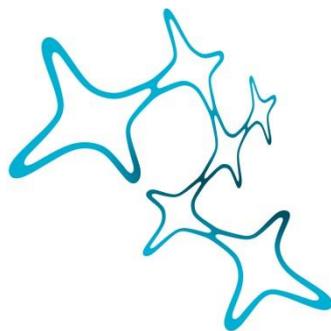

IDENTIFICATION OF A UNIQUE
SPLICE ACCEPTOR SITE ENABLES
NOVEL BIOTECHNOLOGICAL
APPLICATION

Lisa Maria Riedmayr



Graduate School of
Systemic Neurosciences

LMU Munich



Dissertation at the
Graduate School of Systemic Neurosciences
Ludwig-Maximilians-Universität München

July, 2019

Supervisor
Dr. Elvir Becirovic
Department of Pharmacy
Ludwig-Maximilians-Universität München

First Reviewer: Dr. Elvir Becirovic
Second Reviewer: Prof. Michael Kiebler
External Reviewer: Prof. Knut Stieger

Date of Submission: 31.07.2019
Date of Defense: 05.12.2019

TABLE OF CONTENTS

List of Abbreviations.....	7
List of Tables	13
List of Figures	14
Abstract.....	15
1 Introduction.....	17
1.1 The Retina – a Remarkable Sensory Tissue	17
1.1.1 Anatomy of the Retina.....	17
1.1.2 Phototransduction Cascade	19
1.1.3 Inherited Retinal Diseases.....	20
1.2 Splicing – the Art of Removing Introns	23
1.2.1 The Splicing Mechanism.....	23
1.2.2 Mutations Affecting mRNA Splicing.....	25
1.2.3 mRNA <i>Trans</i> -Splicing.....	26
1.3 Gene Therapy – a Weapon Against Genetic Diseases	27
1.3.1 A Brief History of Gene Therapy	27
1.3.2 Recombinant Adeno-Associated Virus-Mediated Gene Supplementation in the Retina	30
1.3.3 rAAV Dual Vector Strategies	32
2 Aim of the Study.....	39
3 Materials & Methods	40
3.1 Materials	40
3.2 <i>In Silico</i> Splice Prediction Analysis	40
3.3 Cloning.....	40
3.3.1 Expression Vectors.....	40
3.3.2 Standard Cloning Techniques	42
3.3.3 Site-Directed Mutagenesis.....	45

Table of Contents

3.4	Cell Culture And Transfection.....	46
3.4.1	HEK293 and HEK293T Cells.....	46
3.4.2	661W Cells.....	47
3.4.3	Mouse Embryonic Fibroblasts.....	48
3.5	Production of Recombinant Adeno-Associated Viruses.....	48
3.5.1	Transfection and Harvest.....	48
3.5.2	Iodixanol Density Gradient Centrifugation.....	50
3.5.3	Anion Exchange Chromatography.....	51
3.5.4	Increase of rAAV Concentration.....	51
3.5.5	rAAV Titer Determination.....	52
3.6	Animals.....	54
3.7	Subretinal Injection.....	54
3.8	Immunohistochemistry.....	54
3.9	Confocal Microscopy.....	56
3.10	RNA Extraction.....	56
3.11	cDNA Synthesis.....	57
3.12	Reverse Transcription PCR.....	58
3.13	Branch Point Analysis.....	60
3.13.1	Nested Lariat RT-PCR.....	60
3.13.2	TOPO Cloning and Analysis of Lariats.....	61
3.14	Protein Extraction.....	61
3.15	Western Blotting.....	62
3.16	Statistics.....	65
4	Results.....	66
4.1	Investigation of Two adRP-Causing Rhodopsin Mutations and Their Effect on mRNA Splicing.....	66

4.1.1	Generation of <i>De Novo</i> Splice Acceptor Site in <i>RHO</i> Exon 3 by c.620T>G Mutation	66
4.1.2	Mislocalization of c.620T>G Minigene-Born Protein <i>in Vitro</i>	68
4.1.3	Missplicing and Mislocalization of the c.620T>G Mutant in Mouse Photoreceptors.....	70
4.1.4	Identification of Core Element Underlying the c.620T>G SAS Strength.....	72
4.1.5	Independence of SAS_620 Strength from Gene Environment.....	74
4.1.6	High Flexibility of the c.620T>G SAS in Branch Point Choice	77
4.2	Development of a Novel Dual Vector Technology Using the SAS_620 Splicing Element	80
4.2.1	Reconstitution of a Fluorescent Reporter Using mRNA <i>Trans</i> -Splicing.....	80
4.2.2	Impact of Splice Acceptor Site vs. Binding Domain on Reconstitution Efficiency	84
4.2.3	Reconstitution via mRNA <i>Trans</i> -Splicing in Different Cell Types	86
4.2.4	Investigation of mRNA <i>Trans</i> -Splicing-Based rAAV Dual Vectors <i>in Vivo</i>	87
4.2.5	Identification of Potent Binding Domain Suitable for Human Gene Therapy.....	89
4.2.6	Proof of Principle for Reconstitution of Large Genes via mRNA <i>Trans</i> -Splicing.....	90
5	Discussion.....	92
5.1	Implications of c.620T>G Mutation Creating a Splice Acceptor Site with Unique Properties	92
5.2	Evaluation of Novel mRNA <i>Trans</i> -Splicing Dual Vector Approach.....	94
6	References.....	98
7	Appendix.....	110
7.1	List of RT-PCR Primers.....	110

Table of Contents

7.2 List of Binding Domain Sequences.....	111
7.3 Reconstitution Efficiencies	113
List of Publications	115
Acknowledgements	116
Affidavit.....	118
Declaration of Author Contributions.....	119

LIST OF ABBREVIATIONS

µg	microgram
µl	microliter
µM	micromolar
aa	amino acid
AAV	adeno-associated virus
<i>ABCA4</i>	human ATP-binding cassette transporter A4 gene
ABCR	ATP-binding cassette transporter
<i>ACADM</i>	human acyl-CoA dehydrogenase medium chain gene
<i>ADGRV1</i>	human adhesion G protein-coupled receptor V1 gene
adRP	autosomal dominant retinitis pigmentosa
APS	ammonium persulfate
arRP	autosomal recessive retinitis pigmentosa
<i>ASCL1</i>	human achaete-scute family bHLH transcription factor 1 gene
ASSP	Alternative Splice Site Predictor
BBS	BES buffered saline
BD	binding domain
BES	N,N-Bis(2-hydroxyethyl)-2-aminoethanesulfonic acid
BP	branch point
bp	base pair
bGH	bovine growth hormone
c	concentration
Ca ²⁺	calcium ion
CaCl ₂	calcium chloride
<i>CACNA1F</i>	human calcium voltage-gated channel subunit alpha1 F gene
Cas9	CRISPR-associated protein 9
<i>CDH23</i>	human cadherin 23 gene
cDNA	complementary DNA
CDS	coding sequence
<i>CEP290</i>	human centrosomal protein of 290 kDa gene
cf.	confer
cGMP	cyclic guanosine monophosphate
C-intein	C-terminal part of the split intein

List of Abbreviations

<i>CLRN1</i>	human clarin-1 gene
CMV	cytomegalovirus
CNG	cyclic nucleotide gated
CNGA1	cyclic nucleotide gated channel subunit A1
CNGB1	cyclic nucleotide gated channel subunit B1
CNGB1a	cyclic nucleotide gated channel subunit B1a
CO ₂	carbon dioxide
<i>CRB1</i>	human crumbs cell polarity complex component 1 gene
CRISPR	clustered regulatory interspaced short palindromic repeats
Ct	cycle threshold
dCas9	dead Cas9
DnaB	replicative DNA helicase
DnaE	DNA polymerase III subunit alpha
DMD	Duchenne muscular dystrophy
DMEM	Dulbecco's Modified Eagle Medium
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
e.g.	<i>exempli gratia</i>
EL2	second extracellular loop
ESE	exonic splice enhancer
ESS	exonic splice silencer
<i>EYS</i>	human eyes shut homolog gene
FBS	fetal bovine serum
FL	full-length
G	gauge
g	gram
GC	guanylyl cyclase
GPCR	G protein-coupled receptor
gRNA	guide RNA
G _T	G-protein transducin
h	hour

List of Abbreviations

H ₂ O	water
H3	transmembrane helix domain 3
H5	transmembrane helix domain 5
<i>HBQ1</i>	human hemoglobin subunit theta 1 gene
HCl	hydrogen chloride
HDR	homology directed repair
HEK293	human embryonic kidney 293
HEK293T	HEK293 derived Lenti-X 293T
HSC	hematopoietic stem cell
HSF	Human Splicing Finder
Hz	hertz
i.e.	<i>id est</i>
IS	inner segment
ISE	intronic splice enhancer
ISS	intronic splice silencer
ITR	inverted terminal repeats
kb	kilobase
KCl	potassium chloride
kDa	kilodalton
kg	kilogram
l	liter
LB	lysogeny broth
LCA	Leber congenital amaurosis
M	molar
MΩ	megaohm
<i>MAPT</i>	human microtubule associated protein tau gene
MEF	mouse embryonic fibroblast
Met207	methionine residue at position 207
mg	milligram
MgCl ₂	magnesium chloride
min	minute
<i>MIP</i>	human major intrinsic protein gene
MM	molar mass

List of Abbreviations

mRNA	messenger ribonucleic acid
<i>MYO7A</i>	human myosin VIIA gene
n	number of independent samples
Na ⁺	sodium ion
Na ₂ HPO ₄	disodium hydrogen phosphate
NaH ₂ PO ₄	sodium dihydrogen phosphate
NaCl	sodium chloride
ng	nanogram
N-intein	N-terminal part of the split intein
nm	nanometer
nr.	number
ONL	outer nuclear layer
OS	outer segment
P21	postnatal day 21
pA	polyadenylation signal
PB	phosphate buffer
PBS	phosphate buffered saline
PBS-MK	PBS-MgCl ₂ -KCl ₂
<i>PCDH15</i>	human protocadherin related 15 gene
PCR	polymerase chain reaction
PDE6	phosphodiesterase 6
PEG	polyethylene glycol
PES	polyethersulfone
PFA	paraformaldehyde
pg	picogram
PMMA	polymethylmethacrylate
PPT	polypyrimidine tract
pre-mRNA	precursor mRNA
PTM	pre-mRNA <i>trans</i> -splicing molecule
rAAV	recombinant adeno-associated virus
RDH	retinol dehydrogenase
<i>RHO</i>	human rhodopsin gene
RNA	ribonucleic acid

List of Abbreviations

RP	retinitis pigmentosa
<i>RP1</i>	human RP1 axonemal microtubule associated gene
RPE	retinal pigment epithelium
RPE65	RPE-specific 65-kDa protein
rpm	rounds per minute
<i>RPS27</i>	human ribosomal protein 27 gene
RT	room temperature
RT-PCR	reverse transcription PCR
s	second
<i>S100A12</i>	human S100 calcium binding protein A12 gene
SAS	splice acceptor site
SAS_620	core element of the c.620T>G SAS
SDS	sodium dodecyl sulfate
SDS	splice donor site
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SL	spliced leader
SMaRT	spliceosome-mediated mRNA <i>trans</i> -splicing
SNP	single nucleotide polymorphism
snRNP	small nuclear ribonucleoprotein
<i>SNRNP200</i>	human small nuclear ribonucleoprotein U5 subunit 200 gene
<i>SpCas9</i>	<i>Streptococcus pyogenes</i> Cas9
SV40	simian virus 40
TBE	tris-boric acid-EDTA
TBST	tris buffered saline with 0.1 % Tween-20
TEMED	tetramethylethylenediamine
Tris	tris-(hydroxymethyl) aminomethane
TX	Triton X-100
U	unit
<i>USH2A</i>	human usherin gene
UV	ultraviolet
V	volt
v	volume
vg	viral genomes

List of Abbreviations

vs.	versus
w	weight
WT	wild type

LIST OF TABLES

Table 1 Canonical consensus sequences of human splice elements.....	23
Table 2 IRD-causing genes exceeding the AAV packaging limit	32
Table 3 List of plasmids used in this study.....	41
Table 4 Herculase II PCR and primer elongation cycling conditions	43
Table 5 KAPA Quickchange PCR cycling conditions	46
Table 6 Calcium phosphate transfection mix	47
Table 7 Fast cycling program for real-time qPCR	53
Table 8 List of all antibodies used for immunohistochemistry.....	55
Table 9 Herculase II RT-PCR cycling conditions.....	59
Table 10 VWR Taq polymerase RT-PCR cycling conditions.....	59
Table 11 Herculase II nested lariat PCR cycling conditions.....	60
Table 12 List of all antibodies used for western blotting	64
Table 13 Predicted SAS strengths using different prediction software tools	75
Table 14 Detected branch points and corresponding prediction scores.....	79

LIST OF FIGURES

Figure 1 Schematic illustration of the retinal anatomy	18
Figure 2 Morphology of a rod photoreceptor and its phototransduction cascade.....	19
Figure 3 Position and function of Met207 in rhodopsin.....	22
Figure 4 Illustration of the splicing mechanism	24
Figure 5 Single nucleotide polymorphisms affecting splicing.....	25
Figure 6 Naturally occurring trans-splicing events.....	27
Figure 7 rAAV-mediated gene supplementation therapy.....	30
Figure 8 Schematic depiction of different rAAV dual vector strategies	34
Figure 9 Splice acceptor site elements in WT and mutant RHO exon 3	67
Figure 10 Generation of a de novo splice acceptor site in RHO exon 3 by c.620T>G.....	68
Figure 11 Live imaging dual fluorescence splicing detection assay	69
Figure 12 RHO minigene expression in mouse rod photoreceptors.....	71
Figure 13 Splicing of c.620T>G SAS sequences in the RPS27 minigene	73
Figure 14 Splicing of SAS_620 in independent gene environments	74
Figure 15 Splicing of SAS_620 in mouse embryonic fibroblasts.....	75
Figure 16 Introduction of SAS_620 into RHO intron 2.....	76
Figure 17 Mutation of potential branch points in RHO c.620T>G exon 3.....	77
Figure 18 Identification of branch points via nested lariat RT-PCR.....	78
Figure 19 Schematic illustration of the reporter assay used for evaluation of gene reconstitution via mRNA trans-splicing.....	81
Figure 20 Binding domain optimization for increase in reconstitution efficiency.....	82
Figure 21 Evaluation of trans-splicing precision and reconstitution efficiency.....	83
Figure 22 Impact of SAS vs. BD on reconstitution efficiency	85
Figure 23 Reconstitution via mRNA trans-splicing in different cell types	86
Figure 24 mRNA trans-splicing rAAV dual vector approach in vivo	88
Figure 25 Identification of suitable BDs originating from lacZ gene.....	89
Figure 26 Reconstitution of SpCas9-VPR.....	90

ABSTRACT

Disease-associated mutations located in the protein coding exonic regions are in general classified as missense mutations supposedly affecting single amino acids. However, exonic mutations can potentially also affect mRNA splicing suggesting that a substantial portion of disease-associated mutations might have been misclassified in the past.

