. D., Piano, F., Corbett, A. H., Carstens, R. P., Frey, B. J., Anderson, R. A., Lynch, K. W., Penalva, L. O. F., Lei, E. P., Fraser, A. G., Blencowe, B. J., Morris, Q. D., and Hughes, T. R. (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177.
- (24) Lunde, B. M., Moore, C., and Varani, G. (2007) {RNA}-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* 8, 479–490.
- (25) Sauer, E. (2013) Structure and RNA-binding properties of the bacterial LSm protein Hfq. *RNA Biol.* 10, 610.
- (26) Noble, W., and Burns, M. P. (2010) Challenges in neurodegeneration research. *Front. Psychiatry*, DOI: 10.3389/fpsy.2010.00007.
- (27) Young, A. B. (2009) Four decades of neurodegenerative disease: How far we have come! *J. Neurosci.* 29, 12722.
- (28) Cao, Y., Jiang, L., Zhao, L., Zhou, X., Wang, N., Zhang, P., Tang, Y., and Zhou, J. (2015) Evaluation of the in vivo safety profiles of Rictor inhibition using a zebrafish model. *Curr. Pharm. Des.* 21, 1645–1653.
- (29) Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* 49, 6177–6196.
- (30) Dahlin, J. L., Nissink, J. W. M., Strasser, J. M., Francis, S., Higgins, L., Zhou, H., Zhang, Z., and Walters, M. A. (2015) PAINS in the assay: Chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J. Med. Chem.* 58, 2091–2113.
- (31) Capuzzi, S. J., Muratov, E. N., and Tropsha, A. (2017) Phantom PAINS: Problems with the Utility of Alerts for PAINS in assay interference Compound S. *J. Chem. Inf. Model.* 57, 417–427.
- (32) Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A. N., and Basu, U. (2014) Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* 514, 389–393.
- (33) Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao, J., Elliott, O., Liu, Z. P., Economides, A. N., Bradner, J. E., Rabadan, R., and Basu, U. (2015) RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* 161, 774–789.
- (34) Ebner, B. a., Ingram, M. A., Barnes, J. A., Duvick, L. A., Frisch, J. L., Clark, H. B., Zoghbi, H. Y., Ebner, T. J., and Orr, H. T. (2013) Purkinje cell ataxin-1 modulates climbing fiber synaptic input in developing and adult mouse cerebellum. *J. Neurosci.* 33, 5806–5820.
- (35) Matilla-Dueñas, A., Goold, R., and Giunti, P. (2008) Clinical, genetic, molecular, and pathophysiological insights into spinocerebellar ataxia type 1. *Cerebellum* 7, 106–14.
- (36) Huang, M., and Verbeek, D. S. (2018) Why do so many genetic insults lead to Purkinje Cell degeneration and spinocerebellar ataxia? *Neurosci. Lett.*, DOI: 10.1016/j.neulet.2018.02.004.
- (37) Levesque, M. P., Krauss, J., Koehler, C., Boden, C., and Harris, M. P. (2013) New Tools for the Identification of Developmentally Regulated Enhancer Regions in Embryonic and Adult Zebrafish. *Zebrafish* 10, 21–29.

Supporting Information

Supplementary Information

A Chemical Biology Approach to Model Pontocerebellar Hypoplasia Type 1B (PCH1B)

Liberty François-Moutal^{1,2}, Shahriyar Jahanbakhsh³, Andrew Nelson⁴, Debashish Ray⁵, David D. Scott^{1,2}, Matthew Hennefarth³, Aubin Moutal¹, Samantha Perez-Miller^{1,2}, Andrew J. Ambrose⁶, Ahmed Al-Shamari¹, Philippe Coursodon¹, Bessie Meechoovet⁷, Rebecca Reiman⁷, Eric Lyons⁴, Mark Beilstein⁴, Eli Chapman⁶, Quaid D. Morris^{5,8-10}, Kendall Van Keuren-Jensen⁷, Timothy R. Hughes^{5,8}, Rajesh Khanna^{1,2}, Carla Koehler³, Joanna Jen¹¹, Vijay Gokhale¹², and May Khanna^{1,2*}

Methods in detail

Materials

All reagents were purchased from Sigma (St. Louis, MO, USA) and Fisher Scientific (Hampton, NH) unless otherwise indicated.

Cloning of EXOSC3

The coding sequence of human EXOSC1, EXOSC2 and EXOSC3 were amplified by PCR using primers listed in Supplementary Table 4 (Eurofins Genomics, Louisville, KY), and subcloned using the CloneTech in-Fusion kit into the pGEX-6-P1 vector between *Bam*HI and *Xho*I restriction sites, resulting in a construct with an N-terminal GST tag. Amplified sequences were verified by DNA sequencing. The coding sequence of EXOSC1 amplified using another set of primers (listed in Supplementary Table 4), was also subcloned into a pET28a+ vector between *Bam*HI and *Eco*RI restriction sites resulting in a construct with an N-terminal His tag. Amplified sequences were verified by DNA sequencing.

Generation of EXOSC3 disease causing mutations

Mutations to plasmid pGEX-6-P1-EXOSC3 utilized in this study were introduced by QuikChange II XL mutagenesis kit (Cat# 200521, Agilent, Santa Clara, CA) in the pGEX-6-P1-EXOSC3 plasmid, using the primers listed in Supplementary Table 5. Plasmids were purified from DH5 α E. coli using the NucleoBond® Xtra Maxi kit (Cat# 740414, Macherey-Nagel, Germany). Amplified sequences and introduced mutations were verified by DNA sequencing.

Purification of recombinant GST-EXOSC2 and GST-EXOSC3

Following a 19 h long induction at 19 °C with 0.2 mM IPTG, bacterial cells expressing recombinant GST-tagged protein were resuspended in 20 mM Tris-HCl pH 8, 500 mM NaCl, 20% sucrose, 0.1% NP-40, 1 mM DTT supplemented with Complete protease inhibitors (Roche, Basel, Switzerland). Disruption of the bacteria was performed by two rounds of high-pressure homogenization at 10,000 PSI with a LM10 microfluidizer (Microfluidics, Westwood, MA), and the lysate was centrifuged 45 min at 4°C. The supernatant was loaded on a GST-Trap column (GE Healthcare, Uppsala, Sweden) equilibrated with 20 mM Tris pH 8, 350 mM NaCl, 20 % sucrose, 1 mM DTT. After a washing step with the same buffer, proteins were eluted with a gradient of glutathione. The fractions of interest were loaded on a HiLoad Superdex size exclusion column (GE Healthcare, Uppsala, Sweden) and eluted with 20mM Tris pH8, 350 mM NaCl, 20 % sucrose, 1 mM DTT. The eluted protein was concentrated with Amicon Ultra 15 centrifugal filters (Regenerated cellulose 10,000 NMWL; Merck Millipore, Darmstadt, Germany) to 5-7mg/mL, flash frozen on dry ice and stored at -80°C until use. Protein concentration was determined by a Pierce assay using bovine serum albumin as a standard. The purity of the proteins was verified with SDS-PAGE.

Purification of recombinant EXOSC1-His

Following a 19 h induction at 16°C with 0.5 mM IPTG, bacterial cells expressing recombinant EXOSC1-His were resuspended in 20 mM Tris-HCl pH 8, 500 mM NaCl, 20% sucrose, 0.1% NP-40, 1mM DTT supplemented with Complete protease inhibitors (Roche, Basel, Switzerland). Disruption of the bacteria was performed as described previously. The supernatant was loaded on a His-Trap (GE Healthcare, Uppsala, Sweden) equilibrated with 20 mM Tris-HCl pH 8, 300 mM NaCl, 20% sucrose, 1 mM DTT, 30 mM imidazole. After a washing step with the same buffer, proteins were eluted with a gradient of imidazole up to 400mM. The fractions of interest were loaded on a HiLoad Superdex size exclusion column (GE Healthcare, Uppsala,

Sweden) and eluted with 20 mM Tris-HCl pH 8, 300 mM NaCl, 20% Sucrose, 1 mM DTT. The eluted protein was concentrated with Amicon Ultra 15 centrifugal filters (Regenerated cellulose 10,000 NMWL; Merck Millipore, Darmstadt, Germany) and flash frozen on dry ice and stored at -80°C until use. Protein concentration was determined by a Pierce assay using bovine serum albumin as a standard. The purity of the protein was verified with SDS-PAGE.

Purification of GST-EXOSC3 mutants

GST-EXOSC3-V80F, Y109N, D132A and G135E mutants were purified using the same protocol as the wild-type protein. Except Y109N, all mutants purified with similar yields, purity and no aggregation was observed at any step (Fig S6). We were able to concentrate those proteins at ~5mg/mL. Despite our best efforts, we were not able to concentrate GST-EXOSC3-Y109N to 5mg/mL, the highest concentration we were able to get is 0.9mg/mL.

Differential scanning fluorimetry

Differential scanning fluorimetry was used to assess thermal stability of wild-type GST-EXOSC3 and disease-causing mutants. Experiments was performed as described in ¹. Briefly, 5 μM of GST-EXOSC3 or GST-EXOSC3 mutant protein was added to 5 μL of 100-times diluted SYPRO Orange stock (5000×) (Sigma-Aldrich, Stockholm, Sweden) in a 20mM MOPS pH7.2 buffer. The melting points of the proteins were determined using GraphPad Prism (GraphPad software, CA, USA). Differential scanning fluorimetry experiments revealed that while EXOSC3-D132A and EXOSC3-G135E were similar to EXOSC3-WT in their melting temperatures, Y109N and V80F induced ~1-2°C shift (vs. wildtype) in the thermal melting point of EXOSC3, which denotes local structural modifications (Supplementary Fig. S2C, S2D).

Synthesis and Blotting of SPOT Membranes

Peptides of 12 amino acids in length were spotted on SPOTs blots ² for EXOSC1, EXOSC5 and EXOSC9 proteins. Peptides were synthesized using standard 9-fluorenylmethoxycarbonyl (Fmoc) chemistry, in 30 × 20 spot arrays using a Multiprep peptide synthesizer adapted for SPOT synthesis (Intavis AG, Cologne, Germany). Membranes were blocked for at least 1 h in Tris-buffered saline containing 0.5% Tween 20 (TBST) with 5% skim milk powder before an overnight incubation with 1μM of GST-EXOSC3 or mutants at 4°C with gentle shaking. Following a series of washes in TBST, the blot was probed for an hour with a GST antibody at 4 °C. The following day, blots were washed three times for 10 min each time in TBST, incubated in secondary antibody (IgG (H+L) Cross Adsorbed Secondary Antibody, DyLight 800 (ThermoFisher, Product # SA5-10176) at dilutions 1:5,000) for 45 min at room temperature, and washed in TBST three more times for 10 min each time before visualizing SPOTs by exposing the membranes.

RNAcompete

The RNA pool generation, RNAcompete pulldown assays, and microarray hybridizations were performed as previously described³⁻⁵. Briefly, RNAcompete experiments employed defined RNA pools that are generated from 244K Agilent custom DNA microarrays. The final RNA pool consists of 241,399 individual sequences 30 to 41 nucleotides in length⁵. The microarray design is detailed elsewhere ⁵ and can be ordered from Agilent Technologies using AMADID# 024519. In RNAcompete assays, 20 pmoles of GST-tagged constructs and RNA pool (1.5 nmoles) are incubated in 1 mL of Binding Buffer (20 mM Hepes pH 7.8, 80 mM KCl, 20 mM NaCl, 10% glycerol, 2 mM DTT, 0.1 μg/μL BSA) containing 20 μL glutathione Sepharose 4B (GE

Healthcare) beads (pre-washed 3 times in Binding Buffer) for 30 minutes at 4°C, and subsequently washed four times for two minutes with Binding Buffer at 4°C. Z-scores were calculated for the motifs as described previously⁵.

Surface Plasmon Resonance

We performed Surface Plasmon Resonance (SPR) with a Biacore 3000 instrument (GE Healthcare, UK). EXOSC3 was covalently immobilized on a CM5 chip using standard amine coupling according to the manufacturer's protocol (GE Healthcare, UK). For comparative analysis, relative binding was normalized based on how much protein was immobilized (RU was divided by total protein immobilized). Once the protein was immobilized (ranging from 1000-3000 RU), we flowed the following sequences: Long G-rich: GGGAGGGAGGGAGAGGGA, Short G-rich: GGGAGGG, Random: AGUCAUUC (Eurofins Genomics, Louisville, KY) with HBS-EP buffer. Binding of RNA onto the immobilized proteins was tested at varying concentrations.

Covalent labelling of GST-EXOSC3 and mutants

Purified GST-EXOSC3 and mutants were fluorescently labelled using the Monolith Protein Labeling Kit RED-NHS (Nanotemper, Germany) according to the manufacturer's instructions. Briefly, GST-EXOSC3 was diluted to 20 µM in labeling buffer and mixed with one molar equivalent of the fluorescent NT-647-NHS dye and incubated for 30min at room temperature in the dark. Unreacted free dye was then eliminated using a desalting column, provided with the kit. Limited aggregation was observed for each protein when covalently labeled, as measured by microscale thermophoresis. Each protein was centrifuged 5min at 10,000xg, 4°C to remove aggregates. Experiments were performed directly after labelling.

Microscale Thermophoresis

Microscale thermophoresis experiments were performed similarly to ⁶. Briefly, 50nM of labeled protein was mixed with ranging concentration of long G-rich RNA (GGGAGGGAGGGAGAGGGA) in a 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 10 mM MgCl₂, 0.05 % Tween-20 buffer. The thermographs were recorded using MST premium capillaries at 20% LED and 40% MST power. Data analysis was performed with the MO Affinity Analysis software (Nanotemper) using the Kd model (standard fitting model derived from law of mass action).