Here, it was hypothesized that surprisingly differential phenotypes caused by similar mutations in the same gene and originally classified as missense mutations might originate from a differential impact on mRNA splicing. This hypothesis was evaluated using the example of two retinitis pigmentosa (RP) mutations in the rhodopsin gene which affect the same nucleotide but are linked to remarkably different phenotypes. The c.620T>G mutation is associated with a severe, early-onset RP, whereas c.620T>A causes a mild, late-onset disease phenotype. In this study, I show that c.620T>G, in contrast to c.620T>A, creates an exceptionally strong splice acceptor site (SAS) resulting in a 90 bp in-frame deletion and protein mislocalization *in vitro* and *in vivo*. Furthermore, I could identify the core element responsible for the c.620T>G SAS strength. Finally, I demonstrate that this SAS is highly flexible in branch point choice, which could explain its remarkable strength.

Apart from its role in pathogenicity, mRNA splicing is also playing an integrative part in some important biotechnological applications. This particularly applies to the process called mRNA *trans*-splicing that is capable of a seamless ligation of two separate mRNA molecules. Here, I investigated, whether mRNA *trans*-splicing can be used as a novel approach to overcome the limited genome packaging capacity of adeno-associated viral (AAV) vectors, which represents the most relevant limitation hindering their broad use for gene therapy. Current approaches addressing this issue focus on dual AAV vector strategies aiming at reconstitution of split genes at the genomic or at the protein level. However, these approaches harbor important disadvantages hampering their broad usage in therapeutically relevant contexts.

To provide a proof of principle for reconstitution of split genes at the mRNA level, I developed a split fluorophore assay which allows for a convenient detection and quantification of the reconstitution efficiencies of the split gene. I could identify two

key elements conferring high reconstitution efficiencies: the binding domain, which enables the tight interaction of two mRNA molecules, and the SAS facilitating the mRNA *trans*-splicing process. Using a combination of internally optimized binding domains and the remarkably strong c.620T>G SAS identified in the first part of this thesis very high reconstitution efficiencies up to 78 % were achieved *in vitro*. Moreover, by successfully applying this approach in the mouse retina, I provide a proof of principle for this technology *in vivo*. Finally, I show first evidence that mRNA *trans*-splicing can be utilized to reconstitute an important coding sequence exceeding the capacity of AAV vectors: the *SpCas9*-VPR module.

This novel technology offers key advantages over currently used approaches and might expand the treatment options for patients with inherited diseases.

1 INTRODUCTION

1.1 THE RETINA – A REMARKABLE SENSORY TISSUE

1.1.1 Anatomy of the Retina

The vertebrate retina is the part of the central nervous system enabling vision by turning light stimuli into electrical signal. This sensory tissue is constituting the innermost part of the eye and consists of five main neuronal cell types and their interconnecting synapses organized in different layers (Figure 1) (Hoon et al., 2014). The cell bodies of the highly specialized, light-sensing photoreceptors are forming the outer nuclear layer (Müller, 1854). In the adjacent outer plexiform layer, they are synapsing on bipolar cells, which are in turn processing the photoreceptor input and relaying the information further to retinal ganglion cells (Euler et al., 2014). The cell bodies of bipolar, horizontal, amacrine and Müller glia cells together form the inner nuclear layer of the retina. Horizontal cells are modulating the synaptic transmission between photoreceptors and bipolar cells. By providing inhibitory feedback to the photoreceptors, they are serving as a local gain control (Herrmann et al., 2011, Masland, 2012). Amacrine cells synapse onto bipolar and retinal ganglion cells thereby shaping and controlling ganglion cell responses (Baccus et al., 2008, Masland, 2001). By contrast, Müller glia cells are providing homeostatic and metabolic support to the retina. In the inner plexiform layer, bipolar cells are forming synapses with retinal ganglion cells, which are located in the ganglion cell layer. The axons of these cells are forming the optic nerve, leave the eye at the optic disc and transmit the visual information to higher processing areas in the brain.

The mammalian retina is highly conserved across species (Wassle and Boycott, 1991, Masland, 1996). One of its remarkable features is its inverted organization with the light-detecting photoreceptors in the outermost layer facing away from the entering light. There are two distinct types of photoreceptors, i.e. the rods and the cones (Schultze, 1866). Rods constitute 95 % of photoreceptors in the human retina and mediate dim light vision due to their high photon sensitivity (Curcio et al., 1990, Sampath and Rieke, 2004). By contrast, cone photoreceptors are facilitating central, high acuity vision. In primates, this type of photoreceptor is mostly localized in a

central spot of the retina called macula. Cones are 100-fold less light sensitive than rods, however, their responses are highly wavelength-specific and are therefore mediating color discrimination (Nikonov et al., 2006).

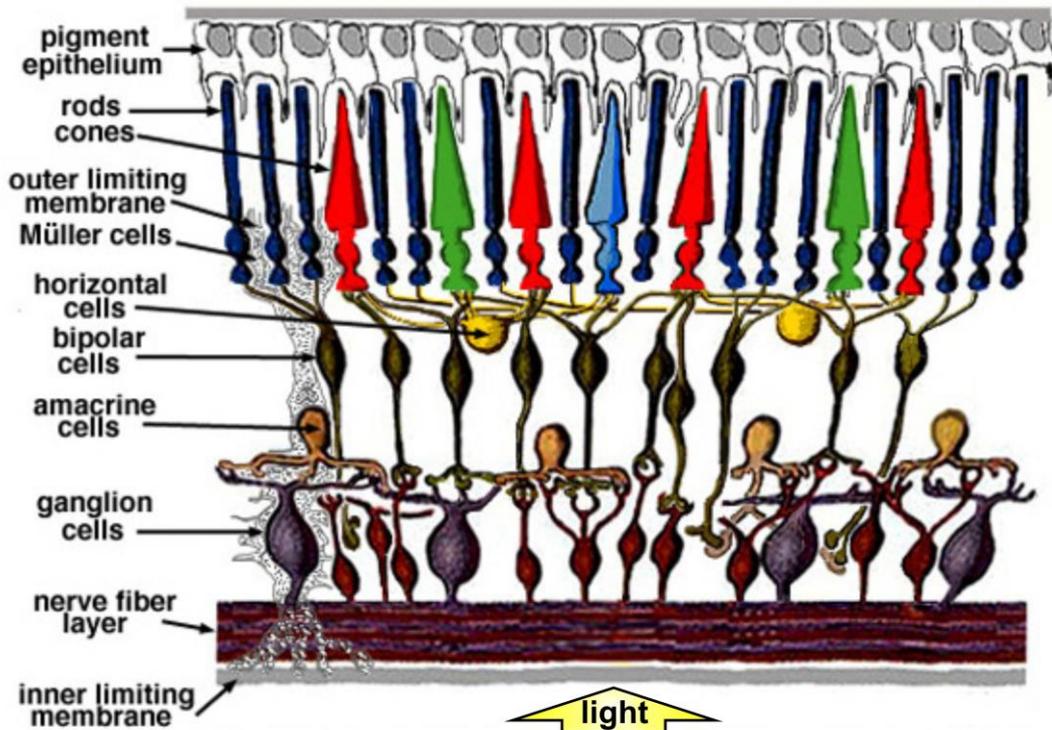


Figure 1 Schematic illustration of the retinal anatomy

The identities of retinal cell types and morphology defining structures are indicated on the left. Adapted from (Webvision, 2011).

Photoreceptors are compartmentalized cells, consisting of the light-sensitive outer segment (OS), an inner segment (IS) containing the organelles necessary for metabolism, a nucleus containing cell body and a synaptic terminal (Figure 2) (Carter-Dawson and LaVail, 1979). OS and IS are linked via a connecting cilium responsible for transport of cargo between these two compartments. The OS of rods and cones can easily be differentiated due to their distinct morphology. While rods are possessing long OS consisting of tightly stacked disks lacking a connection to the cell membrane, the shorter cone OS are composed of lamellar evaginations of the plasma membrane. The photoreceptor OS are embedded within the retinal pigment epithelium (RPE), which plays an important role in the maintenance of visual function (Young and Bok, 1969).

1.1.2 Phototransduction Cascade

The conversion of detected light into an electrical signal taking place in rod and cone photoreceptors is referred to as the phototransduction cascade (Arshavsky et al., 2002). As the signaling process is highly similar in both cell types, the rod phototransduction will be described exemplarily in further detail (Figure 2). During darkness, the constant activity of transmembrane guanylyl cyclases (GCs) maintains high levels of cyclic guanosine monophosphate (cGMP) in the cytosol (Yang et al., 1995). This cGMP is binding to cyclic nucleotide gated (CNG) channels keeping them in an open state and facilitating a constant influx of sodium (Na^+) and calcium ions (Ca^{2+}) called the dark current (Hagins et al., 1970). This continuous depolarization of the photoreceptor elicits a sustained glutamate transmitter release at the synapse during darkness.

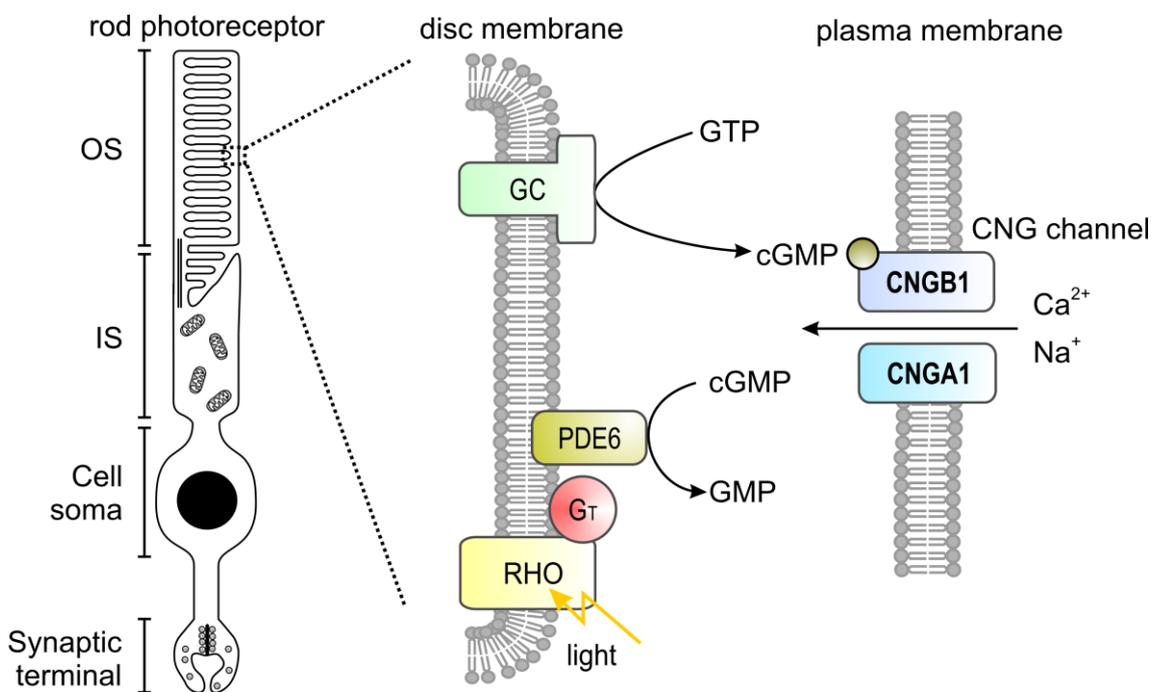


Figure 2 Morphology of a rod photoreceptor and its phototransduction cascade

Left panel, Morphology of a rod photoreceptor. OS, outer segment. IS, inner segment. Right panel, Phototransduction cascade in a rod photoreceptor. High cGMP levels are maintained via guanylyl cyclases (GCs) keeping the CNG channels, composed of CNGA1 and CNGB1 subunits, in an open state. After detection of light, rhodopsin (RHO) is activated, which in turn activates the G-protein transducin (G_T). The dissociation of its α -subunit stimulates the hydrolysis of cGMP by the phosphodiesterase 6 (PDE6), which leads to closure of CNG channels and a hyperpolarization of the photoreceptor. Modified from (Michalakis et al., 2018).

The process of vision begins with the detection of photons by the visual pigment, a G protein-coupled receptor (GPCR) called rhodopsin. This photopigment consists of the protein moiety opsin and the covalently bound chromophore 11-*cis*-retinal (Palczewski et al., 2000). After absorption of the photon, the 11-*cis*-retinal isomerizes into all-*trans*-retinal inducing conformational changes of rhodopsin into its active metarhodopsin II state. This in turn catalyzes an exchange of GDP with GTP at the heterotrimeric G-protein transducin. Subsequently, its α -subunit dissociates and stimulates the downstream effector cGMP phosphodiesterase 6 (PDE6). PDE6-mediated hydrolysis of cGMP leads to closure of the CNG channels and a hyperpolarization of the photoreceptor. As a consequence, the glutamate transmitter release at the synapse is interrupted and responses in the second-order neurons, i.e. bipolar cells, are triggered (Luo et al., 2008, Michalakakis et al., 2018).

After initiating the phototransduction cascade, the bleached photopigment needs to be recovered. To first inactivate metarhodopsin II, the protein is phosphorylated by the rhodopsin kinase and subsequently bound by arrestin. During these steps, the all-*trans*-retinal dissociates from the protein moiety and is regenerated within the retinoid visual cycle (Kiser et al., 2012). For this process, all-*trans*-retinal is reduced to all-*trans*-retinol via retinol dehydrogenases (RDHs). This step is further facilitated via the ATP-binding cassette transporter (ABCR), which moves all-*trans*-retinal bound by phosphatidylethanolamine into reach of the RDHs and thereby accelerates the reduction of toxic retinaldehydes (Quazi et al., 2012). Subsequently, the reduced chromophore diffuses through the interphotoreceptor matrix into the RPE, where it is esterified, transformed into 11-*cis*-retinol via the RPE-specific 65-kDa protein (RPE65) and oxidized into 11-*cis*-retinal via RDHs. Back in the photoreceptor OS, the 11-*cis*-retinal initiates de-phosphorylation of the opsin and again becomes part of a fully functional rhodopsin photopigment.

1.1.3 Inherited Retinal Diseases

Inherited retinal diseases (IRDs) are a genetically and phenotypically heterogeneous group of disorders. They are the leading cause of blindness in people aged 15 to 45 and have an estimated prevalence of 1 in 1,500 – 1 in 3,000 (Bundey and Crews, 1984, Puech et al., 1991). IRDs can be grouped according to the primarily

affected retinal cell type, i.e. rod or cone photoreceptors, and according to the state of disease progression.

The most common progressive IRD predominantly affecting cones is the Stargardt macular dystrophy displaying a prevalence of 1 in 8,000 – 1 in 10,000 (Tanna et al., 2017). It is characterized by an onset in childhood or early adolescence and leads to a degeneration of the macula thereby resulting in a loss of central vision. It is inherited in an autosomal recessive manner and it is mostly caused by mutations in the *ABCA4* gene encoding the retinal transporter ABCR (Fujinami et al., 2013).

With a prevalence of 1 in 4,000, the most common IRD primarily affecting rod photoreceptors is retinitis pigmentosa (RP). RP is a progressive disease leading to night blindness and tunnel vision (Hartong et al., 2006). In later stages, secondary cone photoreceptor cell death is induced eventually resulting in complete vision loss. However, the severity of the symptoms, rate of progression and age of onset show a striking variability in this disease. Depending on the affected gene and the nature of the mutation, retinitis pigmentosa can be inherited in a non-syndromic form as an autosomal dominant (12 – 35 %), autosomal recessive (40 – 60 %) or X-linked trait (4 – 15 %) (Grondahl, 1987, Ayuso et al., 1995, Bravo-Gil et al., 2017). Moreover, this dystrophy can occur in combination with other non-ocular symptoms as part of a syndrome. The most frequent syndromic form is the Usher syndrome accounting for 30 – 40 % of all autosomal recessive retinitis pigmentosa (arRP) cases (Boughman et al., 1983). It is characterized by progressive visual impairment in combination with partial or total hearing loss and, in some cases, abnormalities of the vestibular system.

One of the first loci identified for autosomal dominant retinitis pigmentosa (adRP) was the rhodopsin encoding gene (*RHO*) (Dryja et al., 1990). Up until now more than 150 disease-causing mutations have been identified within this gene and, collectively, they represent the most common cause for adRP by far (Athanasίου et al., 2018). Within this locus, the c.620T>G transversion located in *RHO* exon 3, which is causing a severe, early-onset form of adRP, has been identified early on (Farrar et al., 1992). Assuming that c.620T>G is a missense mutation replacing the methionine at position 207 with an arginine (Met207Arg), several studies addressed the importance of this residue for rhodopsin function and therapy (Padron-Garcia et al.,

2004, Hagemann et al., 2006, Aguila et al., 2009, Audo et al., 2010a, Rakoczy et al., 2011, Maubaret et al., 2012, Zhou et al., 2012, Hernandez-Rodriguez et al., 2012, Hernandez-Rodriguez et al., 2013). Met207 has been identified in other GPCRs as a part of the ligand binding pocket and was found to be critical for binding of 11-*cis*-retinal (Figure 3) (Ahuja et al., 2009, Moitra et al., 2012, Kimata et al., 2016). Two decades later, another adRP mutation affecting the same nucleotide, i.e. the c.620T>A transversion classified as Met207Lys, has been identified (Audo et al., 2010a, Audo et al., 2010b). Intriguingly, this mutation, despite supposedly altering the same amino acid as c.620T>G, results in a mild, late onset disease variant. These two mutations strikingly illustrate the adRP phenotype variability and highlight the importance for elucidating the underlying molecular mechanisms.

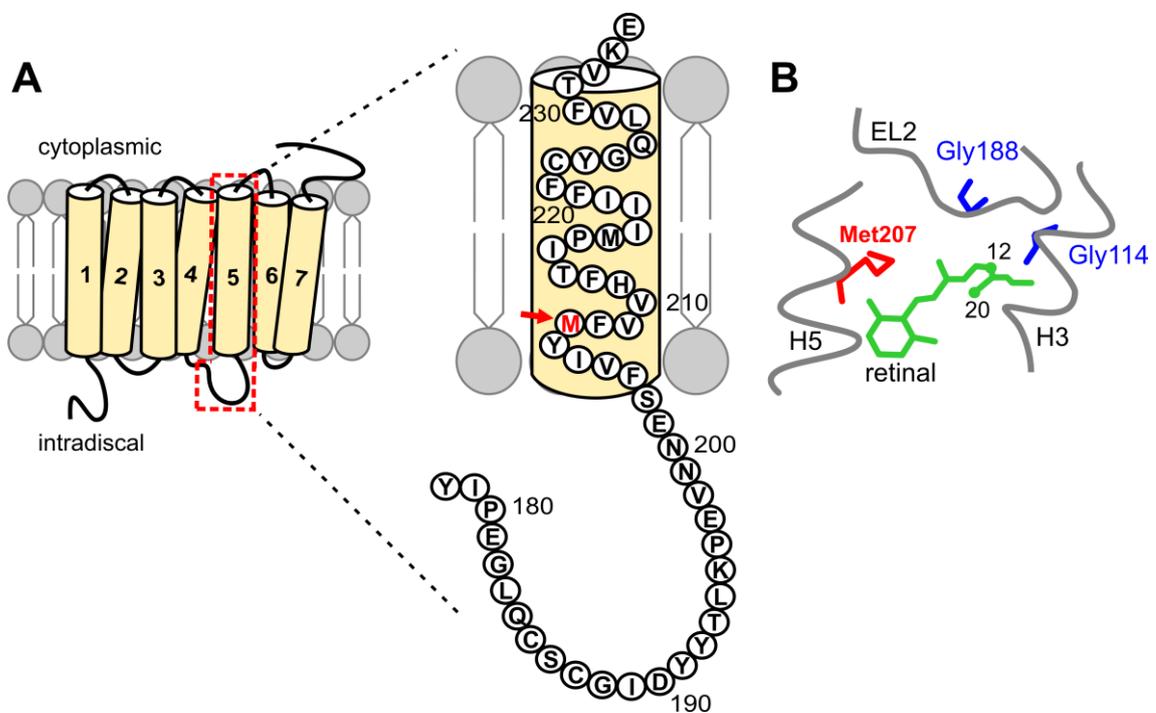


Figure 3 Position and function of Met207 in rhodopsin

A, Rhodopsin topology. Left panel, Schematic illustration of the molecular structure of rhodopsin. The region encoded by *RHO* exon 3, highlighted by a red dashed rectangle, comprises the second extracellular loop (EL2) and the transmembrane helix domain 5 (H5). Right panel, Magnification of the exon 3-encoded region. The position of methionine 207 (Met207) is indicated by a red arrow. B, Stick model showing the 11-*cis*-retinal binding pocket. The glycines at position 114 (Gly114) and 188 (Gly188) as well as Met207 play a crucial role. H3: transmembrane helix domain 3. B was modified from (Ahuja et al., 2009).

1.2 SPLICING – THE ART OF REMOVING INTRONS

1.2.1 The Splicing Mechanism

Most eukaryotic genes contain non-coding introns, which must be removed from the precursor messenger RNA (pre-mRNA) to generate a translatable mature messenger RNA (mRNA) in a process called ‘splicing’. Pre-mRNA splicing has been discovered more than 40 years ago and has since then been an important field of research (Berget et al., 1977). The process has mostly been investigated in *Saccharomyces cerevisiae* and later in *Schizosaccharomyces pombe*, which resembles the complex splicing mechanism found in human cells more closely (Hossain and Johnson, 2014, Stepankiw et al., 2015). Splicing is mediated via a large ribonucleoprotein complex, the spliceosome, which consists of five conserved small nuclear ribonucleoproteins (snRNPs), i.e. U1, U2, U4, U5 and U6 snRNP, and more than 300 proteins. The spliceosome assembles and disassembles for each intron in a highly dynamic process. For this purpose, specific splicing sequences need to be recognized: the 5’ splice donor site (SDS) marking the end of the 5’ exon and the beginning of the intron and the 3’ splice acceptor site (SAS) indicating the end of the intron and the beginning of the 3’ exon. A functional SAS requires three distinct elements, which are usually located within a range of approximately 50 bp (Gao et al., 2008): a branch point (BP) sequence, a polypyrimidine tract (PPT) and the splice acceptor site. The canonical human consensus sequences for each element can be found in Table 1 (Zhang, 1998, Lim and Burge, 2001, Gao et al., 2008).

Table 1 Canonical consensus sequences of human splice elements

Splice element	Consensus sequence
Splice donor site	mAG GUrAGU
Branch point	yUn A y
Polypyrimidine tract and splice acceptor site	(y) ₁₁ nCAG G

y = pyrimidine, r = purine, n = nucleotide, m = adenine or cytosine; | marks the junction between an exon and an intron; the branch point nucleotide is depicted as a bold letter

Each splicing cycle consists of two transesterification reactions (Figure 4) (Shi, 2017). For the first reaction, known as branching, the branch point adenosine attacks the phosphate bond at the 5' exon-intron junction. This results in the formation of an intronic lariat-3' exon intermediate and a free 5' exon end. In the second reaction, called exon ligation, the free end of the 5' exon attacks the phosphate at the intron-3' exon junction causing a ligation of the two exons and a release of the intron lariat structure.

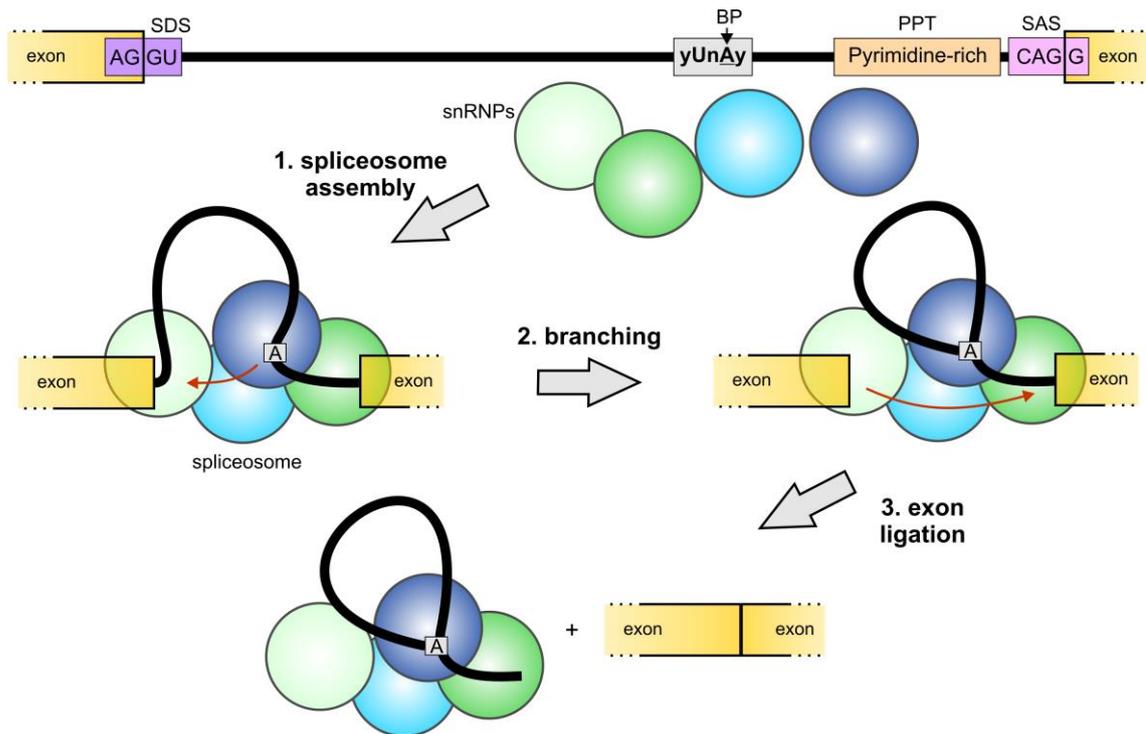


Figure 4 Illustration of the splicing mechanism

For splicing, several sequence elements are required, i.e. a splice donor site (SDS), a branch point (BP), a polypyrimidine tract (PPT) and a splice acceptor site (SAS). Upon recognition of these splicing elements, snRNPs assemble to form a functional spliceosome. Next, the branch point adenosine attacks the end of the 5' exon and forms a lariat in a process called branching. The free OH group of the 5' exon in turn attacks the 3' exon thereby facilitating exon ligation and excision of the lariat.