In silico docking using Schrodinger

Molecular docking studies were performed using Schrodinger suite of programs including Glide docking. X-ray structure of the human RNA exosome (PDB code: 2nn6) was used for virtual screening of small molecule libraries. Initially docking grids were generated using RNA-EXOSC3 and EXOSC1 –EXOSC3 interface residues as active sites. We focused our docking on a 10Å³ pocket on the KH (which included Gln244, Asn251, Ile252, Leu271 and Ser274) and S1 (which included Phe146, Glu147, Arg152, Asn188 and Met190) residues domains that accommodates the residue within the RNA-binding motifs. Virtual screening was performed using stepwise virtual screening protocol in Glide program. DIVERSet-CL library, a small molecule library of 50,000 compounds from Chembridge Inc., was used for as an input for virtual screening. Resulting docking poses were analyzed using XP GScore and the top 20 compounds were selected for further screening. The resulting complexes were ranked using Glide score and other energy related terms⁷.

Saturation Transfer Difference Nuclear Magnetic Resonance (STD-NMR) Spectroscopy for small molecule binding

1D ¹H saturation transfer difference nuclear magnetic resonance (STD NMR) spectra with a spectral width of 12 ppm were collected for samples containing 500 μM of compound and 5 μM proteins in phosphate buffer (1:100 dilution was always maintained) as we have previously done for small molecule binding³⁹. STD NMR spectra were collected with a spectral width of 12 ppm, 16 K data points, and 3 second repetition delay. A saturation of the protein was achieved by a 2 second train of selective 50 ms Gaussian pulses centered at 0.74 ppm (on-resonance) and 30 ppm (off resonance). A 20-ms spin-lock was used to suppress the protein signal, followed by the double PFG spin echo to remove residual water signal. We acquired 512 scans per experiment. The on-resonance and off-resonance spectra were acquired interleaved, and the difference spectrum was acquired by phase cycling. Spectra processing and analysis were performed with the VNMRJ 3.2 (Agilent Technologies, Santa Clara, CA) and MestReNova 7.1 (Mestrelab Research, S.L., Santiago de Compostela, Spain).

ELISA-based determination of EXOSC3 binding to RNA

96-well plates (Nunc MaxiSorp, Thermo Scientific) were coated with anti-GST antibody (10 ng per well, Sigma) and incubated at room temperature overnight. The next day, the plates were washed and blocked with 3% BSA to minimize nonspecific adsorptive binding to the plates. 300 nM of GST-EXOSC3 was then added to the plates. As a negative control, some wells received 300 nM GST. The plates were incubated at room temperature with shaking for 1 hour. The plates were then washed with PBS containing 0.5% Tween-20. Escalating concentration of biotinylated RNA was added to the plate. The plates were then washed with PBS containing 0.5% Tween-20. The bound RNA was detected by streptavidin-HRP (150 ng/mL, Sigma). Tetramethylbenzidine (R&D Systems) was used as the colorimetric substrate. The optical density of each well was determined immediately, using a microplate reader (Multiskan Ascent, Thermo) set to 450 nm with a correction wavelength of 570 nm. Data were analyzed by nonlinear regression analysis using GraphPad Prism 5 (GraphPad, San Diego, CA).

Zebrafish morpholino injection and incubation with compounds

Zebrafish embryos were obtained from University Animal Care, and procedures were approved by the University of Arizona IACUC. Adult fish were maintained at 28 °C using a 14-h light/10-h dark cycle and bred overnight to obtain embryos. Morpholino oligonucleotides (Gene Tools) were designed to block translation initiation zebrafish exosc3 pre-mRNA as described⁸. Glass capillaries with a 1 mm outer diameter were pulled for use on a micro injector to perform injections of embryos between the one and four cell stage. The injection volume ranged from 0.5 to 1.0 nl at a concentration of 3 ng/nl. Embryos were incubated in E2 medium (1.5M NaCl, 50mM KCl, 100mM MgSO₄, 15mM KH₂PO₄, 5mM Na₂HPO₄) at 28 °C. On 3 dpf, water was changed to E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄ and 0.00001% methylene blue; Sigma) for experimental acclimation and general use.

Zebrafish embryos at the one cell stage were incubated with 50 μM compounds in E2 medium (1.5 M NaCl, 50 mM KCl, 100 mM MgSO₄, 15 mM KH₂PO₄, 5 mM Na₂HPO₄) for 5 days. Phenotype of the 5 dpf embryos were observed and compared to a vehicle only control.

Photobleaching measurement

50 nM of labeled EXOSC3 was mixed with increasing concentrations of ERD03 in 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 10 mM MgCl₂, 0.05 % Tween-20 buffer and incubated 10 min at room temperature. The photobleaching measurements were performed on a Monolith NT.115 (Nanotemper, Germany) using MST premium capillaries, at 40% LED and 40% MST power. Measurements were performed in triplicates. Data analysis was performed with the MO Affinity Analysis software (Nanotemper) using the Kd model (standard fitting model derived from law of mass action).

RNAseq analysis

Each isolated RNA sample was DNase treated with TURBO DNA-free (Thermo Fisher), then purified and concentrated with RNeasy MinElute Cleanup Kit (Qiagen). The RNA was measured for quantity with Quant-iT Ribogreen RNA Assay (Thermo Fisher,) and quality with Agilent High Sensitivity RNA Screen Tape and buffer (Agilent).

For each RNA sample, indexed, Illumina-compatible, double-stranded cDNA libraries were synthesized from 10ng total RNA with Takara Bio's SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian kit (Takara Bio). Library preparation included chemical RNA fragmentation (94 °C for 4 min), a 5-cycle Indexing PCR, ribosomal cDNA depletion, and a 12-cycle enrichment PCR. Each library was measured for size with Agilent's High Sensitivity D1000 Screen Tape and reagents (Agilent) and measured for concentration with KAPA SYBR FAST Universal qPCR Kit (Kapa Biosystems). Libraries were then combined into an equimolar pool, and the pool was measured for size and concentration. The pool was clustered onto a paired-end flow cell (Illumina, Cat. No. PE-401-3001) with a 1% v/v PhiX Control v3 spike-in (Illumina) and sequenced on Illumina's HiSeq 2500 with TruSeq v3 chemistry (Illumina). The first and second reads were each 83 bases.

RNA-sequencing reads were mapped, transcripts assembled, and lncRNAs identified using the Evolinc workflow⁹ with HiSat2¹⁰ and Stringtie¹¹ options. Transcripts with less than 3 reads per base were discarded. Final transcripts were merged into one genome annotation file for downstream analyses. To determine differential expression, reads per locus were first quantified using HTseq¹² in CyVerse's Discovery Environment¹³ and then tested for differential expression using DESeq2¹⁴ in R. RNAseq libraries are available on NCBI's sequence read archive (<https://www.ncbi.nlm.nih.gov/sra>) under the bioproject ID PRJNA470927.

Quantitative PCR

RNA from 1dpf zebrafish (n=50) was extracted using the mirVANA kit (Ambion) following manufacturer instructions and then treated with DNA-free Kit (Ambion) in order to remove genomic DNA contamination. 1 µg of RNA was used to synthesize cDNA with the Maxima Reverse transcriptase (ThermoFisher). qPCR (Biorad light-cycler equipped with a MyIQ detection system) was performed in triplicates. qPCR primers described in Supplementary Table 6 were the same as was previously described¹⁵.

Zebrafish cerebellum size determination

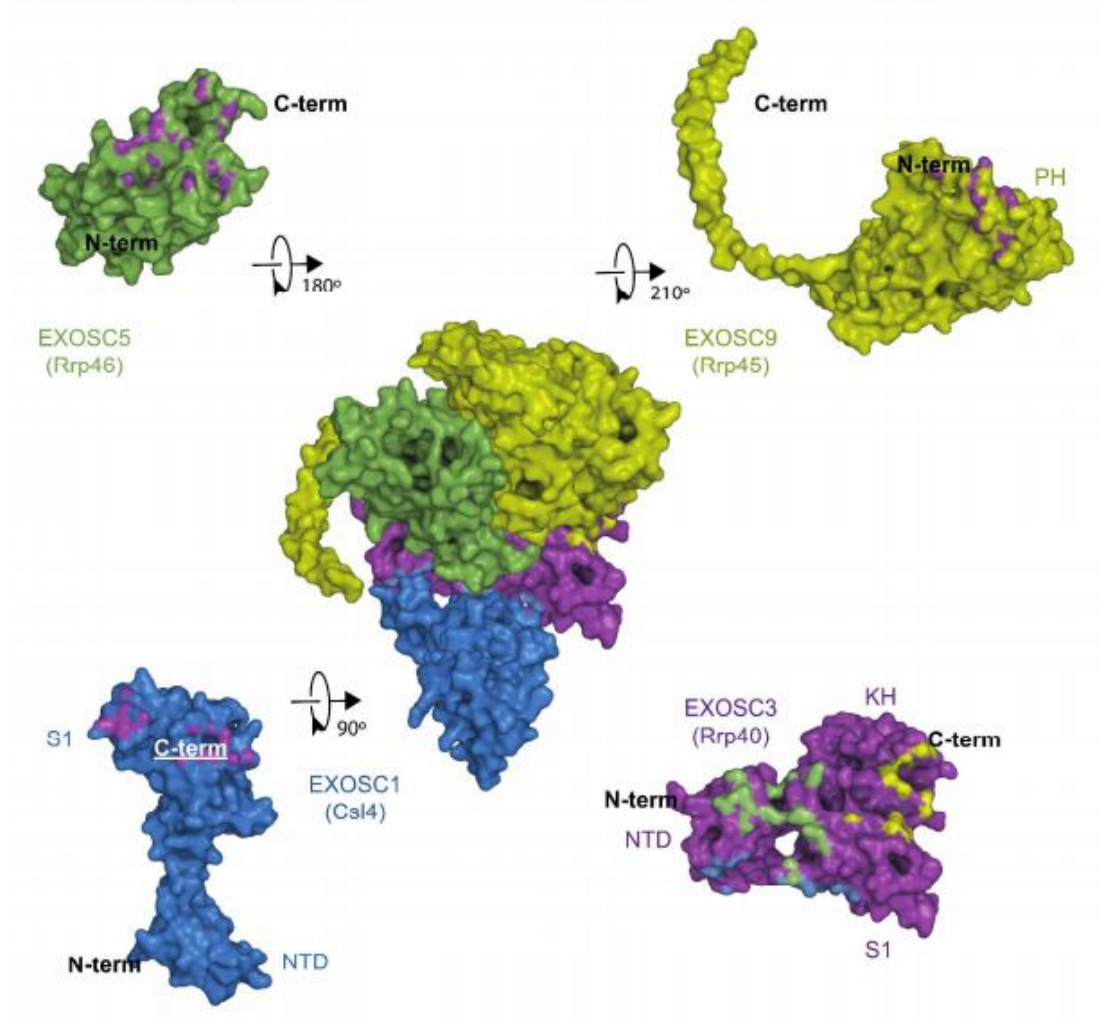
All zebrafish were maintained in accordance with standard laboratory conditions. The University of California, Los Angeles Chancellor's Animal Research Committee approved all experiments performed on zebrafish (ARC #2005-141-32). The TDL6 (Tübingen driver line) zebrafish used in all experiments were identified in a screen for developmentally regulated enhancers that drive tissue-specific expression. Gal4-driven green-fluorescent protein (GFP) expression marks

neurons of the central nervous system (CNS). The Gal4-GFP cassette location was mapped to the first intron of the *gna12* gene on chromosome 6 and is regulated by either the *gna12* or *inka1b* enhancer.

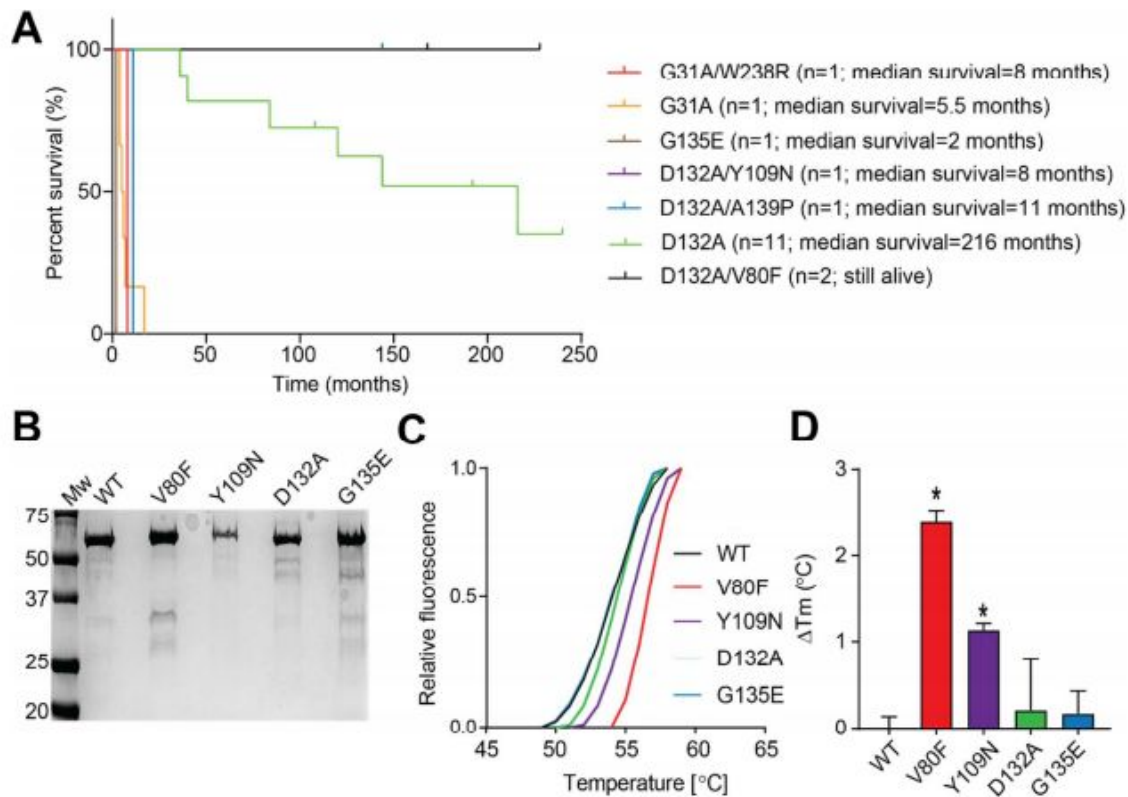
Zebrafish embryos were incubated in E3 buffer (5.0 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄), E3 buffer with 0.5% v/v DMSO, ERD03 (50 μM, 0.5% v/v DMSO in E3 buffer), immediately after fertilization in 50 mm diameter Pyrex dishes at 28.5°C for 3 days. The zebrafish were then fixed for 10 minutes in 4% v/v paraformaldehyde, rinsed 3X with phosphate-buffered saline, mounted in 1% w/v low-melt agarose on 35 mm glass bottom dishes (MatTek Corporation, Ashland, MA), and imaged using a Zeiss Axio Observer Z1 spinning disk confocal microscope (Carl Zeiss Microscopy, Thornwood, NY). Embryos were viewed under a 10X objective lens and captured with SlideBook 6.0 software (Intelligent Imaging Innovations, Denver, CO). All images were then exported to ImageJ Fiji Edition³¹⁶ Version 2.0 (continuous release version; National Institutes of Health, Bethesda, MD) and analyzed. Images were de-identified and analyzed blindly by an investigator different from the investigator who acquired the images.