Moreover, the splicing reaction is regulated by auxiliary *cis*-acting splicing regulatory elements composed of up to ten nucleotides. According to their function and position, they can be categorized into exonic splice enhancers (ESEs), exonic splice silencers (ESSs), intronic splice enhancers (ISEs) and intronic splice silencers (ISSs) (Wang et al., 2012). These elements are able to recruit *trans*-acting proteins

for promotion or prevention of exon inclusion and are therefore proposed to play an important role in alternative splicing (Wang et al., 2006).

1.2.2 Mutations Affecting mRNA Splicing

Disease-causing single nucleotide polymorphisms (SNPs) are commonly assumed to be missense or nonsense mutations located in coding regions and leading to single amino acid exchanges or premature stop codons, respectively. However, SNPs can also lie in exon-intron boundaries thereby affecting pre-mRNA splicing by abolishing existing *cis*-acting splicing elements (Figure 5) (Cartegni et al., 2002, Abramowicz and Gos, 2018). For example, SNPs can eliminate splice donor or acceptor sites, which often results in exon skipping. Such a mutation has been identified in the *MIP* gene causing autosomal dominant congenital cataract (Abramowicz and Gos, 2018).

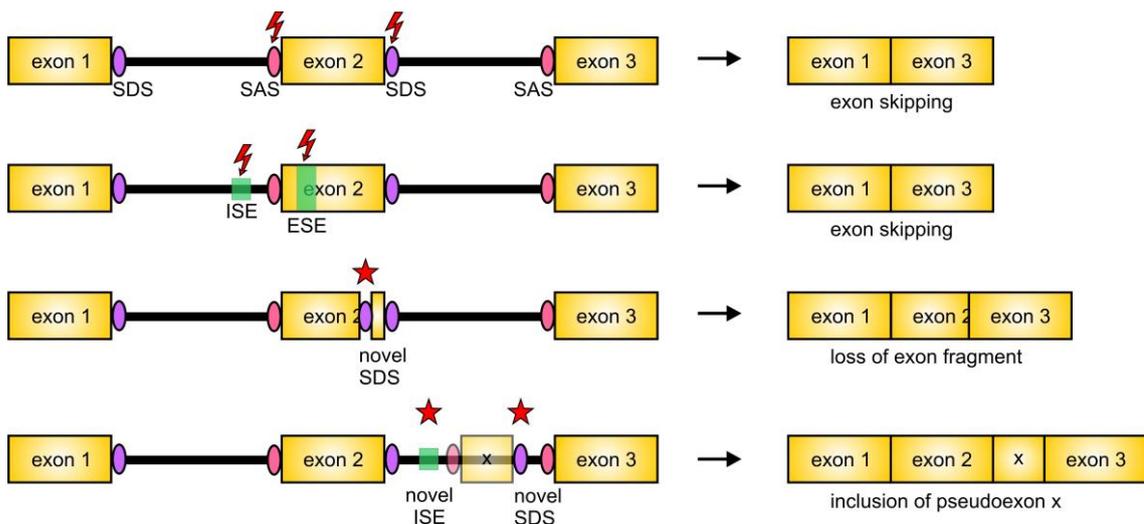


Figure 5 Single nucleotide polymorphisms affecting splicing

Left panel, Positions of single nucleotide polymorphisms (SNPs) abolishing existing (red lightning) or creating novel (red star) *cis*-acting splicing elements. SAS, splice acceptor site (pink oval). SDS, splice donor site (purple oval). ISE, intronic splice enhancer (green box). ESE, exonic splice enhancer (green box). Right panel, possible effect of the respective SNPs on pre-mRNA splicing.

Moreover, SNPs can disrupt existing exonic or intronic splice enhancers or silencers, which may lead to exon skipping, activation of cryptic splice sites or alterations in the frequency of specific splice isoforms. Such mutations have been identified for medium-chain acyl-CoA dehydrogenase deficiency in the *ACADM* gene, in which a

SNP is leading to exon 5 skipping, and in frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17), in which the frequency of different *MAPT* splice isoforms is disturbed (Nielsen et al., 2007, Niblock and Gallo, 2012). Another mechanism by which SNPs influence splicing is the creation of novel splice sites within coding sequences leading to loss of exon fragments and possibly to frame-shifts (Figure 5). An example is a mutation in the *PRPH2* gene creating an alternative SDS and a 90 nucleotide in-frame deletion in exon 2 (Becirovic et al., 2016b). Finally, splicing mutations can create novel splice sites and regulatory elements deep within intronic regions (Scotti and Swanson, 2016) like in *CLRN1*, where an intronic Usher syndrome-causing SNP has been identified, which creates a novel SDS leading to the inclusion of a pseudoexon and a premature stop codon (Khan et al., 2017).

1.2.3 mRNA *Trans*-Splicing

Apart from the regular *cis*-splicing events described in 1.2.1, splicing can also occur in *trans* thereby combining two separate pre-mRNA molecules to create a non-co-linear chimeric RNA (Lei et al., 2016). This process was first discovered in trypanosomes, in which a short leader was spliced to the 5' end of a pre-mRNA (Van der Ploeg et al., 1982). Since then, *trans*-splicing events have been identified in many more species including *Caenorhabditis elegans*, *Mus musculus* and *Homo sapiens* (Krause and Hirsh, 1987, Hirano and Noda, 2004, Chuang et al., 2018).

Depending on the nature of the two pre-mRNA molecules involved in the process, *trans*-splicing events can be categorized into three types: spliced leader (SL) *trans*-splicing, intragenic *trans*-splicing and intergenic *trans*-splicing (Figure 6). SL *trans*-splicing is a mechanism occurring at high frequencies, however, it is mainly present in lower eukaryotes such as nematodes (Hastings, 2005). Intragenic and intergenic *trans*-splicing events combine two pre-mRNA molecules originating from the same or from different genes, respectively. They are mainly present in non-SL species like insects or vertebrates and occur at much lower rates (Horiuchi et al., 2003, Li et al., 2008a). In vertebrates, *trans*-spliced chimeric mRNAs are proposed to serve several physiological functions as protein coding or noncoding RNAs. For example, they are thought to play a role in gene expression regulation and pluripotency maintenance (Ni et al., 2011, Wu et al., 2014).

The *trans*-splicing process is expected to share most characteristics with *cis*-splicing. For instance, the same splice sites can be used for *trans*- as well as *cis*-splicing and splicing factors known to promote alternative splicing can also enhance *trans*-splicing frequency (Shao et al., 2012, Chuang et al., 2018). The *trans*-splicing mechanism has already been applied successfully in two independent treatment strategies for inherited diseases. The first approach is centered around a group I intron-based ribozyme, which is used to replace defective 3' sequences of a pre-mRNA with a contained non-mutated 3' exon tag via *trans*-splicing (Byun et al., 2003). For the second approach, called spliceosome-mediated mRNA *trans*-splicing (SMaRT), an exogenous pre-mRNA *trans*-splicing molecule (PTM) is introduced into the target cell and replaces a part of the mutated endogenous pre-mRNA (Dooley et al., 2018).

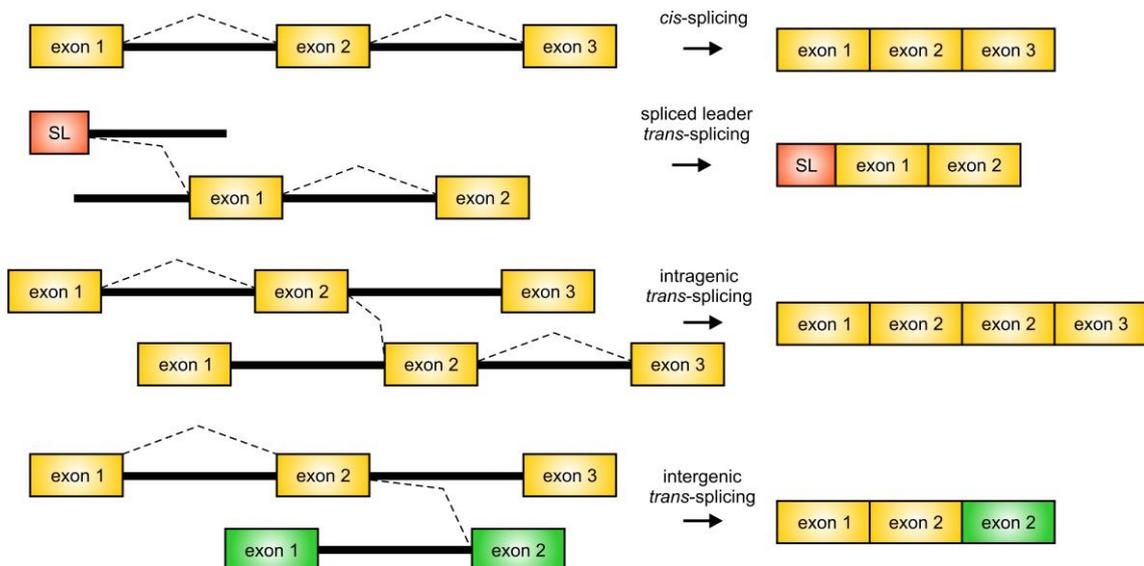


Figure 6 Naturally occurring trans-splicing events

Shown are *cis*-splicing as well as spliced leader (SL), intragenic and intergenic *trans*-splicing events. Dashed lines represent splicing events. Mature mRNA is shown on the right side.

1.3 GENE THERAPY – A WEAPON AGAINST GENETIC DISEASES

1.3.1 A Brief History of Gene Therapy

Nearly 50 years ago, gene therapy has for the first time been proposed as a treatment for human genetic diseases (Friedmann and Roblin, 1972). Its fundamental principle is the delivery of exogenous DNA into a diseased target cell

for genetic modification. In contrast to protein-based drugs, which require multiple infusions, gene therapy offers the advantage of a sustained protein expression in long-lived cells (Dunbar et al., 2018). First approaches mainly focused on the alleviation or correction of a genetic defect by delivering a wild type version of the mutated gene into the target cell. For this purpose, efficient and non-toxic gene delivery vectors were required. Pioneers in the field recognized the potential of mammalian viruses for efficient gene transfer and soon replication-deficient γ -retroviral vectors were used for gene delivery into hematopoietic stem cells (HSCs) (Mann et al., 1983, Riviere et al., 1995). For a more efficient and controlled transduction, cells have been isolated and treated *ex vivo* followed by an engraftment back into the patient. However, due to their inability of transducing non-dividing cells and their increased risk of oncogenic mutagenesis, γ -retroviral vectors have nowadays been largely replaced by the less genotoxic lentiviruses for *ex vivo* gene therapy (Montini et al., 2009, Ribeil et al., 2017).

Targeting organs more directly via *in vivo* delivery thereby avoiding the exertive and delicate collection, culture and transplantation of target cells has always been an attractive strategy. However, *in vivo* approaches depend on a safe and efficient delivery to the target organ and a tissue-specific gene expression. These requirements were met by the adeno-associated virus (AAV), a non-enveloped member of the *Parvoviridae* family. It consists of an icosahedral capsid containing single-stranded DNA and is natively replication-deficient (Kotterman et al., 2015). In the 1990s, AAVs have been shown to efficiently transduce skeletal muscle cells and yield a long-term transgene expression (Kessler et al., 1996).

Building on this work, the ability of different naturally occurring AAV serotypes to specifically and efficiently transduce a variety of tissues including muscle, liver, brain and retinal cells has been discovered and improved by using artificially optimized capsids (Zhong et al., 2008). Another advantageous feature of AAVs was the fact that their DNA was predominantly non-integrating and instead existing as an episome in the target cell, which substantially decreases the risk of malignant transformation due to integration. A leap in the field has been achieved, when large scale recombinant AAV (rAAV) production protocols have been established further allowing preclinical studies in large animals and clinical trials (Xiao et al., 1998). One

of the biggest drawbacks of rAAV based gene therapy, however, is their low genome packaging capacity of about 4.7 kb precluding them from treatment of defective genes with large coding sequences (Wu et al., 2010). Nevertheless, rAAV-mediated gene delivery is widely used for *in vivo* gene therapy and has already been applied successfully for treatment of several genetic diseases like for hemophilia B and for spinal muscular atrophy (Rangarajan et al., 2017, Mendell et al., 2017).

In recent years, new possibilities have opened up in the field of gene therapy since the discovery of the universal applicability of the CRISPR-Cas system for targeted genome editing in 2012 (Jinek et al., 2012). With this technology, a Cas9 DNA endonuclease can be targeted to a specific locus in the genome using a programmable guide RNA (gRNA) and induce a double strand break *in situ*. When an additional exogenous donor DNA repair template is provided, homology-directed repair (HDR) can be induced offering the possibility to exchange and correct mutated regions of a gene (Mali et al., 2013). This prospect is especially important for treatment of dominant negative mutations, which need to be eliminated to achieve sufficient rescue. However, HDR is an event of noticeably low efficiency and is virtually absent in postmitotic cells (Lino et al., 2018).

Lately, another promising strategy based on CRISPR-Cas9 technology has been developed. The catalytically inactive “dead” Cas9 (dCas9) has been fused to various effector domains like adenine base editors or transcriptional regulators to modulate the human genome (Chavez et al., 2015, Yeo et al., 2018, Ryu et al., 2018). These novel techniques could be used to specifically modify single base pairs or to artificially switch genes off and on, which holds the potential to revolutionize the field of gene therapy once more. However, an application of these approaches *in vivo* still faces obstacles, as the coding sequences of the Cas9 fusion proteins evidently exceed the packaging limit of AAVs. Moreover, despite these new technologies offering exciting opportunities for human gene therapy, it needs to be stressed that CRISPR-Cas9-based approaches are still in their translational infancy and need to be thoroughly tested.

1.3.2 Recombinant Adeno-Associated Virus-Mediated Gene Supplementation in the Retina

Currently, rAAV-mediated gene supplementation is the most successful and most widely used gene therapy approach. For this treatment strategy, an entire gene expression cassette containing a promoter, a healthy copy of the mutated gene of interest and a polyadenylation signal (pA) flanked by inverted terminal repeats (ITR) is packaged into a rAAV and delivered into the target cell (Figure 7) (Boye et al., 2013). Subsequently, the wild type gene will be expressed, which renders this method particularly suitable for therapy of loss-of-function mutations.

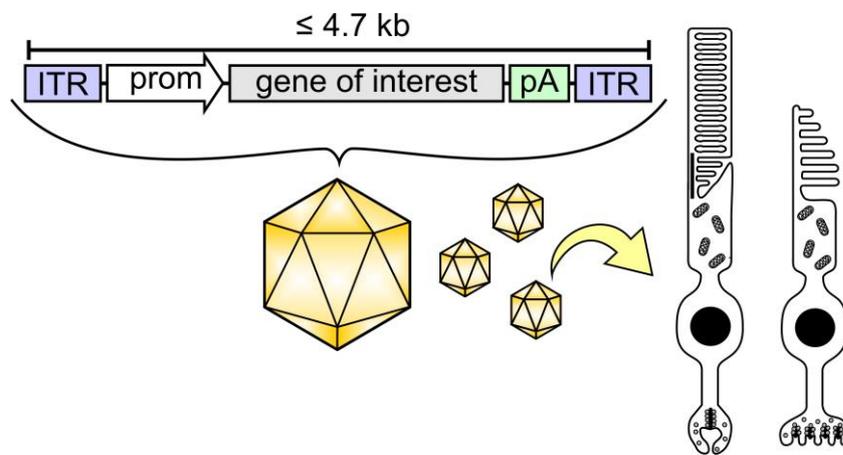


Figure 7 rAAV-mediated gene supplementation therapy

For gene supplementation therapy, the gene of interest including all regulatory elements, i.e. a promoter (prom) and a polyadenylation signal (pA), flanked by inverted terminal repeats (ITR) is packaged into an AAV and delivered into the target cell, e.g. rod or cone photoreceptors. The size of the packaged sequence should not exceed 4.7 kb.

There are several characteristics that make the eye an attractive target for gene therapy. First, the organ is small and highly compartmentalized, which allows for a local delivery of only a small vector dose (Smith et al., 2012). Second, the eye is immune privileged due to its anatomical structure, the presence of a blood-retinal-barrier and the ocular microenvironment (Willett and Bennett, 2013, Trapani and Auricchio, 2018). Consequently, the risk of systemic vector dissemination and therefore the threat of immune responses against the vector or the transgene is decidedly reduced. Third, the cell population in the retina largely consists of post-

mitotic cells, which allows for a stable and long-term transgene expression from non-integrating vectors after successful transduction.

As most IRDs are caused by gene defects affecting photoreceptors or RPE cells (c.f. 1.1.3), they represent the main target cell population for rAAV-mediated gene supplementation in the retina. To sufficiently transduce these cell types, two delivery routes are primarily used. Subretinal injections have already been shown early on to result in an efficient transduction, as they deliver the virus directly into proximity of RPE cells and photoreceptors (Ali et al., 1996). This transduction efficiency could even be improved by using pseudotyped viruses containing the AAV 2 genome packaged into capsids of other serotypes, e.g. of AAV5 or AAV8, and by mutating specific tyrosine residues within the capsid (Natkunarahah et al., 2008, Petrs-Silva et al., 2011). Such variants like the AAV2/8 Y733F vector can efficiently transduce photoreceptors throughout the injected area after subretinal delivery (Koch et al., 2012). Intravitreal injection, the second route of administration, is considered to be technically less challenging and less invasive, as it does not rely on detachment of the retina from the RPE (Ochakovski et al., 2017). However, the virus needs to cross several retinal layers in order to infect photoreceptors and the highly concentrated vectors are diluted within the vitreous cavity (Igarashi et al., 2013). Therefore, AAV capsid variants mediating a highly efficient transduction while still possessing the ability to diffuse across the retinal layers are needed for this approach.

rAAV-mediated gene supplementation therapy has been shown to successfully correct IRD phenotypes in various disease models, ranging from *Gucy2e*^{-/-} mice, a model for the severe IRD Leber congenital amaurosis (LCA), to CNGB1-deficient dogs displaying a RP-like phenotype (Mihelec et al., 2011, Petersen-Jones et al., 2018). Building on some of this work, the first gene supplementation therapy product, a rAAV delivering a wild type copy of *RPE65*, has been approved by the FDA for treatment of LCA in December 2017 (NatureBiotechnologyNews, 2018). By being successfully translated into an approved therapy product and by achieving proof-of-principle in more than 20 different IRD animal models, rAAV-mediated gene supplementation has substantially advanced the field of retinal gene therapy and remains the gold standard for treatment of IRDs (Smith et al., 2012).

1.3.3 rAAV Dual Vector Strategies

Despite its success story, rAAV-mediated gene supplementation is not suitable for treatment of all inherited diseases caused by loss-of-function mutations yet. As mentioned above, one major limitation of this approach is the AAV's restricted packaging capacity (Wu et al., 2010). Therefore, diseases caused by mutated genes with coding sequences (CDSs) larger than 4 kb are currently precluded from treatment (McClements and MacLaren, 2017). However, many inherited diseases including hemophilia A, Duchenne muscular dystrophy (DMD) and IRDs can be caused by loss-of-function mutations in large genes (Oldenburg and El-Maarri, 2006, Goldstein and McNally, 2010, Tanna et al., 2017). For instance, a selection of IRD-causing genes exceeding the packaging limit and consequently precluded from gene supplementation therapy can be found in Table 2 (Kersey et al., 2018, RetNet, 2019). Patients with mutations in one of these genes would therefore benefit greatly from novel gene therapeutic treatment strategies overcoming these restrictions.

Table 2 IRD-causing genes exceeding the AAV packaging limit

Gene	CDS Length	Associated IRD
<i>ABCA4</i>	6.8 kb	Stargardt macular dystrophy, arRP
<i>ADGRV1</i>	19 kb	Usher syndrome
<i>CACNA1F</i>	5.9 kb	Congenital stationary night blindness
<i>CDH23</i>	10.1 kb	Usher syndrome
<i>CEP290</i>	7.4 kb	Leber congenital amaurosis
<i>CRB1</i>	4.2 kb	arRP, Leber congenital amaurosis
<i>EYS</i>	9.4 kb	arRP
<i>KIAA1549</i>	5.9 kb	arRP
<i>MYO7A</i>	6.6 kb	Usher syndrome
<i>PCDH15</i>	5.9 kb	Usher syndrome
<i>RP1</i>	6.5 kb	adRP, arRP
<i>SNRNP200</i>	6.4 kb	adRP
<i>USH2A</i>	15.6 kb	Usher syndrome, arRP

One option to circumvent this limitation would be to utilize other vectors such as lentiviruses, which possess a higher cargo capacity of about 8.5 kb while still yielding moderate titers (Kumar et al., 2001). However, these vectors show less ability to selectively transduce a subset of target cells when compared to AAVs (Peng et al., 2001). Moreover, they are less efficient in transducing the post-mitotic cells of the central nervous system and, due to their larger size, are less capable of diffusing through multi-layered structures such as the retina (Lipinski et al., 2014). Given these restrictions combined with the increased potential of lentiviruses to induce host immune responses and their mutagenic potential due to their integration into the host genome, it is considered worthwhile to develop suitable rAAV systems to accommodate the need for larger cargo size.

Several groups have attempted to enforce packaging of genes larger than 4.7 kb into an AAV capsid thereby creating so-called oversized vectors (Wu et al., 2007, Allocca et al., 2008). However, this results in the packaging of fragmented transgenes, which will contain varying overlapping regions and by chance undergo homologous recombination in the target cell (Dong et al., 2010, Hirsch et al., 2010). Due to the heterogeneity of the packaged genomes, which impedes a definite characterization of the vectors, further progress of this approach towards a clinical application is rather unlikely (McClements and MacLaren, 2017). As a more attractive alternative, rAAV dual vector technologies have been developed, for which the CDS of a transgene is split into at least two parts and packaged into two or more separate rAAV vectors. After co-transduction of the target cell by the split genome vectors, the full-length CDS is reconstituted. Since AAVs show a high co-transduction efficiency of > 90 % in the retina, the efficient delivery of both halves into the same target cell is not expected to represent a limiting issue (Palfi et al., 2012, Becirovic et al., 2016a, Bohm et al., 2017). An efficient reconstitution of the two transgene halves, however, remains challenging. For several years since the development of rAAV dual vector systems, reconstitution has only been addressed at the DNA level and several strategies have been explored to improve this approach (Figure 8) (McClements and MacLaren, 2017).

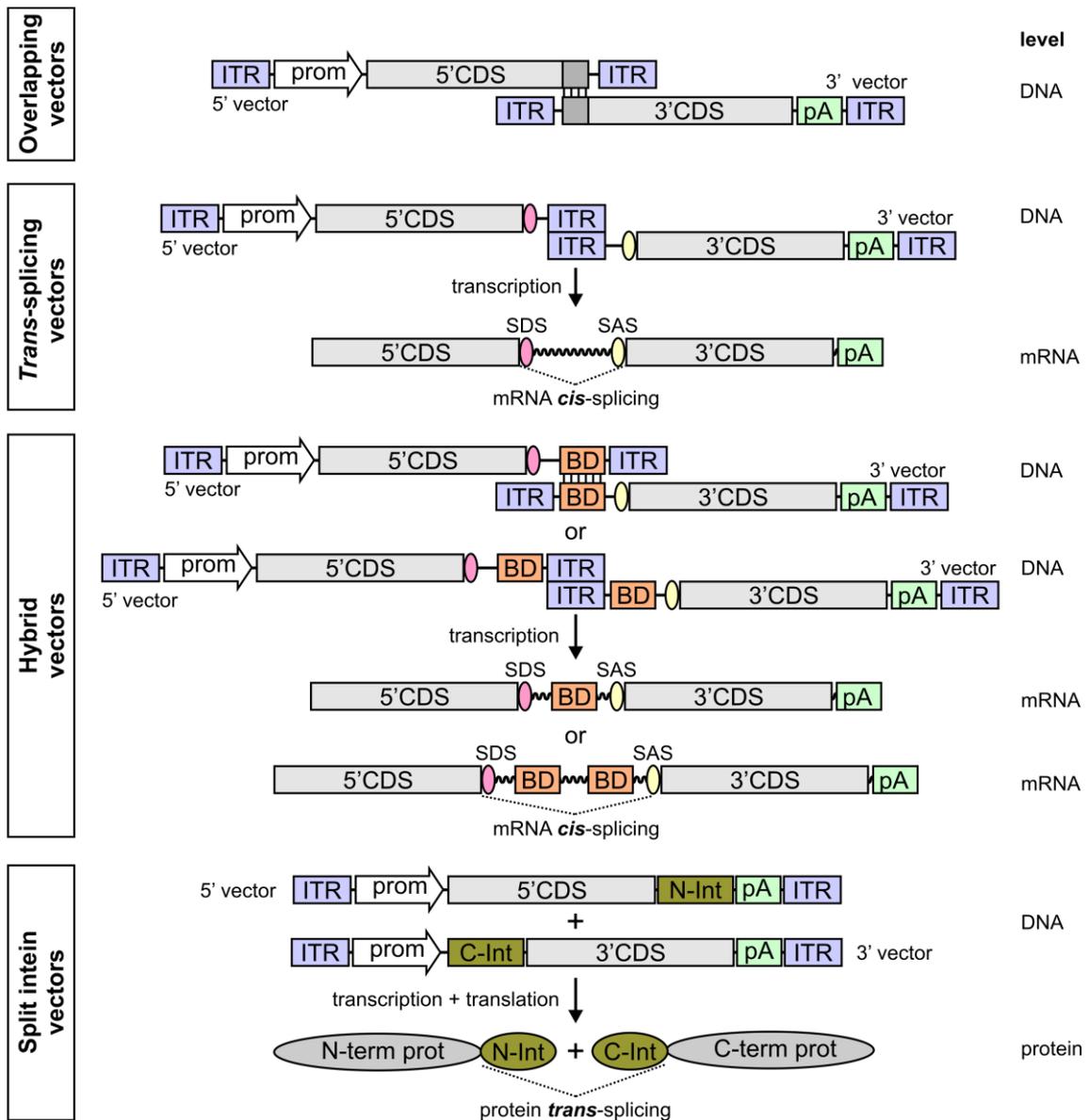


Figure 8 Schematic depiction of different rAAV dual vector strategies

Illustration of the different reconstitution strategies of the currently employed rAAV dual vector approaches. Gene of interest coding sequences are split into two halves (5' CDS and 3' CDS) and packaged into separate vectors (5' vector and 3' vector). For overlapping vectors, homologous recombination is facilitated via an overlapping sequence part. For *trans*-splicing vectors, genomic reconstitution is enabled via concatemerization. Remaining ITR sequences are removed at the mRNA level via *cis*-splicing mediated by a splice donor (SDS) and acceptor site (SAS). Hybrid vectors combine the two strategies described above by reconstituting via a recombinogenic binding domain (BD) or concatemerization. Remaining ITR and/or BD sequences are removed at mRNA level via *cis*-splicing. For split intein vectors, both halves are transcribed and translated separately. Reconstitution is facilitated at the protein level via split intein mediated protein *trans*-splicing. ITR, inverted terminal repeats. prom, promoter. pA, polyadenylation signal. N/C-term prot, N/C-terminal protein half. N/C-Int, N/C-terminal part of the split intein.