In order to determine the relative size of the zebrafish embryo cerebellum, the fluorescent image stacks were opened in ImageJ. The images were calibrated to ensure that the area (in μm²) was being calculated instead of the number of pixels. The region containing the cerebellum was cropped from the rest of the image stack and then an intensity threshold was applied to select the 2D cross section of the cerebellum on every image. From this, the average area was calculated per fish and statistical analysis was performed to determine any significant differences in size between treatment groups.

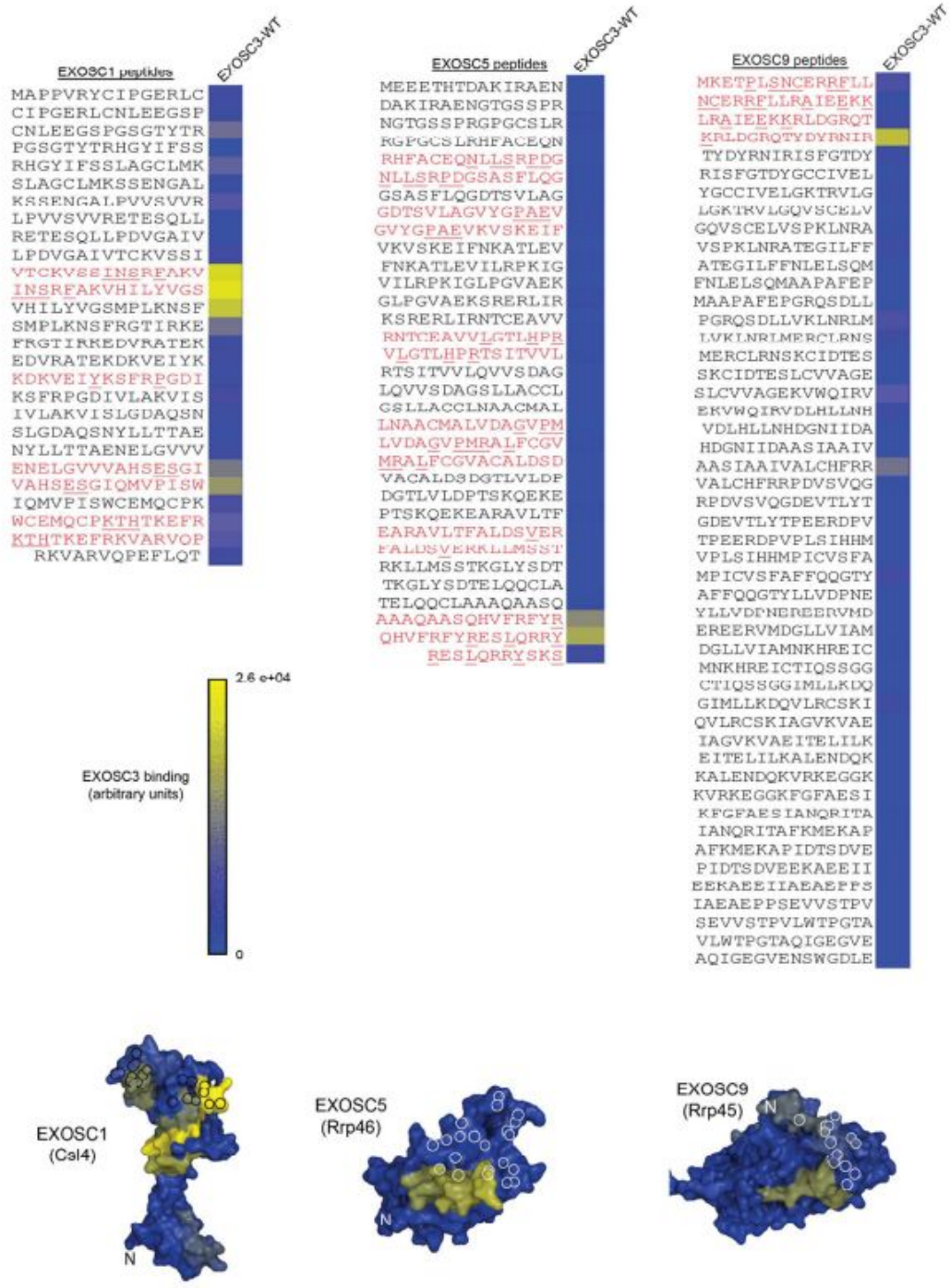
Supplementary Figures



Supplementary Figure S1: Interaction of EXOSC3 with EXOSC1, EXOSC5 and EXOSC9. Surface representation of EXOSC3 (purple) and the subunits in contact with EXOSC3: EXOSC1 (blue), EXOSC5 (green) and EXOSC9 (yellow) (extracted from PDB code: 2nn6). Points of contact between EXOSC3 and EXOSC1, EXOSC5 and EXOSC9 have been mapped on every protein surface using the color coding described previously. For each protein the N-terminus (N-term) and the C-terminus (C-term) are shown. The C-terminus of EXOSC1 (underlined, in white) is located on the back of the protein. NTD: N-terminus Domain, PH: RNase Pleckstrin homology domain, S1: ribosomal protein S1 domain and KH: ribonucleoprotein K homology domain.

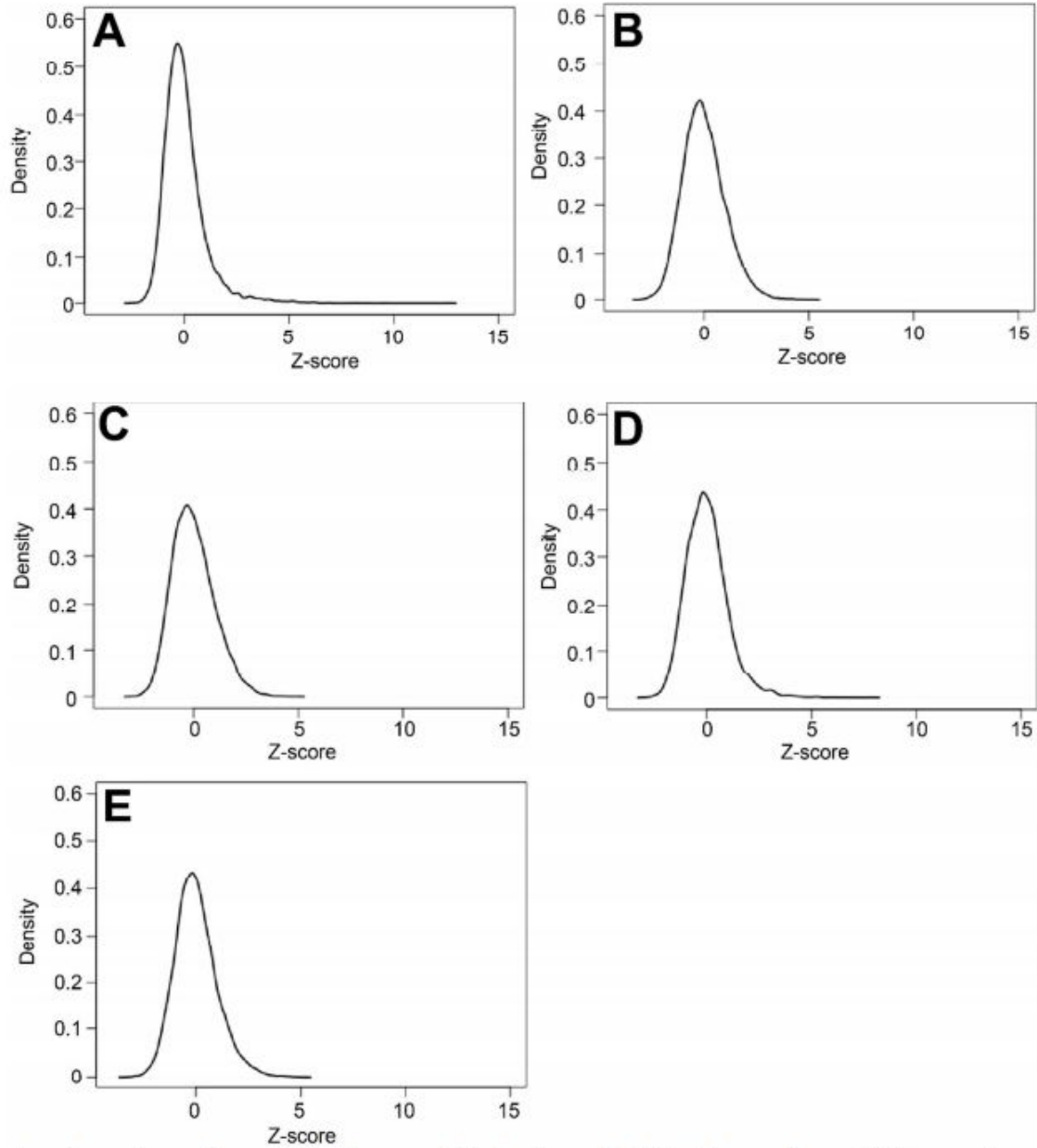


Supplementary Figure S2: Stability of disease causing EXOSC3 mutants. **A.** Survival of Pontocerebellar Hypoplasia 1B patients based on EXOSC3 mutations. Data extracted from ^{10,12-14} was used to assess PCH1B patients' survival. Kaplan-Meier survival curves is shown for each homozygous and heterozygous mutations of EXOSC3, along with the number of patients and the median survival. Patients exhibiting the heterozygous D132A/V80F mutations (black), and three patients with the homozygous D132A mutation (green) were still alive at the time of publication. **B.** SDS PAGE gel using 5 μ g of purified GST-EXOSC3 along with EXOSC3 mutants V80F, Y109N, D132A and G135E. **C.** Representative normalized fluorescence thermal melting curves of GST-EXOSC3-WT (black), GST-EXOSC3-V80F (red), GST-EXOSC3-Y109N (dash line), GST-EXOSC3-D132A (blue) and GST-EXOSC3-G135E (green). **D.** Bar graph illustrating the average ΔT_m for GST-EXOSC3-WT and disease-causing mutants. Asterisks indicate statistical significance ($p < 0.05$, Kruskal-Wallis non-parametric test with a Dunnett's post hoc analysis, $n=6$).

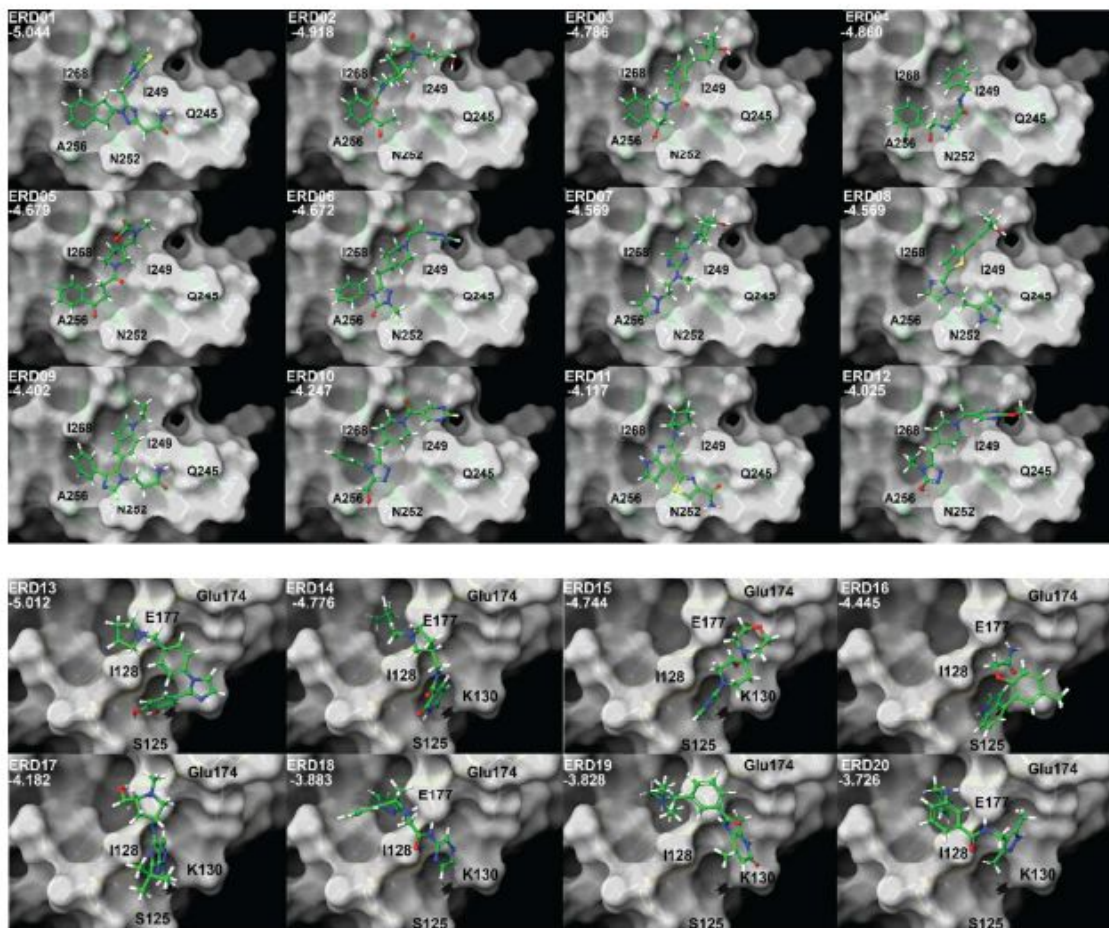


Supplementary Figure S3: EXOSC3 binding to adjacent subunits through peptide array. Western assay of EXOSC3-WT on immobilized 12-mer EXOSC1 (A), EXOSC5 (B) and EXOSC9 (C) peptides, in overlapping 8 amino acids steps. The blots were scanned and spot intensities were

quantified and raw data was represented as a matrix generated with the Matrix2png software⁴⁹. Each row represents the peptides (12-mer) tiled with overlap to cover the entire protein sequences.