One of these strategies is referred to as the overlapping approach. For this strategy, the two defined transgene halves contain a homologous overlapping region at the 3' end of the 5' half and at the 5' end of the 3' half, which initiates homologous recombination and thereby facilitates reconstitution of the full-length transgene (Halbert et al., 2002, Lopes et al., 2013). Here, the 5'-vector contains the 5' half of the transgene driven by an appropriate promoter. The 3'-vector lacks any promoter, but contains the 3' half of the transgene and a pA signal for mRNA stabilization and proper expression of the reconstituted protein. Another strategy called *trans-splicing* approach does not possess a region of overlap, but instead is based on concatemerization, i.e. the formation of stable episomal structures by joining the flanking ITRs. This process can occur intramolecularly between the flanking ITRs of one episome as well as intermolecularly between different episomes (Chen et al., 2001). Following this concatemerization, the ITR sequence still resides in between the two transgene halves and would therefore disrupt the open reading frame. To overcome this problem, a splice donor site (SDS) is placed at the 3' end of the 5' CDS and a splice acceptor site (SAS) is placed at the 5' end of the 3' CDS. After transcription, the full-length transgene will still contain (partial) sequences of the ITRs flanking the two CDSs, however, due to the presence of the splicing elements, this sequence will be removed from the mRNA via regular *cis*-splicing (Reich et al., 2003, Lai et al., 2005). Therefore, although this strategy has been termed '*trans-splicing*' approach, mRNA splicing never actually takes place in *trans*. The last strategy based on genomic reconstitution is a fusion of the overlapping and the *trans-splicing* approach termed 'hybrid approach'. For this purpose, a designated recombinogenic sequence, e.g. derived from the alkaline phosphatase or the F1-phage, is employed. Moreover, the two transgene halves contain appropriate splice sites to remove the recombinogenic sequence from the mRNA after successful recombination (Ghosh et al., 2008, Trapani et al., 2015).

Several studies set out to determine the differential efficiencies of the three strategies based on genomic reconstitution. All approaches have been reported to result in transgene reconstitution and expression *in vivo* and thereby mitigate disease phenotypes (Lai et al., 2005, Trapani et al., 2014, McClements et al., 2019). However, when comparing the strategies directly, contradictory results have been obtained. For instance, Duan et al. demonstrated that the *trans-splicing* approach

yields a more efficient reconstitution than the overlapping vectors, whereas Carvalho et al. showed that the overlapping approach dominates the *trans*-splicing strategy (Duan et al., 2001, Carvalho et al., 2017). Nevertheless, most studies concluded that the hybrid approach results in the best reconstitution efficiencies (Dyka et al., 2014, Colella et al., 2014, Carvalho et al., 2017).

Still, all three approaches have limitations impeding an efficient reconstitution. The overlapping strategy, for instance, is most likely depending on reconstitution via homologous recombination – an event, which is rarely taking place in post-mitotic cells such as photoreceptors (Duan et al., 2001, Halbert et al., 2002, Fishel et al., 2007). Accordingly, many studies have reported inefficient expression of the full-length transgene in photoreceptors when using this approach (Lopes et al., 2013, Trapani et al., 2014, Carvalho et al., 2017). The *trans*-splicing strategy, on the other hand, is based on reconstitution via ITR-mediated concatemerization. However, in this process, self-circularization is preferred over intermolecular interaction (Duan et al., 1998, Choi et al., 2005). Moreover, the two transgene halves can concatemerize incorrectly, i.e. 5' half – 5' half, 3' half – 3' half or 3' half – 5' half, instead of forming the correct head-to-tail 5' half – 3' half concatemer (Schnepp et al., 2005). To overcome this issue, heterologous ITRs have been developed for an orientation-directed concatemerization. However, this strategy seems to reduce the production yield of the vectors (Trapani et al., 2015). As the hybrid approach doubles the likelihood of successful reconstitution by combining both strategies, it also achieves the highest transgene expression levels. Still, the success rates obtained for the DNA-based reconstitution systems are varying depending on the targeted cell type, the delivered transgene or the transduction efficiency of the AAV serotype (Duan et al., 2001, Colella et al., 2014). Moreover, when the reconstitution efficiencies of all approaches are estimated in a quantitative, systematic experimental design, even the best approaches achieve only 17 % of the full-length transgene expression *in vitro* and 10 % *in vivo* (Carvalho et al., 2017). Although higher reconstitution efficiencies have been reported when using serotypes for ideal transduction in optimized model systems (Colella et al., 2014), the overall reported transgene expression levels remain disappointingly low (Duan et al., 2001, Trapani et al., 2014, Carvalho et al., 2017).

Recently, another option for transgene reconstitution has been explored: split intein-mediated protein *trans*-splicing for reconstitution at the protein level (Figure 8). Inteins are small internal protein elements found in microorganisms like bacteria, viruses or yeast. They are capable of excising themselves post-translationally from their host proteins, analogous to intron splicing in the pre-mRNA (Novikova et al., 2014). A subset of this group, called split inteins, possesses the ability to mediate the scarless ligation of two host proteins via protein *trans*-splicing (Wu et al., 1998). For this, the N-terminal part of the intein (N-intein) is located on the C-terminal end of one protein, while the C-terminal part of the intein (C-intein) is positioned at the N-terminal end of the other protein. In 2008, split inteins have already been applied successfully for gene supplementation therapy of large genes (Li et al., 2008b). In this study, sequences encoding the cyanobacterial DnaB N-intein or C-intein have been fused to a split dystrophin cDNA. Both transgene halves were transcribed and translated separately, reconstituted at the protein level via protein *trans*-splicing and resulted in expression of dystrophin in a DMD mouse model. Recently, this method has been optimized further and successfully applied for treatment of IRDs yielding high transgene expression levels and outperforming DNA-based reconstitution systems (Tornabene et al., 2019). Moreover, this strategy has been used to deliver large fusion proteins such as the transcriptional activator Cas9-VPR into target cells *in vivo* (Chew et al., 2016).

However, this system again comes with limitations. For split intein-mediated protein *trans*-splicing, the split site must contain either a cysteine, serine or threonine adjacent to the C-intein limiting the split site options. Moreover, the second amino acid residue adjacent to the split site also seems to influence the efficiency of the protein *trans*-splicing reaction (Stevens et al., 2017). Another important requirement for the selection of a suitable split site is that it is localized outside of structural domains to avoid improper protein folding. Consequently, several different split sites need to be screened for each transgene (Tornabene et al., 2019) and it seems very likely that for many proteins, such split sites are not present in the central region of the respective coding sequences. Moreover, split inteins yielding the highest reconstitution efficiencies are derived from bacterial proteins, i.e. a subunit of DNA Polymerase III (DnaE) or a DNA helicase (DnaB) found in *Nostoc punctiforme* or *Rhodothermus marinus*, respectively (Stevens et al., 2016, Chew et al.,

2016). As transgene reconstitution via protein *trans*-splicing results in an equimolar production of inteins, this approach requires a thorough investigation regarding its potential to elicit host immune responses or to cause detrimental long-term toxic effects, e.g. due to accumulation in the target cell. In conclusion, despite recent advancement in the field of rAAV dual vector technology, there is still an unmet need for efficient and safe reconstitution of large genes.

2 AIM OF THE STUDY

Disease-associated mutations can act at the mRNA or at the protein level. Mutations located in the protein-coding exonic regions are generally classified as missense mutations expected to alter single amino acids. However, exonic mutations can potentially also affect mRNA splicing, which indicates that in the past a substantial portion of disease-associated mutations might have been misclassified. In this study, the hypothesis that differential phenotypes caused by similar mutations in the same gene originate from differential impact on mRNA splicing should be investigated (1). Moreover, apart from its role in pathogenicity, mRNA splicing is playing an integrative part in some important biotechnological applications. This particularly applies to the process called mRNA *trans*-splicing, which is enabling a seamless ligation of two separate mRNA molecules. In this study, it should be examined, whether mRNA *trans*-splicing could be used to overcome the limited genome packaging capacity of AAV vectors in a novel biotechnological approach (2).

- (1) The hypothesis should be evaluated using the example of two adRP mutations in the rhodopsin gene which affect the same nucleotide but are linked to remarkably different phenotypes. The c.620T>G mutation is associated with a severe, early-onset RP, whereas c.620T>A causes a mild, late-onset disease phenotype. The differential impact of both mutations should be tested at the mRNA and at the protein level *in vitro* and *in vivo*.
- (2) The limited genome packaging capacity of AAV vectors represents the most relevant limitation hampering their broad use for gene therapy. Current approaches addressing this issue focus on dual AAV vector strategies aiming at reconstitution of split genes at the genomic or at the protein level. However, these approaches harbor important disadvantages which hinder their broad usage in therapeutically relevant contexts. In this study, a novel dual vector technology should be developed aiming at a reconstitution at the mRNA level using mRNA *trans*-splicing. For this purpose, a split fluorophore assay will be established for testing and optimization *in vitro*. If successful, this method should be tested *in vivo* in the mouse retina and a proof of principle should be provided for the reconstitution of an important sequence exceeding the AAV packaging limit, i.e. the *SpCas9*-VPR module.

3 MATERIALS & METHODS

3.1 MATERIALS

All chemicals used in this study hold a quality suitable for molecular biological applications and were obtained from Sigma-Aldrich, VWR, Roth, Merck or Fluka unless stated otherwise. All cell culture media and supplements were acquired from Gibco™ (Thermo Fisher Scientific) unless stated otherwise. For all buffers and reactions, highly pure deionized water obtained from the Milli-Q Plus System (Millipore) purified with 18.2 MΩ x cm (25 °C) was used. All materials used for cell culture applications were obtained in cell culture grade sterile wrapping or sterilized by autoclaving (Sterilicator, Münchener Medizin Mechanik). Bacterial growth medium was autoclaved prior to use.

3.2 *IN SILICO* SPLICE PREDICTION ANALYSIS

To predict the potential effects of single point mutations on splicing, the open access software tools Human Splicing Finder (HSF, <http://www.umd.be/HSF/>) and NNSplice (Berkeley Drosophila Genome Project, http://www.fruitfly.org/seq_tools/splice.html) were used. To predict the strength of different splice acceptor sites, the HSF, NNSplice and the Alternative Splice Site Predictor (ASSP, <http://wangcomputing.com/assp/>) were used. To obtain scores for branch point sequences, the HSF was used.

3.3 CLONING

3.3.1 Expression Vectors

As expression system for *in vitro* experiments, the mammalian expression vector pcDNA3.1(+) containing a cytomegalovirus (CMV) promoter and a bovine growth hormone (bGH) polyadenylation signal (pA) was used. For *in vivo* expression, the genes of interest were cloned into a pAAV2.1 plasmid (Michalakis et al., 2010) containing a murine rhodopsin or CMV promoter, a SV40 or synthetic pA and flanking inverted terminal repeats (ITR). A list of the plasmids used in this study can be found in Table 3. The size of the transgene expression cassette within the flanking

ITRs is indicated for each pAAV2.1 plasmid. The SP-dCas9-VPR plasmid was a gift from George Church (Addgene plasmid nr. 63798; <http://n2t.net/addgene:63798>; RRID: Addgene_63798).