Supplementary Figure S4: Z-score distribution of RNA 7-mers from RNAcompete on EXOSC3 proteins. The graph represents density plots for RNAcompete-derived 7mer Z-scores for EXOSC3-WT **A**, EXOSC3-V80F **B**, Y109N **C**, D132A **D**, G135E **E**.

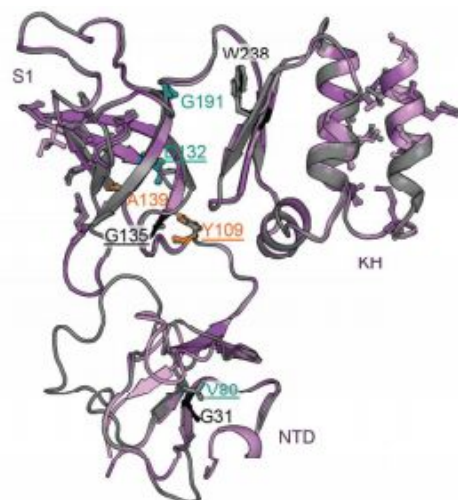


Supplementary Figure S5: Top hits from *in silico* docking poses for each domain. EXOSC3 RNA Disruptors (ERD) 1 to 20 docking poses on 2nn6 along with the corresponding XP Glide docking score, which is a semi quantitatively ranking of candidate ligands to bind to a specified conformation of the protein, are shown. ERD 01-12 compounds target the KH domain, while ERD13-20 target the S1 domain.

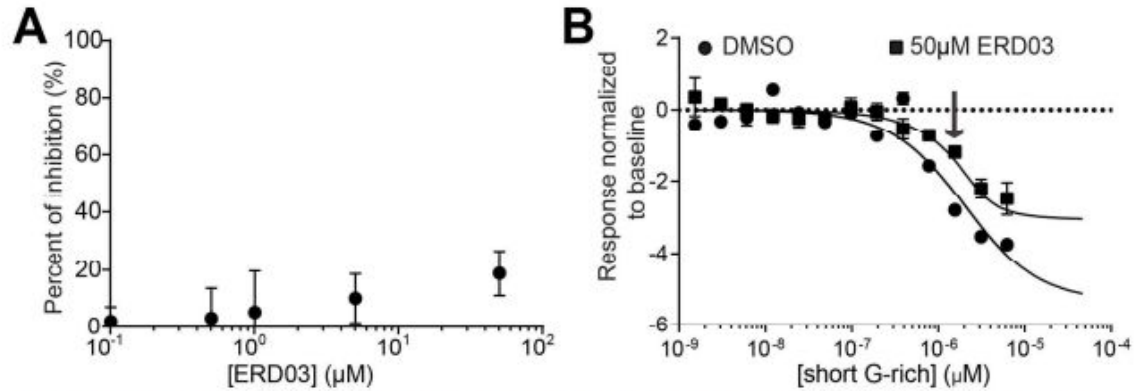
Human	1	MAEPASVAAESLAGSRARAARTVLGQVVLPG EE LLL---PEQEDAEGPGG	47
Zebrafish	1	-----MDSSVHTSLLERIGDVVLP GD LLFSFSPPEAGDA-----	34
Human	48	AVER PLSLNARACSRVRVVC GP GLRRC GD RL L VTKCGRLRHKEPGSGSGG	97
Zebrafish	35	-----NPKA---DRLIC GP GLRRSGAEIR V CRAGVLKHKQP-----	67
Human	98	CVY WVDS Q Q KR Y VPVKCDHVICIVTAK SGD IF KV D V GC SEPA S LSYLSPE	147
Zebrafish	68	NMYWVNCQ QR R Y VPK GES VIGIVTAK SGD V FK V D V GC SEQASLSYL A PE	117
Human	148	GATKRRNPVQ VD LIY Q PV V ANKDM EP EMV C IDSCGRANGM GV IGQDG	197
Zebrafish	118	GATKRRNPVQ VD LV F Q P TIANKDM EP ELV C IDSCGRANGM GV FGDG	167
Human	198	LLFKVTLGLIRKLLAPDCE II Q EV GKLYPLEIVFGMNGRI V WKARTIQQT	247
Zebrafish	160	LLFKVSLGLVRRLLAQSDLVSDLE RM PFCEM V GMNGR V WKARTVQHT	217
Human	248	IL LN IL E ACE MTSD QR K Q IF S RL AE S--	275
Zebrafish	218	IL SN LL E ACE N M STA QR HT L FR R V AD GGI	247

S1 Pocket

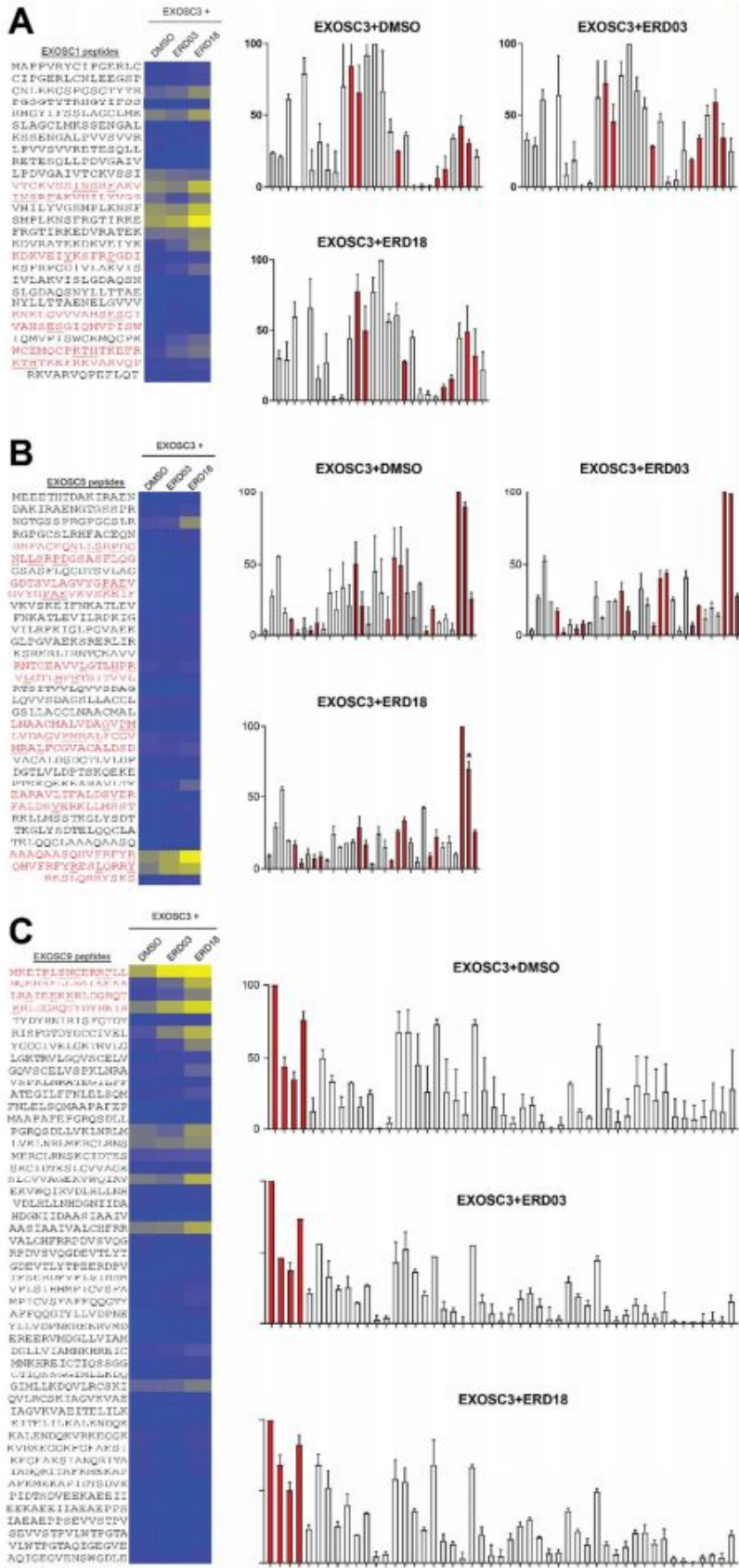
KH Pocket



Supplementary Figure S6: Comparison of human structures of EXOSC3 with homology model of Zebrafish EXOSC3. **A.** Sequence alignment for human EXOSC3 (Uniprot entry: Q9NQT5) and zebrafish EXOSC3 (Uniprot entry: Q4KMD2). Residues in the S1 and KH pockets are boxed with amino acids that potentially contact the small molecules bolded and underlined. PCH1B-associated residues are shown in bolded italics. Differences between human and zebrafish EXOSC3 are shown in magenta. **B.** The structure of human EXOSC3 protein (extracted from PDB: 2nn6) shown as ribbon diagram. Differences in the primary sequence are shown in magenta. A homology model of the zebrafish EXOSC3 generated from human EXOSC3 is also shown in grey. The S1 and KH small-molecule binding sites and the PCH1B-associated residues are shown as sticks. The zebrafish EXOSC3 was modeled in SWISS-MODEL using the top search hit, human EXOSC3 (PDB code: 2nn6, chain G, 67% identical and GMQE score of 0.73) using default parameters⁵⁴.

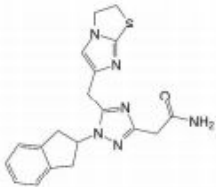
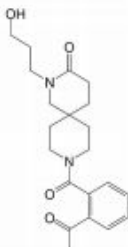
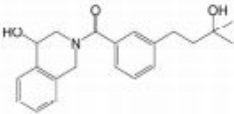
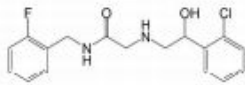
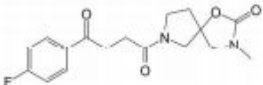
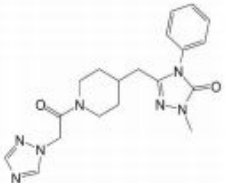
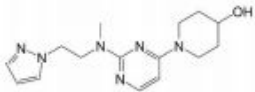


Supplementary Figure S7: ERD03 disrupts EXOSC3-RNA interaction in a dose dependent fashion. **A.** Concentration-dependent curves were obtained for ERD03's disruption of RNA-EXOSC3 interaction using ELISA. Measurements were performed in sextuplicate and are represented as mean \pm SD. **B.** MST values of EXOSC3 with 0.012-6.25µM of short G-rich RNA and 1% (v/v) DMSO or 50µM ERD03. Arrow denotes 57 \pm 3% inhibition of EXOSC3-RNA by ERD03. Some error bars are smaller than the symbols.



Supplementary Figure S8: Effect of ERD03 and ERD18 on EXOSC3 binding to adjacent subunits assessed by peptide array. Western blot assay of EXOSC3-WT incubated with 50 μ M of either ERD03 or ERD18 or DMSO (2.5%) on immobilized 12-mer EXOSC1 (A), EXOSC5 (B) and EXOSC9 (C) peptides with 8 amino acids overlap. The blots were scanned and spot intensities were quantified and raw data was represented as a matrix generated with the Matrix2png software⁵¹. Each row represents the peptides (12-mer) tiled with overlap to cover the entire protein sequences. Each column represents EXOSC3-WT and different mutations used to test binding to proteins (tiled on each row). On the right, the data was represented as normalized to the maximum. Predicted contacts with EXOSC3 are shown in red. Data: means \pm SD (n=2)

Supplementary Table 2

Compound	Structure	Chemical Name	Molecular weight
ERD01		2-(1-(2,3-dihydro-1H-inden-2-yl)-5-((2,3-dihydroimidazo[2,1-b]thiazol-6-yl)methyl)-1H-1,2,4-triazol-3-yl)acetamide	378.5
ERD02		9-(2-acetylbenzoyl)-2-(3-hydroxypropyl)-2,9-diazaspiro[5.5]undecan-3-one	372.5
ERD03		(4-hydroxy-3,4-dihydroisoquinolin-2(1H)-yl)(3-(3-hydroxy-3-methylbutyl)phenyl)methanone	339.4
ERD04		2-((2-(2-chlorophenyl)-2-hydroxyethyl)amino)-N-(2-fluorophenyl)acetamide	336.8
ERD05		1-(4-fluorophenyl)-4-(3-methyl-2-oxo-1-oxa-3,7-diazaspiro[4.4]nonan-7-yl)butane-1,4-dione	334.3
ERD06		5-((1-(2-(1H-1,2,4-triazol-1-yl)acetyl)piperidin-4-yl)methyl)-2-methyl-4-phenyl-2,4-dihydro-3H-1,2,4-triazol-3-one	381.4
ERD07		1-(2-((2-(1H-pyrazol-1-yl)ethyl)(methyl)amino)pyrimidin-4-yl)piperidin-4-ol	302.4

Compound	Structure	Chemical Name	Molecular weight
ERD08		4-(5-(1-(2-(1H-imidazol-4-yl)ethyl)-1H-imidazol-2-yl)thiophen-2-yl)-2-methylbut-3-yn-2-ol	326.4
ERD09		3-(5-(1-ethyl-1H-indol-5-yl)-4-phenyl-1H-imidazol-1-yl)propanamide	358.4
ERD10		4-((5-(hydroxymethyl)-4-phenyl-4H-1,2,4-triazol-3-yl)methyl)piperidin-1-yl(pyrimidin-5-yl)methanone	378.4
ERD11		2-(6-(cyclopentylamino)-1H-pyrrolo[2,3-b]pyridin-4-yl)thiazole-4-carboxamide	327.4
ERD12		4-ethyl-5-((1-((2-methoxypyrimidin-5-yl)methyl)piperidin-4-yl)methyl)-2,4-dihydro-3H-1,2,4-triazol-3-one	332.4
ERD13		4-(1-(4-(pyrrolidin-1-ylmethyl)phenyl)-1H-imidazol-2-yl)benzene-1,2-diol	335.4
ERD14		N-((1-((5-ethylpyridin-2-yl)methyl)pyrrolidin-3-yl)methyl)-2-hydroxynicotinamide	340.4

Compound	Structure	Chemical Name	Molecular weight
ERD15		(3R,4R) 4 morpholino 1 (9H purin 6 yl)piperidin 3 ol	304.3
ERD16		2-(2-(1H-indol-4-yl)-4-methylphenoxy)acetamide	280.3
ERD17		((2S,4S)-4-((6-(tert-butyl)-1-methyl-1H-pyrazolo[3,4-d]pyrimidin-4-yl)methyl)-1-methylpyrrolidin-2-yl)methanol	318.4
CRD18		2-((1-(2-(1H-pyrazol-1-yl)phenyl)ethyl)amino)-N-(pyrazin-2-yl)acetamide	322.4
ERD19		N-(2,4-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-3-((4-(2-hydroxyethyl)piperazin-1-yl)methyl)benzamide	384.5
ERD20		N-((3,5-dimethyl-1H-pyrazol-4-yl)methyl)-3-((thiophen-2-ylmethyl)amino)benzamide	340.4

Supplementary Table 4: Primers used to clone the cap proteins.

Protein	Primer Fw	Primer Rv	Plasmid	DNAsu accession number
EXOSC2	5'- <u>tccaggggccctggatcctccg</u> cgatggagatgagg-3'	5'- tcacgatgcggccgctc <u>gagctattatcc</u> ctctgtccaaa-3'	pGEX-6-P1	HsCD00 084012
EXOSC3	5'- <u>tccaggggccctggatcctccg</u> ccgaacctgcgtct-3'	5'- tcacgatgcggccgctc <u>gagctatcaac</u> tttctgccaatctgg-3'	pGEX-6-P1	HsCD00 0429597
EXOSC1	5'- cagcaaatgggtcgcggatccga aaacctgtactccagtccgcgcc acctgtgagata-3'	5'- ttgtcgacggagctcga <u>attcctattaggt</u> ctgcaagaattcgggtgtactcgggc- 3'	pET28a(+)	HsCD00 084353

EXOSC2 and EXOSC3 sequences were amplified from pVP16 plasmid, purchased from DNAsu (Tempe, Arizona) using the primers described in the table. EXOSC2 and EXOSC3 were inserted into a pGEX-6P-1 plasmid using BamHI (GGATCC) and HindIII (CTCGAG) restriction sites (underlined in the primer sequences) generating constructs with an N-terminal GST tag. EXOSC1 was inserted into pET28a(+) plasmid using BamHI (GGATCC) and EcoRI (GAATTC) restrictions sites generating a construct with an N-terminal His tag.

Supplementary Table 5: Primers used to generate EXOSC3 mutants. Mutations V80F, Y109N, G135E and D132A were introduced in pGEX-6P-1- EXOSC3 using the primers described in the table.

Mutation	Primer Forward	Primer Reverse
V80F	5'-ccgcacttggtgaacagcagggcggcc-3'	5'-ggaccgcctgctgttcaccaagtgcgg-3'
Y109N	5'-tccttttactggaacattccgcttctgctgagagt-3'	5'-actctcagcagaagcggaatgtccagtaaaagga-3'
D132A	5'-ggctcactccctccaacagcaacttgaatatactcca-3'	5'-tggagatatattcaaagttgctgttggaggagtgaggacc-3'
G135E	5'-aagaagctggctcactctctccaacatcaacttg-3'	5'-caaagttgatgttggagagagtgagccagcttctt-3'

Supplementary Table 6: qPCR primers used to assess mRNA levels in zebrafish embryos.
Those primers are from¹⁵.

Gene	Primer Forward	Primer Reverse
atxn1a	5'-GGGTGGAAGACCTGAAAACA-3'	5'-GCCGAACACAAAGAAAGGAT-3'
atxn1b	5'-TACAGACATCGCCACAGAG-3'	5'-CAGCGGCACTCCTAATGCT-3'
β -actin	5'-CGAGCTGTCTTCCCATCCA-3'	5'-TCACCAACGTAGCTGTCTTTCTG-3'

Supplemental references:

- (1) Moutal, A., François-Moutal, L., Perez-Miller, S., Cottier, K., Chew, L. A., Yeon, S. K., Dai, J., Park, K. D., Khanna, M., and Khanna, R. (2016) (S)-Lacosamide Binding to Collapsin Response Mediator Protein 2 (CRMP2) Regulates CaV2.2 Activity by Subverting Its Phosphorylation by Cdk5. *Mol. Neurobiol.* 53, 1959–1976.
- (2) Frank, R. (2002) The SPOT-synthesis technique: Synthetic peptide arrays on membrane supports - Principles and applications. *J. Immunol. Methods* 267, 13–26.
- (3) Ray, D., Ha, K. C. H., Nie, K., Zheng, H., Hughes, T. R., and Morris, Q. D. (2017) RNAcompete methodology and application to determine sequence preferences of unconventional RNA-binding proteins. *Methods* 118–119, 3–15.
- (4) Ray, D., Kazan, H., Chan, E. T., Castillo, L. P., Chaudhry, S., Talukder, S., Blencowe, B. J., Morris, Q., and Hughes, T. R. (2009) Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nat. Biotechnol.* 27, 667–670.
- (5) Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L. H., Dale, R. K., Smith, S. A., Yarosh, C. A., Kelly, S. M., Nabet, B., Mecnas, D., Li, W., Laishram, R. S., Qiao, M., Lipshitz, H. D., Piano, F., Corbett, A. H., Carstens, R. P., Frey, B. J., Anderson, R. A., Lynch, K. W., Penalva, L. O. F., Lei, E. P., Fraser, A. G., Blencowe, B. J., Morris, Q. D., and Hughes, T. R. (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177.
- (6) Wilson, S. M., Moutal, A., Melemedjian, O. K., Wang, Y., Ju, W., François-Moutal, L., Khanna, M., and Khanna, R. (2014) The functionalized amino acid (S)-Lacosamide subverts CRMP2-mediated tubulin polymerization to prevent constitutive and activity-dependent increase in neurite outgrowth. *Front. Cell. Neurosci.* 8.
- (7) Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* 49, 6177–6196.
- (8) Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schöneborn, S., Menezes, M. P., Hong, J. E., Leong, D. W., Senderek, J., Salman, M. S., Chitayat, D., Seeman, P., Von Moers, A., Graul-Neumann, L., Kornberg, A. J., Castro-Gago, M., Sobrido, M. J., Sanefuji, M., Shieh, P. B., Salamon, N., Kim, R. C., Vinters, H. V., Chen, Z., Zerres, K., Ryan, M. M., Nelson, S. F., and Jen, J. C. (2012) Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.* 44, 704–708.
- (9) Nelson, A. D. L., Devisetty, U. K., Palos, K., Haug-Baltzell, A. K., Lyons, E., and Beilstein, M. A. (2017) Evolinc: A tool for the identification and evolutionary comparison of long intergenic non-coding RNAs. *Front. Genet.* 8.
- (10) Kim, D., Langmead, B., and Salzberg, S. L. (2015) HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360.

- (11) Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* *33*, 290–295.
- (12) Anders, S., Pyl, P. T., and Huber, W. (2015) HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* *31*, 166–169.
- (13) Merchant, N., Lyons, E., Goff, S., Vaughn, M., Ware, D., Micklos, D., and Antin, P. (2016) The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. *PLoS Biol.* *14*.
- (14) Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*.
- (15) Giunta, M., Edvardson, S., Xu, Y., Schuelke, M., Gomez-Duran, A., Boczonadi, V., Elpeleg, O., Müller, J. S., and Horvath, R. (2016) Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. *Hum. Mol. Genet.* *25*, 2985–2996.
- (16) Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: An open-source platform for biological-image analysis. *Nat. Methods*.
- (17) Pavlidis, P., and Stafford Noble, W. (2003) Matrix2png: A utility for visualizing matrix data. *Bioinformatics* *19*, 295–296.

Chapter 3:

A Small Molecule Candidate for Prevention or Treatment of Amyotrophic Lateral Sclerosis and Other TDP-43 Proteinopathies

Chapter 3 contains Shahriyar Jahanbakhsh's and Matthew R. Hennefarth's contributions to an unpublished manuscript. The other co-authors are Liberty François-Moutal, David D. Scott, Razaz Felemban, Victor Miranda, Melissa Sayegh, Samantha Perez-Miller, Vijay Gokhale, Daniela C. Zarnescu, and May Khanna, all at the University of Arizona, Tucson; these co-authors discovered the small molecule rTRD01 and performed the biophysical characterization of rTRD01.

Introduction

As I discussed in the opening chapter, TDP-43 is a DNA- and RNA-binding protein whose aggregation in cytosolic inclusions is a pathological hallmark of ALS. The architecture of TDP-43 consists of an N-terminal domain, two tandem RNA recognition motifs (RRMs), and a C-terminal glycine-rich, prion-related domain¹. The C-terminal region contains most of the disease-associated mutations observed among patients with familial forms of ALS² (**Figure 3-1**).

In addition to its nuclear functions in regulation of transcription and alternative splicing, TDP-43 is recruited to cytoplasmic RNA granules that are formed after exposure to oxidative, osmotic, temperature, viral, and other environmental stressors. These so-called stress granules (SGs) are membrane-less organelles believed to promote cell survival by storing non-essential mRNAs, translation factors and RNA-binding proteins during times of stress. SGs disassemble as the cell recovers. Depletion of TDP-43 reduces nucleation of SGs and greatly reduces cell survival under stress conditions³.

While SGs are a vital component of the cellular response to stress, perturbations to SG dynamics (assembly or disassembly) can potentially result in neurodegenerative disease. One popular hypothesis posits that mutations or cellular conditions that increase SG persistence, either due to excessive assembly or impaired clearance, increase the stochastic conversion of SG-localized TDP-43 into a toxic aggregate⁴. A 2019 report, in contrast, found that SG components are predominantly excluded from TDP-43 pathology in motor neurons from ALS patients. Rather, SGs were shown to sequester histone deacetylase 6, a facilitator of protein degradation, resulting in reduced TDP-43 clearance and enhanced TDP-43 aggregation⁵. In addition, the ability of TDP-43 to bind RNA is an inherent requirement for TDP-43 toxicity⁶.

Thus, decreasing the association of TDP-43 with RNA, be it in stress granules or elsewhere, under the oxidative cellular conditions of ALS⁷, is a plausible therapeutic approach.

Given the absence of effective therapies for ALS and other TDP-43 proteinopathies, our colleagues at May Khanna's laboratory at the University of Arizona sought to apply the chemical biology approach described in Chapter 2 to identify new small molecule candidates that inhibit association of TDP-43 with RNA. *In silico* docking yielded a molecule, rTRD01, that was selected for additional *in vitro* biophysical and *in vivo* biological analysis.

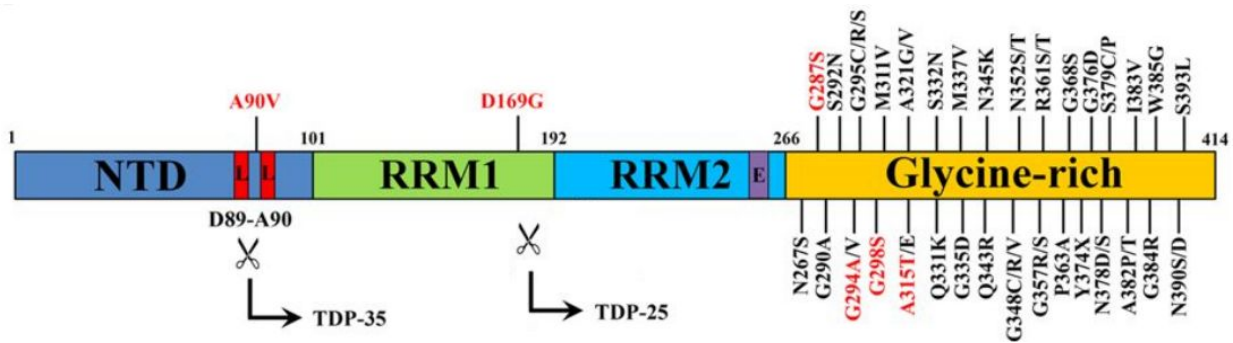


Figure 3-1². Domain architecture of TDP-43. TDP-43 consists of an NTD, two tandem RNA recognition motifs (RRM1 and RRM2), and a C-terminal glycine-rich region. Approximately 50 missense mutations have been identified in TDP-43 in sporadic and familial ALS patients, mostly in the glycine-rich region. Scissors denote caspase cleavage sites that generate C-terminal 35 kDa and 25 kDa fragments. L, nuclear localization signal. E, nuclear export signal. Adapted with permission from Springer Nature under a Creative Commons–Attribution 4.0 International license.

Paraquat as a model for ALS and Parkinson's disease

Paraquat (*N,N'*-dimethyl-4,4'-bipyridinium dichloride) is a dicationic herbicide that induces oxidative stress and neuronal death in zebrafish^{8,9} and other model organisms^{10,11}.

Paraquat is a structural analog of the acute neurotoxin 1-methyl-4-phenylpyridinium, and exposure to paraquat is associated with development of Parkinson's disease and ALS in humans^{12,13,14}. Paraquat also induces accumulation of TDP-43 in SGs in cell culture¹⁵.