Table 3 List of plasmids used in this study

Plasmid
pcDNA3.1- <i>RHO</i> _WT-Citrine
pcDNA3.1- <i>RHO</i> _c.620T>G-Citrine
pcDNA3.1- <i>RHO</i> _c.620T>A-Citrine
pcDNA3.1-Citr- <i>RHO</i> _WT ^{mod} -mCherry
pcDNA3.1-Citr- <i>RHO</i> _c.620T>G ^{mod} -mCherry
pcDNA3.1- <i>RHO</i> _c.532-68_insSAS_620-Citrine
pcDNA3.1- <i>RHO</i> _c.620T>G_mut9A-Citrine
pcDNA3.1- <i>RPS27</i> _WT
pcDNA3.1- <i>RPS27</i> _ins <i>RHO</i> _(a-g)
pcDNA3.1- <i>HBQ1</i> _nat
pcDNA3.1- <i>HBQ1</i> _insSAS_620
pcDNA3.1- <i>S100A12</i> _nat
pcDNA3.1- <i>S100A12</i> _insSAS_620
pcDNA3.1- <i>CLRN1</i> _exon0b-1b_nat
pcDNA3.1- <i>CLRN1</i> _exon0b-1b_insSAS_620
pcDNA3.1- <i>CNGB1</i> _exon14-16_nat
pcDNA3.1- <i>CNGB1</i> _exon14-16_insSAS_620
pcDNA3.1-5'Cerulean_SDS_BD
pcDNA3.1-BD_SAS_3'Cerulean
pcDNA3.1-Cerulean_insSDS_intron_SAS
pAAV2.1-mRho- <i>RHO</i> _WT-myc-SV40pA (4.55 kb within ITRs)
pAAV2.1-mRho- <i>RHO</i> _c.620T>G-myc-SV40pA (4.55 kb within ITRs)
pAAV2.1-mRho- <i>RHO</i> _c.620T>A-myc-SV40pA (4.55 kb within ITRs)
pAAV2.1-CMV-Citrine-5'Cerulean_SDS_BD-synpA (1.70 kb within ITRs)
pAAV2.1-CMV-BD_SAS_3'Cerulean-mCherry-synpA (2.07 kb within ITRs)
pAAV2.1-4x-hU6-sgASCL1-CMV-dCas9N_SDS_BD-SV40pA (4.77 kb within ITRs)

pAAV2.1-CMV-BD_SAS_dCas9C_VPR-synpA (4.62 kb within ITRs)

SP-dCas9-VPR (Addgene plasmid nr. 63798)

3.3.2 Standard Cloning Techniques

For cloning, the classical restriction digest as well as the restriction enzyme independent NEBuilder cloning kit (New England Biolabs) and Gibson cloning (Gibson et al., 2009) were used. The genes of interest, i.e. the inserts to be cloned into the respective expression vectors, were either synthesized commercially (BioCat GmbH) or produced from existing plasmids via restriction digest or PCR. A detailed description for cloning and construction of minigenes has been published previously (Riedmayr et al., 2018). Moreover, small inserts up to 250 bp were synthesized via primer elongation. For this purpose, the insert sequence was split into two halves and constructed to contain a 20 bp overlapping region. The sequences were acquired from Eurofins Genomics with one sequence ordered in forward direction, the other in reverse direction. For primer elongation, the Herculase II fusion DNA polymerase (Agilent Technologies) was used in the following reaction mix:

Herculase II primer elongation reaction mix

5x Herculase II Reaction Buffer	5 µl
dNTPs (10 mM)	1 µl
Forward primer (100 µM)	1 µl
Reverse primer (100 µM)	1 µl
Herculase II DNA polymerase	0.5 µl
H ₂ O	<i>ad</i> 25 µl

For amplification of the insert from an existing plasmid via PCR, primers containing appropriate restriction sites on the respective 5' end were obtained from Eurofins Genomics. The Herculase II fusion DNA polymerase (Agilent Technologies) was used in the following reaction mix:

Herculase II PCR reaction mix

5x Herculase II Reaction Buffer	5 µl
dNTPs (10 mM)	1 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Herculase II DNA polymerase	0.5 – 1 µl
DNA template	1 – 30 ng
H ₂ O	<i>ad</i> 25 µl

The primer elongation and PCR were run in a ProFlex PCR System Cycler (Applied Biosystems, Thermo Fisher Scientific) using the cycling conditions described in Table 4. The annealing temperature was calculated using the Thermo Fisher Scientific T_m Calculator.

Table 4 Herculase II PCR and primer elongation cycling conditions

If the parameter used for the PCR differs from the one used for the primer elongation, the primer elongation parameter is indicated in brackets.

Step	Number of cycles	Temperature	Duration
Initial Denaturation	1	95 °C	2 min
Denaturation	25 – 35 (2)	95 °C	20 s
Annealing		50 °C – 72 °C	30 s
Elongation		72 °C	30 s/kb (30 s)
Final Elongation	1	72 °C	3 min (1 min)
Cooling	1	10 °C	15 min

For the classical restriction digest, the desired vector as well as the insert were incubated using the appropriate restriction enzymes (Fermentas, Thermo Fisher Scientific) according to the manufacturer's instructions. The resulting DNA fragments were separated on a 0.7 – 2 % agarose gel via gel electrophoresis run at 150 V in 1x TBE buffer.

1x TBE buffer

0.5 M EDTA, pH 8.0	20.0 ml
Boric acid	27.5 g
Tris-(hydroxymethyl) aminomethane (Tris)	54.0 g
H ₂ O	<i>ad</i> 5 l

The appropriate bands were excised and the DNA was extracted using the Qiaquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions. If necessary, the vector was dephosphorylated using the Rapid DNA Dephos & Ligation Kit (Roche) to prevent re-ligation. The same kit was used for ligation of the insert and the vector according to the provided protocol. The molar ratios of vector DNA to insert DNA used for the ligation ranged from 1:3 to 1:10. Weight to mole conversions were calculated using the open source software NEBioCalculator (New England Biolabs).

For Gibson cloning and the NEBuilder cloning kit, the desired vector was digested using restriction enzymes, the fragments were separated on an agarose gel and the DNA was purified using the Qiaquick gel extraction kit. The insert was obtained commercially (Integrated DNA Technologies) or by primer elongation. Subsequently, the vector and insert were incubated in a self-made (Gibson cloning) or commercially acquired (NEBuilder cloning kit) reaction mix containing a 5' exonuclease, a DNA polymerase and a ligase for 1 h at 50 °C. The molar ratios of vector to insert DNA ranged from 1:3 to 1:10.

The ligated plasmids were transformed into chemically competent 10-beta *Escherichia coli* K12 bacteria (New England Biolabs). 5 µl of the ligation reaction or 3 µl of the NEBuilder or Gibson cloning reaction mix were added to the bacteria, the solution was gently mixed by pipetting and chilled on ice for 30 min. Afterwards, a heat shock was applied for 30 – 45 s by placing the mixture in a Thermomixer compact (Eppendorf) at 42 °C. Subsequently, the bacteria were chilled on ice for 2 min, plated onto a LB(+) agar plate containing 100 µg/ml ampicillin and incubated either at 37 °C in an incubator (Heraeus, Thermo Fisher Scientific) overnight or at RT for two days.

<u>LB(+) medium*</u>		<u>LB(+) agar plate</u>	
Peptone	10 g	Agar	15 g
Yeast extract	5 g	LB(+) medium	<i>ad</i> 1 l
NaCl	5 g	Ampicillin*	100 mg
D-(+)-Glucose	1 g	*added after autoclaving	
H ₂ O	<i>ad</i> 1 l		

*pH adjusted to 7.2 – 7.5

For small-scale production of plasmid DNA, 5 ml LB(+) medium containing 100 µg/ml ampicillin were inoculated with the respective bacteria and incubated at 37 °C in a shaking incubator (Certomat IS, B. Braun Biotech International) at 225 rpm overnight. The plasmid DNA was extracted using an alkaline lysis-based protocol. For medium- and large-scale production of plasmid DNA, 100 ml and 200 ml LB(+) medium containing 100 µg/ml ampicillin were inoculated, respectively, and incubated at 37 °C in a shaking incubator overnight. For plasmid DNA extraction, the PureLink™ HiPure Plasmid Midiprep or Maxiprep Kit (Invitrogen, Thermo Fisher Scientific) was used, respectively. The entire gene of interest was sequenced by Eurofins Genomics prior to use.

3.3.3 Site-Directed Mutagenesis

For site-directed mutagenesis, primers spanning the site-to-be-mutated in forward and reverse direction were commercially obtained (Eurofins Genomics). Both primers contained the desired point mutation in the center of their sequence and ranged from 29 to 33 bp in size. Next, the complete plasmid was amplified thereby incorporating the mutated primer sequence into the novel plasmids. For this purpose, the KAPA HiFi HotStart ReadyMix PCR kit (KAPA Biosystems) was used in the following pipetting scheme:

<u>KAPA Quickchange PCR reaction mix</u>	
2x KAPA Hotstart ReadyMix	12.5 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Template plasmid	10 ng
H ₂ O	<i>ad</i> 25 µl

The PCR was run in a ProFlex PCR System Cycler (Applied Biosystems, Thermo Fisher Scientific) using the cycling conditions described in Table 5.

Table 5 KAPA Quickchange PCR cycling conditions

Step	Number of cycles	Temperature	Duration
Initial Denaturation	1	95 °C	3 min
Denaturation	25	98 °C	20 s
Annealing		68 °C	15 s
Elongation		72 °C	60 s/kb
Final Elongation	1	72 °C	60 s/kb
Cooling	1	10 °C	15 min

After successful amplification, the template plasmid was removed by incubating the PCR reaction mix with 1 µl FastDigest *DpnI* (Fermentas, Thermo Fisher Scientific) and 2 µl FastDigest buffer in a final volume of 30 µl at 37 °C for 2 h. The enzyme can be inactivated afterwards by heating the reaction to 80 °C for 5 min. 5 µl of the mixture were transformed into chemically competent bacteria as described in chapter 3.3.2.

3.4 CELL CULTURE AND TRANSFECTION

3.4.1 HEK293 and HEK293T Cells

Human embryonic kidney 293 (HEK293) cells (DMSZ) were maintained in DMEM medium + GlutaMAX + 1 g/l glucose + pyruvate + 10 % FBS (Biochrom) + 1 % penicillin/streptomycin (Biochrom) in a CO₂ incubator (Heraeus, Thermo Fisher Scientific) at 37 °C and 10 % CO₂. The HEK293 derived Lenti-X 293T (HEK293T) cells (Clontech, Takara) were cultured in DMEM medium + GlutaMAX + 4.5 g/l glucose + 10 % FBS + 1 % penicillin/streptomycin under the same conditions. Both cell lines were passaged two times per week at a confluence of approximately 90 %.

Transient transfections were performed using the calcium phosphate method. For this purpose, cells were seeded onto 6 cm cell culture plates. When transfecting

SpCas9-containing plasmids for western blotting, 10 cm cell culture plates were used. The cells were incubated overnight until they reached the desired confluence of approximately 70 %. The transfection mix components shown in Table 6 were added to a 15 ml Falcon tube in the indicated order. 2x BBS was added dropwise during vortexing.

Table 6 Calcium phosphate transfection mix

Culture plate diameter	6 cm	10 cm
DNA	6 µg	10 µg
H ₂ O	<i>ad</i> 270 µl	<i>ad</i> 450 µl
CaCl ₂ (2.5 M)	30 µl	50 µl
2x BBS	300 µl	500 µl

The transfection mix was incubated for 3 – 4 minutes at RT and added dropwise to the culture medium. The cells were incubated in a 5 % CO₂ setting for 3 – 4 h, the culture medium was replaced and the cells were maintained at 10 % CO₂ for approximately 48 h.

2x BBS solution*

BES	8.52 g
NaCl	13.08 g
Na ₂ HPO ₄	0.17 g
H ₂ O	<i>ad</i> 800 ml

*pH adjusted to 7.080 and filtrated through a 0.2 µm pore size syringe filter (VWR)

When fluorophore-containing plasmids were transfected, successful transfection and expression was evaluated via the EVOS® FL cell imaging system (Life Technologies, Thermo Fisher Scientific).

3.4.2 661W Cells

The 661W cells were kindly provided by Prof. Muayyad Al-Ubaidi (University of Houston). This cell line was cloned from murine retinal tumors and was found to exhibit molecular characteristics of cone photoreceptors (al-Ubaidi et al., 1992, Tan et al., 2004). 661W cells were maintained in DMEM medium + GlutaMAX + 1 g/l

glucose + pyruvate + 10 % FBS (Biochrom) + 1 % Antibiotic-Antimycotic in a CO₂ incubator (Heraeus, Thermo Fisher Scientific) at 37 °C and 10 % CO₂. They were passaged twice per week at a confluence of approximately 90 %. Transient transfections were performed using the calcium phosphate method as described in

3.4.3 Mouse Embryonic Fibroblasts

Mouse embryonic fibroblasts (MEFs) were generated as described (Jat et al., 1986, Xu, 2005). The cells were maintained in DMEM medium + GlutaMAX + 1 g/l glucose + pyruvate + 10 % FBS (Biochrom) + 1 % penicillin/streptomycin (Biochrom) in a CO₂ incubator (Heraeus, Thermo Fisher Scientific) at 37 °C and 5 % CO₂. They were passaged once per week at a confluence of approximately 90 %.

MEFs were transiently transfected using the TurboFect™ Transfection Reagent (Thermo Fisher Scientific). Cells were seeded onto a 6 cm cell culture plate and incubated until 70 – 90 % confluence was reached. The reaction mix was prepared in the following order.

TurboFect™ transfection mix

DNA	6 µg
Serum-free DMEM	<i>ad</i> 600 µl
TurboFect	12 µl

After addition of each component, the solution was mixed vigorously by vortexing. The transfection mix was incubated for 15 – 20 min at RT and subsequently added dropwise to the culture plate. The medium was exchanged after 3 h and the cells were harvested 48 h post transfection. When fluorophore-containing plasmids were transfected, successful transfection and expression was evaluated via the EVOS® FL cell imaging system (Life Technologies, Thermo Fisher Scientific).

3.5 PRODUCTION OF RECOMBINANT ADENO-ASSOCIATED VIRUSES

3.5.1 Transfection and Harvest

Recombinant adeno-associated viruses (rAAVs) were produced by a triple calcium phosphate transfection of a pAAV2.1 plasmid containing the gene of interest, a pAD Helper plasmid and a plasmid encoding the desired capsid. For subretinal injections

into the murine retina, the 2/8Y733F capsid variant was selected due to its high efficiency in transducing photoreceptors and the RPE (Petrus-Silva et al., 2009). HEK293T cells were seeded onto 15x 15 cm cell culture plates and incubated overnight until a confluence of 60 – 80 % was reached. Prior to transfection, the FBS-containing medium was replaced by serum-free medium. The transfection reagents were added to a 50 ml Falcon tube in the indicated order. CaCl₂ and 2x BBS were added dropwise during vortexing.