Upon entering the cell, paraquat undergoes an NADPH-dependent, one-electron reduction to form its free radical, which then reacts immediately with molecular oxygen to reform the cation and produce superoxide anion, which in turn dismutates into hydrogen peroxide. The reduction of paraquat is mediated by the mitochondrial electron transport chain¹⁶. Except for its perpetual reduction and oxidation¹⁷, paraquat is otherwise not metabolized¹⁸.

rTRD01: a candidate small molecule for treatment or prophylaxis of TDP-43 proteinopathies

Using the chemical biology approach detailed in Chapter 2, our colleagues at the University of Arizona discovered a small molecule, rTRD01 (RRM TDP-43 RNA Disruptor 01; 6-[(3*R*)-3-[(4-fluorophenyl)methyl]-3-(hydroxymethyl)piperidin-1-yl]pyrazine-2-carboxamide), capable of binding TDP-43 in the RRM1 and RRM2 domains and partially disrupting TDP-43's interaction with (GGGGCC)₄ RNA, the hexanucleotide repeat expansion mutation of the c9orf72 gene that is the most common genetic cause of ALS and frontotemporal dementia¹⁹. My team at UCLA assayed rTRD01's neuroprotective potential *in vivo* in the zebrafish paraquat model.

Materials and Methods

Zebrafish culture

All zebrafish (*Danio rerio*) were maintained in accordance with standard laboratory conditions²⁰. The University of California, Los Angeles Chancellor's Animal Research

Committee approved all experiments performed on zebrafish (ARC #2005-141-32). The TDL6 (Tübingen driver line) zebrafish used in all experiments were identified in a screen for developmentally regulated enhancers that drive tissue-specific expression. Gal4-driven green fluorescent protein (GFP) expression marks neurons of the central nervous system (CNS). The Gal4-GFP cassette location maps to the first intron of the *gnal2* gene on chromosome 6, and is regulated by the *inka1b* enhancer²¹.

Treatment conditions

Zebrafish embryos ($N > 15$ per condition) were incubated in E3 buffer (5.0 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄) ± 0.075 percent v/v DMSO ± 0.04 ppm paraquat dichloride (Sigma) ± 50 μM rTRD01 ± 100 nM rotenone (Sigma) ± 0.003 percent w/v phenylthiourea (PTU, an inhibitor of pigment synthesis) (Sigma), as indicated in the text. All solutions with rTRD01 or paraquat also contain 0.075 percent v/v DMSO. Media was refreshed every 24 hrs. Zebrafish embryos used in **Figure 3-4** were enzymatically dechorionated by Pronase (Sigma) treatment at 1 hpf.

Imaging

All imaging was performed using live embryos anesthetized in 0.02 percent w/v tricaine (ethyl 3-aminobenzoate methanesulfonate) and mounted in 1 percent w/v low-melt agarose on 35 mm glass bottom dishes (MatTek Corporation, Ashland, MA). Brightfield images were captured on a Leica EZ4D stereo microscope (Leica Microsystems, Buffalo Grove, IL) using Leica LAS EZ 3.4 imaging software. Fluorescence micrographs were obtained using a Zeiss Axio Observer Z1 spinning disk confocal microscope (Carl Zeiss Microscopy, Thornwood, NY) under a 10X objective lens and captured with SlideBook 6.0 software (Intelligent Imaging Innovations,

Denver, CO). Micrographs were then de-identified and exported to ImageJ Fiji Edition²² Version 2.0 (continuous release version; National Institutes of Health, Bethesda, MD) for analysis. Images were analyzed blindly by an investigator different from the investigator who acquired the images.

Zebrafish brain size determination

In order to determine the relative size of the zebrafish embryo brain, the fluorescence image stacks were opened in ImageJ Fiji Edition. The images were calibrated to ensure that the area (in μm^2) was being calculated instead of the number of pixels. Thereafter, the background was subtracted using a 30 pixel rolling ball radius. The region containing the brain was then cropped from the rest of the image stack and then an intensity threshold was applied to select the 2D cross section of the brain on every image. From this, the average area, mean fluorescence, and integrated fluorescence density was calculated per fish and statistical analysis—as described below—was performed to determine any significant differences in size between treatment groups.

Statistical analysis

Statistical analysis was performed for non-binary discrete variables. Results are expressed as mean \pm standard error of the mean. Bartlett's variance test was used to confirm equal variance among groups and then a one-way analysis of variance (ANOVA) with Tukey HSD post-hoc tests was used to determine statistical significance. All statistical analysis was completed with a variety of statistical and graphical packages in R ^{23,24,25,26}. All results were considered significant if $p < 0.05$ on a two-tailed significance test.

Results

rTRD01 protects against paraquat-induced neurodegeneration in zebrafish.

Having demonstrated rTRD01 to have neuroprotective properties *in vitro*, we proceeded to assess the neuroprotective effects of rTRD01 *in vivo* using zebrafish as our model system. rTRD01 was well-tolerated *in vivo* and produced no discernable morphological defects at concentrations up to 75 μ M (the highest tested) (**Figure 3-2**). We proceeded to test rTRD01's ability to protect against paraquat-induced neurotoxicity. Consistent with previous reports^{8,9}, treatment of zebrafish embryos at 18 hpf with 0.04 ppm paraquat produced gross spinal curvature (**Figure 3-3 A**), a phenotype correlated with neurodegeneration. Co-treatment with 50 μ M rTRD01 at 12 hpf rescued the curved spine morphology (**Figure 3-3 A, B**), as did sequential treatment with rTRD01 (i.e. rTRD01, washout, then add paraquat) (**Figure 3-4**). Whereas paraquat treatment at 18 hpf induced neuronal death (**Figure 3-3 D**), as indicated by small but statistically significant reductions in brain cross-sectional area (**Figure 3-3 E**), mean fluorescence (**Figure 3-3 F**), and fluorescence density (**Figure 3-3 G**) compared to DMSO-treated controls, no difference in these metrics was observed between DMSO controls and zebrafish co-treated with 50 μ M rTRD01 at 12 hpf and paraquat at 18 hpf.

rTRD01 is not a general antioxidant.

To exclude a general antioxidant effect by rTRD01, we treated zebrafish with varying concentrations of rTRD01 for 7 days and then exposed the fish to rotenone, an inhibitor of mitochondrial respiratory chain complex I and a potent inducer of reactive oxygen species (ROS) formation²⁷. rTRD01 had no protective effect over DMSO in preventing brain death at all concentrations tested (**Figure 3-5**), indicating that rTRD01 does not act as a general antioxidant

(unlike probucol²⁸, for example). In summary, rTRD01 protects against paraquat-induced neurotoxicity *in vivo* and is not a general antioxidant.

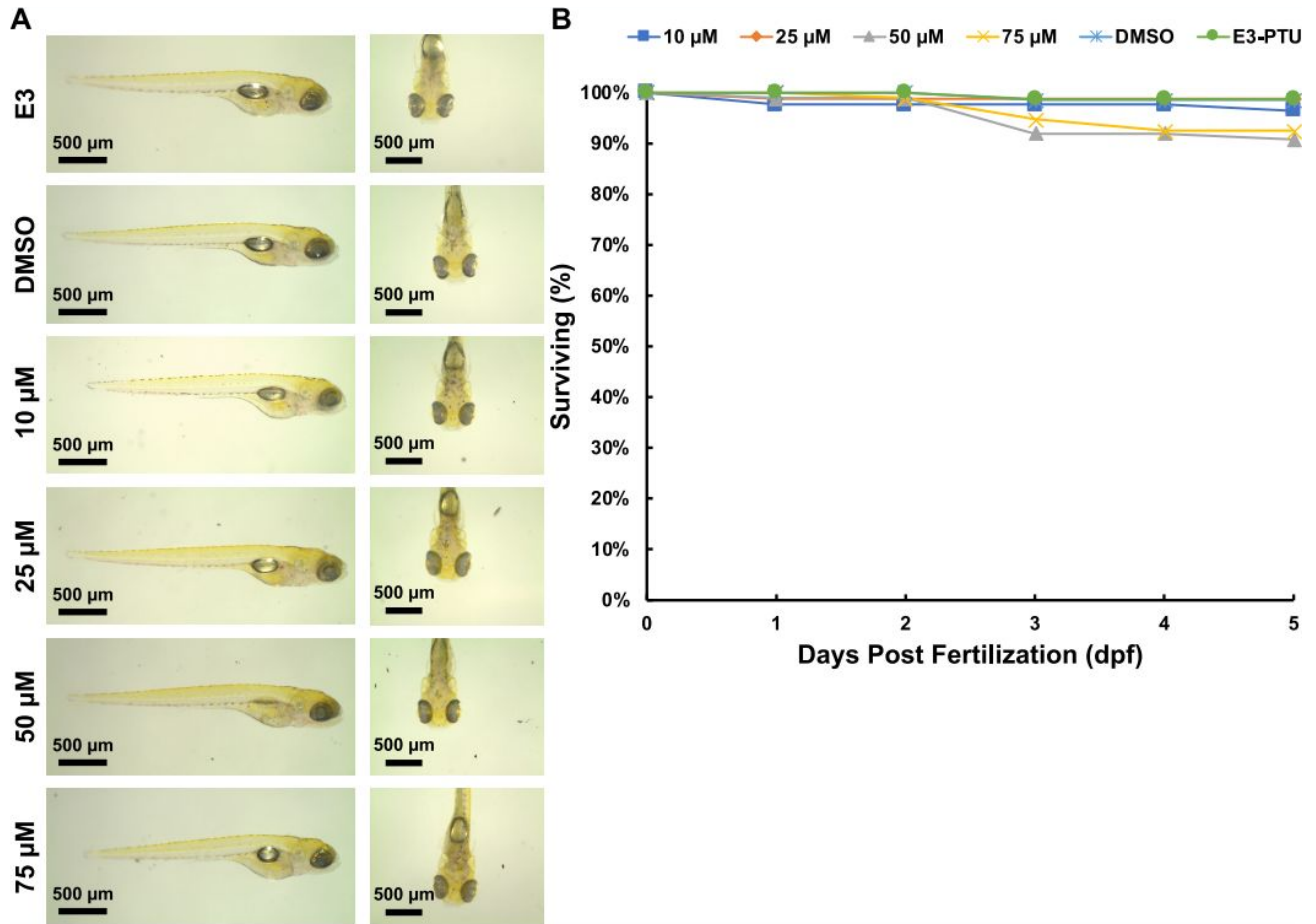


Figure 3-2. Toxicity titration and survival curve for rTRD01 exposure in zebrafish.

(A) Gross morphology (lateral and dorsal) of 5.5 dpf zebrafish incubated at 12 hpf in E3 medium containing DMSO or 10–75 μM rTRD01, as indicated. (B) Surviving zebrafish as percent of total, from 1 dpf to 5 dpf.

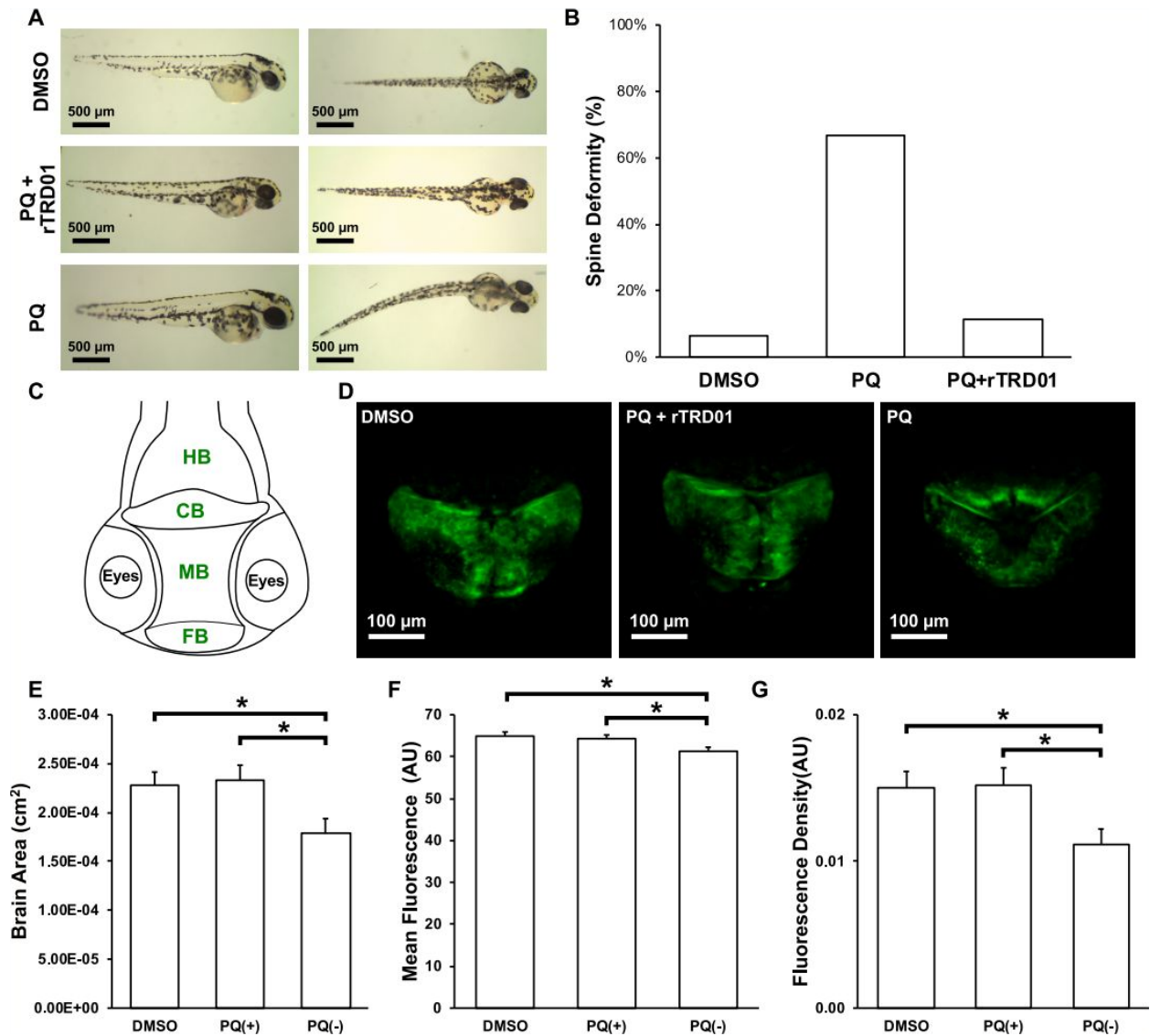


Figure 3-3. rTRD01 rescues paraquat-induced neurodegenerative phenotypes in zebrafish.

(A) Gross morphology of 2.5 dpf zebrafish embryos incubated in E3 medium containing (i) DMSO; (ii) DMSO added at 12 hpf and 0.04 ppm PQ added at 18 hpf; (iii) 50 μ M rTRD01 added at 12 hpf and 0.04 ppm PQ added at 18 hpf. (B) Quantification of curved spine phenotype prevalence in treatment conditions of (A). (C) Schematic depicting the spatial localization of the zebrafish brain regions. Green text represents tissues expressing GFP. (D) Fluorescence images of brain of 3.25 dpf zebrafish embryos incubated in E3-PTU medium containing (i) DMSO; (ii)

0.04 ppm PQ added at 18 hpf; (iii) 50 μ M rTRD01 added at 12 hpf and 0.04 ppm PQ added at 18 hpf. (E–G) Quantification of (E) maximal cross-sectional area, (F) mean fluorescence, and (G) fluorescence density of zebrafish brains subjected to the treatment conditions in (D).

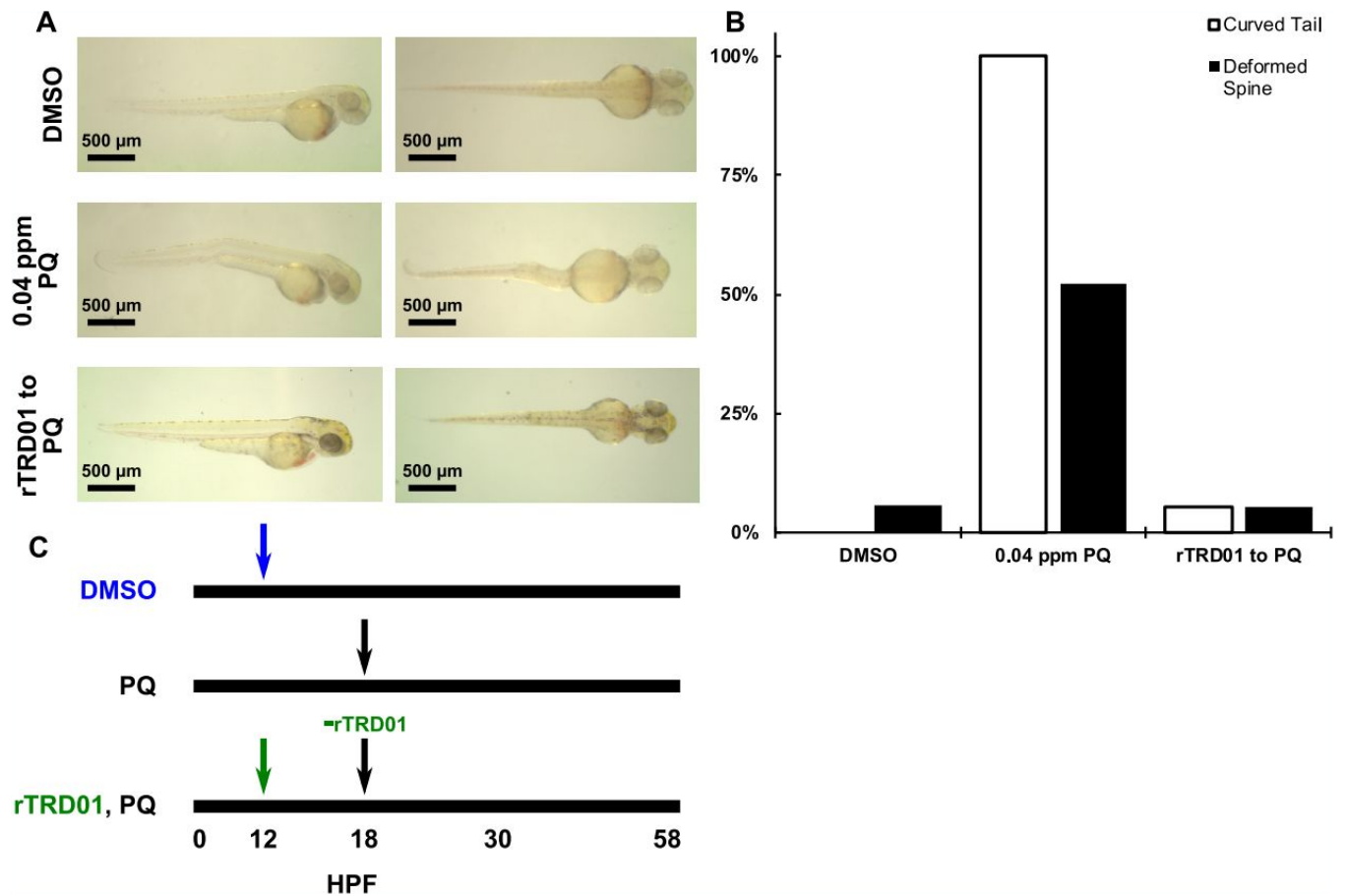


Figure 3-4. Neuroprotection by rTRD01 is not the result of chemical inactivation of paraquat by rTRD01.

(A) Gross morphology (lateral and dorsal) of 2.5 dpf zebrafish incubated in E3-PTU medium containing (i) DMSO added at 12 hpf; (ii) DMSO added at 12 hpf and 0.04 ppm PQ added at 18 hpf; (iii) 50 μ M rTRD01 added at 12 hpf and removed at 18 hpf, followed by addition of 0.04 ppm PQ at 18 hpf. (B) Quantification of phenotype prevalence in treatment conditions of (A). (C) Schematic of small molecule addition times. rTRD01 was removed by washing 3X with E3-PTU medium prior to addition of PQ.

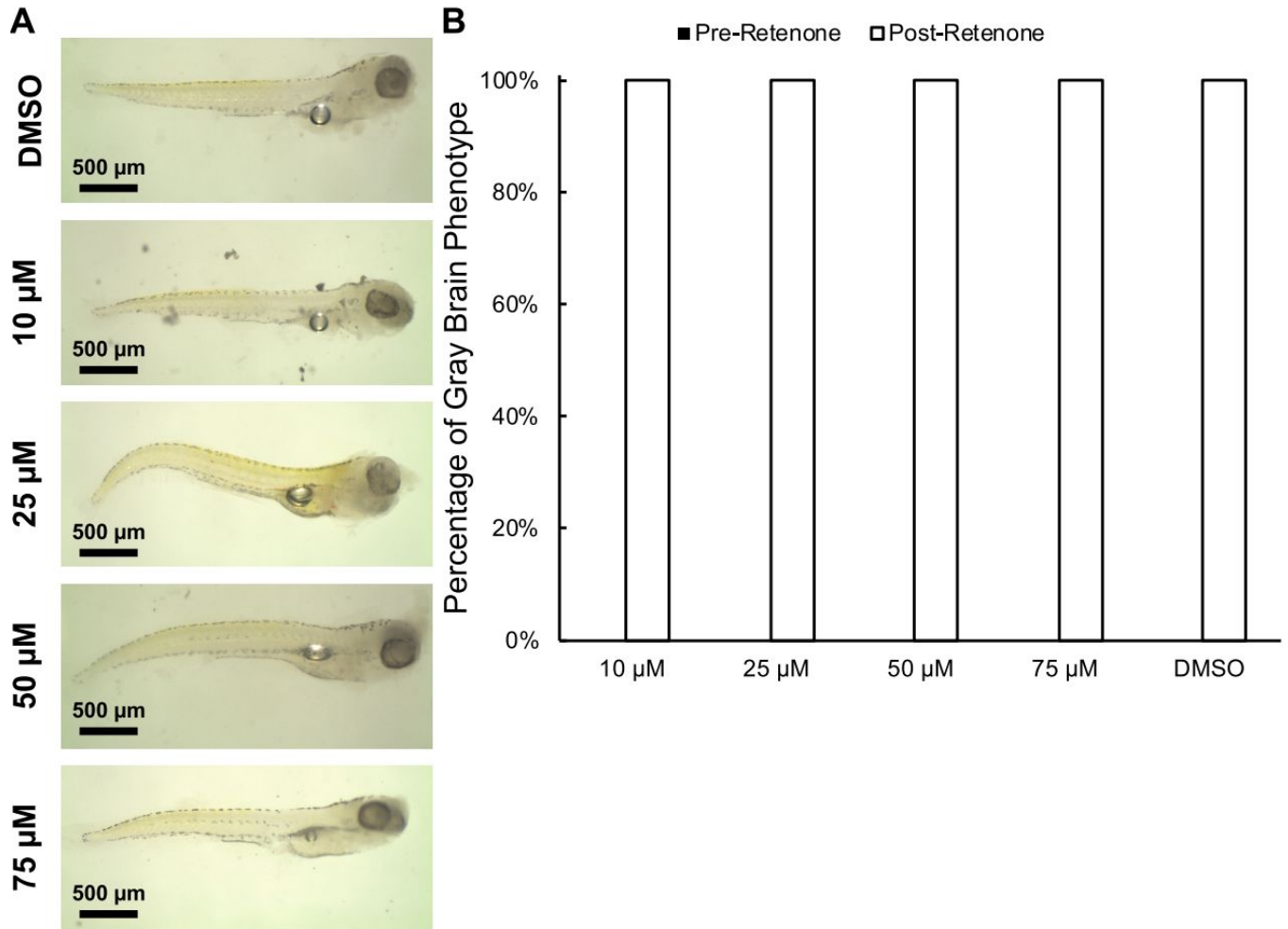


Figure 3-5. rTRD01 does not rescue rotenone-induced neuronal death in zebrafish.

(A) Gross morphology and neuronal transparency of 7 dpf zebrafish embryos after 4 hr incubation in 100 nM rotenone. Zebrafish were incubated at 12 hpf in E3-PTU medium containing either DMSO or 10–75 μM rTRD01, as indicated, prior to the 4 hr rotenone challenge at 7 dpf. (B) Quantification of gray brain phenotype prevalence in treatment conditions of (A).

Discussion

Using a chemical biology approach analogous to the workflow detailed in Chapter 2, we successfully identified a small molecule capable of [1] binding the RRM domains of TDP-43; [2] disrupting the association of TDP-43 with a pathologic RNA repeat expansion sequence; and [3] exerting *in vivo* neuroprotective effects in the zebrafish paraquat model. Future studies will evaluate the specificity of the association of rTRD01 with the RRM domains of TDP-43 (the secondary structures of RRM domains in general are highly conserved²⁹) and, likewise, whether rTRD01 attenuates ALS disease phenotypes *via* a TDP-43-specific mechanism or through multiple modes of action.

References

1. Kuo, P.-H., Doudeva, L. G., Wang, Y.-T., Shen, C.-K. J. & Yuan, H. S. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res.* **37**, 1799–1808 (2009).
2. Chiang, C.-H. *et al.* Structural analysis of disease-related TDP-43 D169G mutation: linking enhanced stability and caspase cleavage efficiency to protein accumulation. *Sci. Rep.* **6**, 21581 (2016).
3. Aulas, A., Stabile, S. & Vande Velde, C. Endogenous TDP-43, but not FUS, contributes to stress granule assembly via G3BP. *Mol. Neurodegener.* **7**, 54 (2012).
4. Fernandes, N., Eshleman, N. & Ross Buchan, J. Stress Granules and ALS: A Case of Causation or Correlation? *Advances in Neurobiology* 173–212 (2018). doi:10.1007/978-3-319-89689-2_7
5. Chen, Y. & Cohen, T. J. Aggregation of the nucleic acid-binding protein TDP-43 occurs via distinct routes that are coordinated with stress granule formation. *J. Biol. Chem.* **294**, 3696–3706 (2019).
6. Voigt, A. *et al.* TDP-43-mediated neuron loss in vivo requires RNA-binding activity. *PLoS One* **5**, e12247 (2010).
7. Barber, S. C. & Shaw, P. J. Oxidative stress in ALS: key role in motor neuron injury and therapeutic target. *Free Radic. Biol. Med.* **48**, 629–641 (2010).
8. Nellore, J. & P, N. Paraquat exposure induces behavioral deficits in larval zebrafish during the window of dopamine neurogenesis. *Toxicol Rep* **2**, 950–956 (2015).

9. Wang, Q. *et al.* Identification of apoptosis and macrophage migration events in paraquat-induced oxidative stress using a zebrafish model. *Life Sci.* **157**, 116–124 (2016).
10. Chen, Q. *et al.* The toxic influence of paraquat on hippocampus of mice: Involvement of oxidative stress. *NeuroToxicology* **31**, 310–316 (2010).
11. Niveditha, S., Ramesh, S. R. & Shivanandappa, T. Paraquat-Induced Movement Disorder in Relation to Oxidative Stress-Mediated Neurodegeneration in the Brain of *Drosophila melanogaster*. *Neurochem. Res.* **42**, 3310–3320 (2017).
12. Spivey, A. Rotenone and paraquat linked to Parkinson's disease: human exposure study supports years of animal studies. *Environ. Health Perspect.* **119**, A259 (2011).
13. Tanner, C. M., Goldman, S. M., Ross, G. W. & Grate, S. J. The disease intersection of susceptibility and exposure: chemical exposures and neurodegenerative disease risk. *Alzheimers. Dement.* **10**, S213–25 (2014).
14. Islam, R., Kumimoto, E. L., Bao, H. & Zhang, B. ALS-linked SOD1 in glial cells enhances β -N-Methylamino L-Alanine (BMAA)-induced toxicity in *Drosophila*. *F1000Res.* **1**, 47 (2012).
15. Parker, S. J. *et al.* Endogenous TDP-43 localized to stress granules can subsequently form protein aggregates. *Neurochem. Int.* **60**, 415–424 (2012).
16. Castello, P. R., Drechsel, D. A. & Patel, M. Mitochondria are a major source of paraquat-induced reactive oxygen species production in the brain. *J. Biol. Chem.* **282**, 14186–14193 (2007).
17. Smith, L. L. Paraquat toxicity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **311**, 647–657 (1985).

18. Vandenbogaerde, J., Schelstraete, J., Colardyn, F. & Heyndrickx, A. Paraquat poisoning. *Forensic Science International* **26**, 103–114 (1984).
19. Majounie, E. *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* **11**, 323–330 (2012).
20. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).
21. Levesque, M. P., Krauss, J., Koehler, C., Boden, C. & Harris, M. P. New tools for the identification of developmentally regulated enhancer regions in embryonic and adult zebrafish. *Zebrafish* **10**, 21–29 (2013).
22. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
23. Fox, J. & Weisberg, S. *An R Companion to Applied Regression*. (SAGE Publications, 2018).
24. Wickham, H. The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software* **40**, (2011).
25. R Core Team. R: A Language and Environment for Statistical Computing. (2018).
26. Wickham, H., François, R., Henry, L. & Müller, K. dplyr: A Grammar of Data Manipulation. (2018).
27. Li, N. *et al.* Mitochondrial Complex I Inhibitor Rotenone Induces Apoptosis through Enhancing Mitochondrial Reactive Oxygen Species Production. *J. Biol. Chem.* **278**, 8516–8525 (2003).

28. Byrnes, J. *et al.* Pharmacologic modeling of primary mitochondrial respiratory chain dysfunction in zebrafish. *Neurochemistry International* **117**, 23–34 (2018).
29. Afroz, T., Cienikova, Z., Cléry, A. & Allain, F. H. T. One, Two, Three, Four! How Multiple RRM_s Read the Genome Sequence. *Methods Enzymol.* **558**, 235–278 (2015).

Chapter 4:

Toward a zebrafish model of pontocerebellar hypoplasia type 1D

Chapter 4 describes the joint work of Gary Shmorgon, Kayla White, and Shahriyar Jahanbakhsh. Gary Shmorgon and Kayla White generated the mutation described using CRISPR/Cas9 mutagenesis. Shahriyar Jahanbakhsh characterized the mutation and isolated the mutant line.

Introduction

PCH1D is caused by mutations in SLC25A46, a transmembrane protein of the outer mitochondrial membrane. Seven disease-causing point mutations (five missense and two nonsense) are known, five of which localize to the fifth of SLC25A46's seven predicted transmembrane helices¹. Clinically, PCH1D is distinguishable from other PCH1 subtypes by optic nerve involvement, respiratory failure and early death, and is at the most severe end of the broad spectrum of SLC25A46-related diseases².

The sole laboratory model of PCH1D, a SLC25A46 knockout mouse model³, offers no means for studying the *in vivo* perturbations in mitochondrial dynamics central to the pathophysiology of PCH1D. Thus, an unmet need exists for new *in vivo* models of PCH1D. In this closing chapter, I describe progress toward a zebrafish knockout model of PCH1D.

MitoTimer: a fluorophore for assaying mitochondrial turnover

Given that SLC25A46 is a transmembrane protein located in the outer mitochondrial membrane, any model of PCH1D must provide for mitochondrial visualization. The oxidation-sensitive, matrix-targeted fluorescent protein marker MitoTimer has emerged as a valuable tool for the study of mitochondrial morphology, network dynamics, import, and turnover^{4,5,6}. MitoTimer is a matrix-targeted derivative of DsRed1-E5, a mutated form of DsRed whose fluorescence shifts from green to red as the protein matures. MitoTimer, like DsRed1-E5, is sensitive to oxygen but insensitive to pH, protein concentration, and ionic strength⁴.

Because MitoTimer fluoresces green when newly synthesized and shifts irreversibly to red fluorescence as it matures and oxidizes, MitoTimer serves as a molecular clock to give temporal and spatial information on protein turnover. A decrease in the number of mitochondria

importing cytosolic preproteins, including freshly synthesized, green MitoTimer, is therefore reflected by a decrease in green MitoTimer signal.

Materials and Methods

Zebrafish culture

AB zebrafish (*Danio rerio*) were maintained in accordance with standard laboratory conditions⁷. The University of California, Los Angeles Chancellor's Animal Research Committee approved all experiments performed on zebrafish (ARC #2005-141-32).

Zebrafish mutagenesis

Nascently-fertilized zebrafish embryos were microinjected with ~500 pL of a solution containing 200 ng/ μ L sgRNA (sequence: UUGGAAUUGGUCUCGCAAG) and 100 ng/ μ L Cas9 nuclease (Synthego, Redwood City, CA). F₀ founders were genotyped using an RFLP assay, and those founders bearing a mutation in SLC25A46 were isolated and bred to yield F₁ animals. F₁ animals were raised to maturity and genotyped to determine zygosity, and the sequence of the mutation was determined using Sanger sequencing (Genewiz, South Plainfield, NJ). The sequences of the PCR primers used to generate the amplicons for genotyping were forward: GTAAGGCAAGGAGAGCATACTTTGG and reverse: ACACTGATGGCTGTCATTGG. Sequences were aligned using the EMBOSS Needle pairwise sequence alignment tool⁸.

Results

PCR-RFLP genotyping of F₁ zebrafish identified four heterozygous animals that were confirmed by Sanger sequencing to harbor the same two base-pair (bp), frameshift-inducing insertion in exon 2 of SLC25A46 (**Figure 4-1**). Analysis of the amino acid sequence indicated

that the insertion, which occurs in an early exon of SLC25A46 (exon 2 of 8⁹), corrupts 78 percent of the translated protein. The heterozygous animals were isolated as mating pairs for propagating the mutant line.

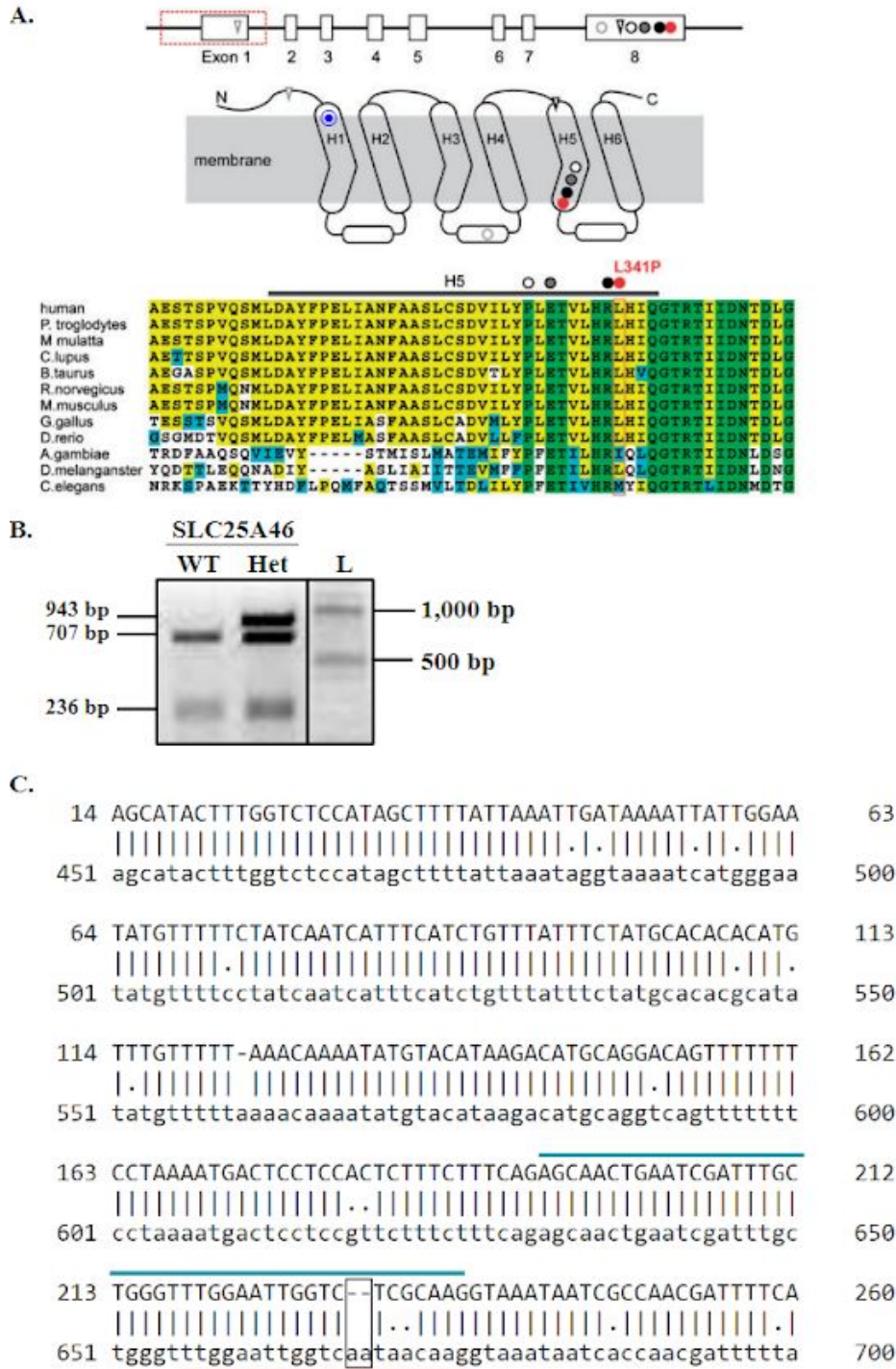


Figure 4-1. Heterozygous knockout of SLC25A46 in the zebrafish. (A)¹ (Top) Predicted membrane architecture of SLC25A46. Circles indicate missense point mutations. Arrowheads indicate nonsense point mutations. The dashed red box indicates a disease-causing deletion. The 2 bp insertion (blue target) localizes to the first transmembrane helix of SLC25A46. (Bottom) Alignment of protein sequences encoded by SLC25A46 orthologs in the human, zebrafish, and other eukaryotes. Adapted with permission from Oxford University Press. (B) RFLP-based screening for CRISPR/Cas9 mutagenesis. Two bands indicate no mutation (wild-type genotype). Three bands indicate a heterozygous mutation. (C) Sequence alignment identifying the mutation as a 2 bp insertion (boxed) in exon 2 (overlined in turquoise). Uppercase letters are the reference sequence. Lowercase letters are the sample sequence. The sgRNA binding site is underlined in red.

Discussion

We have established a zebrafish SLC25A46 knockout model for the purpose of interrogating the pathomechanism of PCH1D *in vivo* and in real time. When used in combination with MitoTimer, our model is poised to advance our understanding of the mitochondrial basis of PCH1D.

Knockout of SLC25A46 is hypothesized to prevent mitochondrial fission, and therefore also prevent proper motor-dependent transport on actin or microtubule networks¹⁰. By promoting mitochondrial hyperfusion, SLC25A46 knockout is also expected to attenuate the selective sequestration of mitochondria by autophagosomes^{11,12}, known as mitophagy. Loss of function missense mutations in SLC25A46 are known to disrupt interactions between mitochondria and the ER membrane and alter mitochondrial membrane lipid composition [i.e. loss of many species

of phosphatidylethanolamine (PE), phosphatidylserine, and phosphatidic acid]¹³. Depletion of PEs in particular is known to decrease the activity of the translocase of the outer membrane complex and reduce translocation of beta-barrel precursors across the OMM¹⁴. Loss of PEs may also decrease supercomplexation of respiratory chain complexes III and IV¹⁵, resulting in less efficient catalysis^{15,16} and a reduction in membrane potential. Therefore, loss of SLC25A46 is also expected to reduce mitochondrial protein import, thus increasing red fluorescence from MitoTimer.

References

1. Wan, J. *et al.* Loss of function of SLC25A46 causes lethal congenital pontocerebellar hypoplasia. *Brain* **139**, 2877–2890 (2016).
2. van Dijk, T. *et al.* Pontocerebellar hypoplasia with spinal muscular atrophy (PCH1): identification of SLC25A46 mutations in the original Dutch PCH1 family. *Brain* **140**, e46 (2017).
3. Terzenidou, M. E. *et al.* Novel insights into SLC25A46-related pathologies in a genetic mouse model. *PLoS Genet.* **13**, e1006656 (2017).
4. Hernandez, G. *et al.* MitoTimer: a novel tool for monitoring mitochondrial turnover. *Autophagy* **9**, 1852–1861 (2013).
5. Gottlieb, R. A. & Stotland, A. MitoTimer: a novel protein for monitoring mitochondrial turnover in the heart. *J. Mol. Med.* **93**, 271–278 (2015).
6. Laker, R. C. *et al.* A novel MitoTimer reporter gene for mitochondrial content, structure, stress, and damage in vivo. *J. Biol. Chem.* **289**, 12005–12015 (2014).
7. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).
8. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz268
9. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).
10. Lackner, L. L. Shaping the dynamic mitochondrial network. *BMC Biology* **12**, (2014).

11. Gomes, L. C., Di Benedetto, G. & Scorrano, L. During autophagy mitochondria elongate, are spared from degradation and sustain cell viability. *Nat. Cell Biol.* **13**, 589–598 (2011).
12. Sgarbi, G. *et al.* Mitochondria hyperfusion and elevated autophagic activity are key mechanisms for cellular bioenergetic preservation in centenarians. *Aging* **6**, 296–310 (2014).
13. Janer, A. *et al.* SLC25A46 is required for mitochondrial lipid homeostasis and cristae maintenance and is responsible for Leigh syndrome. *EMBO Mol. Med.* **8**, 1019–1038 (2016).
14. Becker, T. *et al.* Role of phosphatidylethanolamine in the biogenesis of mitochondrial outer membrane proteins. *J. Biol. Chem.* **288**, 16451–16459 (2013).
15. Tasseva, G. *et al.* Phosphatidylethanolamine deficiency in Mammalian mitochondria impairs oxidative phosphorylation and alters mitochondrial morphology. *J. Biol. Chem.* **288**, 4158–4173 (2013).
16. Milenkovic, D., Blaza, J. N., Larsson, N.-G. & Hirst, J. The Enigma of the Respiratory Chain Supercomplex. *Cell Metab.* **25**, 765–776 (2017).