Transfection mix for AAV production

pAAV2.1 plasmid	270 µg
pAD Helper plasmid	X µg
Capsid plasmid	Y µg
H ₂ O	<i>ad</i> 11.85 ml
Polybrene (8 mg/ml)	15 µl
Dextran (10 mg/ml)	1.5 ml
CaCl ₂ (2.5 M)	1.5 ml
2x BBS	15 ml

The required amounts of pAD Helper and capsid plasmid were calculated using the following formulas:

$$X \mu\text{g} = \frac{270 \mu\text{g} \times \text{MM of pAD Helper}}{\text{MM of pAAV2.1}}$$

$$Y \mu\text{g} = \frac{270 \mu\text{g} \times \text{MM of capsid plasmid}}{\text{MM of pAAV2.1}}$$

2 ml of the transfection mix were added dropwise to each of the 15 culture plates. The plates were rocked gently and subsequently placed in a 5 % CO₂ setting for 24 h. The medium was exchanged afterwards and the plates were placed in a 10 % CO₂ setting for another 48 h.

The virus-containing culture medium was harvested twice. The first harvest took place 72 h post transfection by collecting the entire medium of all plates and adding fresh medium. The second harvest took place after another 72 h incubation period. The culture medium was collected into a 500 ml centrifugation tube. The residual cells were removed from the medium by centrifugation at 4,000 rpm and 4 °C for 15

min (JA-10 rotor, J2-MC High speed centrifuge, Beckman Coulter) and filtration of the supernatant through a 0.45 µm PES filter unit (Nalgene, Thermo Fisher Scientific). A 40 % polyethylene glycol (PEG) solution was added to the flow-through to a final concentration of 8 % and kept at 4 °C overnight to precipitate the viral particles. Subsequently, the solution was centrifuged at 4,000 rpm and 4 °C for 15 min (JA-10 rotor, J2-MC High speed centrifuge, Beckman Coulter). The supernatant was discarded and the virus-containing pellet was stored at -20 °C until further processing.

3.5.2 Iodixanol Density Gradient Centrifugation

The pellet was re-suspended in 7.5 ml sterilely filtrated PBS and incubated with Benzonase® (VWR) at a final concentration of 50 U/ml in a water bath (Haake) at 37 °C for 30 min to remove residual, non-packaged DNA. Next, the virus suspension was pipetted into a Quick-Seal polypropylene tube (39 ml, Beckman Coulter) and a density gradient was established by adding solutions with different iodixanol concentrations below the virus suspension in the following order: 7 ml of 15 %, 6 ml of 25 %, 5 ml of 40 % and 6 ml of 60 % iodixanol solution. For this purpose, the MINIPULS 3 peristaltic pump (Gilson) and a long glass pipette were used. Subsequently, the tubes were sealed with the Beckman Tube Topper and centrifuged at 70,000 rpm and 18 °C for 1 h 45 min in an Optima L-80K ultracentrifuge (70 Ti rotor, Beckman Coulter). Afterwards, the top of the tube was perforated with a cannula to ensure air flow. The 40 % iodixanol phase enriched with viral particles was collected from the gradient using a 20 G cannula and 20 ml syringe by puncturing the tube laterally at the border between the 40 % and the 60 % phase. The virus-containing solution was stored at -80 °C until further processing.

<u>15 % iodixanol solution*</u>		<u>25 % iodixanol solution*</u>	
10x PBS	5 ml	10x PBS	5 ml
MgCl ₂ (1 M)	50 µl	MgCl ₂ (1 M)	50 µl
KCl (2.5 M)	50 µl	KCl (2.5 M)	50 µl
NaCl (5 M)	10 ml	Optiprep™ (Progen)	20.9 ml
Optiprep™ (Progen)	10 ml	phenol red (1 %, w/v)	50 µl
phenol red (1 %, w/v)	37.5 µl	H ₂ O	<i>ad</i> 50 ml
H ₂ O	<i>ad</i> 50 ml		

<u>40 % iodixanol solution*</u>		<u>60 % iodixanol solution*</u>	
10x PBS	5 ml	MgCl ₂ (1 M)	50 µl
MgCl ₂ (1 M)	50 µl	KCl (2.5 M)	50 µl
KCl (2.5 M)	50 µl	Optiprep™ (Progen)	50 ml
NaCl (5 M)	10 ml	phenol red (1 %, w/v)	35.5 µl
Optiprep™ (Progen)	33.3 ml		
H ₂ O	<i>ad</i> 50 ml		

*filtrated through a 0.2 µm pore size syringe filter (VWR)

3.5.3 Anion Exchange Chromatography

To further purify the virus, an anion exchange chromatography was performed employing the ÄKTAprime plus chromatography system (GE Healthcare), the 5 ml HiTrap™ Q FF anion exchange chromatography column (GE Healthcare) and the PrimeView 5.31 software (GE Healthcare). Before starting, the column was equilibrated with buffer A and the virus-containing solution was diluted with this buffer in a 1:1 ratio. The solution was loaded onto the column via a loop injector (50 ml Superloop, GE Healthcare). The UV light-absorbance and conductance properties of the collected fractions were monitored providing information about the amount of contained virus. Remaining bound molecules were removed from the column using a 2.5 M NaCl solution. All virus-containing fractions were pooled and used for subsequent processing.

Buffer A *

Tris	20 mM
NaCl	15 mM
H ₂ O	<i>ad</i> 500 ml

*pH adjusted to 8.5

3.5.4 Increase of rAAV Concentration

To increase the virus concentration, Amicon® Ultra-4 centrifugation filter units (Merck) with a molecular weight cut-off of 100 kDa were used. The virus-containing solution was loaded on top of the filter unit and centrifuged at 4,000 rpm (JA-10 rotor, J2-MC High speed centrifuge, Beckman Coulter) and 4 °C in 20 min intervals until the volume was reduced to 500 µl. Subsequently, the filter unit was washed

with 1 ml 0.014 % Tween/PBS-MK. The solution was further centrifuged under the same conditions until the volume was reduced to 100 μ l of concentrated viral solution. 10 μ l aliquots were prepared and stored at -80 °C until use.

0.014 % Tween/PBS-MK solution

10x PBS	50 ml
MgCl ₂ (1 M)	500 μ l
KCl (2.5 M)	500 μ l
Tween-20	70 μ l
H ₂ O	<i>ad</i> 500 ml

3.5.5 rAAV Titer Determination

To determine the titer of the produced rAAVs, a qPCR was performed using the StepOnePlus Real-Time PCR system (Applied Biosystems, Thermo Fisher Scientific). A standard curve was produced to serve as a reference. For this purpose, a fragment containing a part of the ITR was amplified by PCR using the following primers.

ITR2 forward: 5' GGAACCCCTAGTGATGGAGTT 3'

ITR2 reverse: 5' CGGCCTCAGTGAGCGA 3'

Next, the amplicon was purified and the concentration was determined using the Nanodrop™ 2000c spectrophotometer (Thermo Fisher Scientific). To obtain a standard solution containing 10¹⁰ copies of the amplicon in 5 μ l, the following equation was used.

$$c \left[\frac{\text{pg}}{\mu\text{l}} \right] = \frac{10^{10} \times 660 \times 10^{12} \frac{\text{pg}}{\text{mol}} \times \text{ITR fragment size}}{6.022 \times 10^{23} \frac{1}{\text{mol}} \times 5 \mu\text{l}}$$

After preparation of a standard solution with concentration *c*, a dilution series ranging from 10¹⁰ to 10¹ copies was made. To attain the standard curve, a qPCR was run with three technical replicates of the standard dilution series. For this purpose, a MicroAmp™ Fast Optical 96-Well Reaction Plate (Applied Biosystems, Thermo Fisher Scientific) and the PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific) were used. The virus solution was diluted 100-fold in H₂O and run on the

same reaction plate with three technical replicates. The reaction mix was prepared as follows.

qPCR reaction mix

PowerUp™ Master Mix	10 µl
ITR2 F primer (10 µM)	0.6 µl
ITR2 R primer (10 µM)	0.6 µl
Virus or standard	5 µl
H ₂ O	<i>ad</i> 20 µl

The qPCR was run using the program depicted in Table 7.

Table 7 Fast cycling program for real-time qPCR

Step	Number of cycles	Temperature	Duration
Uracil DNA glycosylase activation	1	50 °C	2 min
Dual-Lock™ DNA polymerase activation	1	95 °C	2 min
Denaturation	40	95 °C	5 s
Annealing		58 °C	5 s
Extension		60 °C	30 s
Melt curve acquisition	1	95 °C	15 s
	1	60 °C	1 min
	1	Ramp to 95 °C (0.3 °C/s)	2 min
	1	95 °C	15 s

The obtained data was analyzed using the StepOnePlus Real-Time PCR system software (Applied Biosystems, Thermo Fisher Scientific). Baseline settings and cycling threshold position were adjusted manually if necessary. The standard curve was attained by plotting the resulting cycle threshold (Ct) values against the logarithm of the dilution. The number of viral genomes per µl (vg/µl) of the produced rAAVs could be inferred from the standard curve.

3.6 ANIMALS

All animal protocols were performed with permission of local authorities (District Government of Upper Bavaria) and in accordance with the German laws on animal welfare (Tierschutzgesetz). For all experiments, C57Bl6/J mice were used. The mice were maintained at a 12 h light/dark cycle and received food and water *ad libitum*.

3.7 SUBRETINAL INJECTION

For subretinal injections, C57Bl6/J mice at postnatal day 21 (P21) were anesthetized by an intraperitoneal injection of ketamine (40 mg/kg body weight) and xylazine (20 mg/kg body weight). After a complete absence of the paw withdrawal reflex the pupils were dilated by administration of atropine (1 %)- and tropicamide (0.5 %)-containing eye drops (Mydriaticum Stulln, Pharma Stulln GmbH). The fundus of the eye was focused using a surgical microscope (OPMI 1 FR pro, Zeiss). 1 μ l containing 10^{10} rAAV particles was injected subretinally by a single injection with a NANOFIL 10 μ l syringe (World Precision Instruments) and a 34 G beveled needle (World Precision Instruments). The injected eyes were treated with an eye ointment containing 5 mg/g gentamicin and 0.3 mg/g dexamethasone. The mice were kept on a heating plate (Leica HI1120, Leica Biosystems) at 37 °C until full recovery from anesthesia. Two to four weeks post-injection, all injected retinas were harvested and processed for RT-PCR analysis or immunohistochemistry.

3.8 IMMUNOHISTOCHEMISTRY

Subretinally injected mice were euthanized via cervical dislocation. The eyes were removed and placed in 0.1 M phosphate buffer (PB). Subsequently, the eyeball was punctured at the *ora serrata* with a 21 G cannula and fixed in 4 % paraformaldehyde (PFA, Sigma Aldrich, pH adjusted to 7.4) for 5 min. Next, the eye was positioned below a stereomicroscope (Stemi 2000, Zeiss) on a filter paper drenched with 0.1 M PB. By cutting alongside the *ora serrata* using surgical scissors (SuperFine Vannas, World Precision Instruments), the cornea, lens and vitreous body were removed. The remaining part of the eyeball containing the retina was fixed in 4 % PFA for 45 min at RT and subsequently washed 3 times for 5 min in 0.1 M PB. For

cryopreservation, the eyeball was placed in a 30 % sucrose solution (w/v) overnight at 4 °C.

The next day, the eyeball was embedded in tissue freezing medium (Sakura) and chilled on dry ice until the medium solidified. The retina was sectioned using a cryostat (Leica CM3050 S, Leica Biosystems) into 10 µm thick slices, collected on coated glass object slides (Superfrost Plus microscopic slides, Thermo Fisher Scientific) and stored at -20 °C.

For immunohistochemical staining, the retinal sections were thawed at RT and encircled using a Super PAP Pen Liquid Blocker (Science Services). Subsequently, the sections were rehydrated with 0.1 M PB for 5 min and fixed with 4 % PFA for 10 min. After washing the sections three times for 5 min each with 0.1 M PB, a solution containing the primary antibodies, 5 % ChemiBLOCKER (Merck) and 0.3 % Triton X-100 in 0.1 M PB was applied. The cryosections were incubated with the primary antibody solution overnight at 4 °C. The next day, the retinas were washed three times with 0.1 M PB for 5 min and incubated with a solution containing the secondary antibodies and 2 % ChemiBLOCKER in 0.1 M PB for 1.5 h at RT. After subsequent washing with 0.1 M PB three times for 5 min, the cell nuclei were stained with 5 µg/ml Hoechst 33342 solution (Invitrogen). Finally, the sections were washed with 0.1 M PB, embedded in Fluoromount-G Mounting Medium (Thermo Fisher Scientific), covered by a cover slip and stored at 4 °C. All antibodies used for immunohistochemistry are listed in Table 8.

0.1 M phosphate buffer (PB)*

Na₂HPO₄ x 2H₂O 28.48 g
 NaHPO₄ x H₂O 5.52 g
 H₂O ad 2 l

*pH adjusted to 7.4

Table 8 List of all antibodies used for immunohistochemistry

Antibody	Source	Catalog nr.	Dilution
Myc-Tag (9B11) Mouse mAb	Cell Signaling Technology	2276	1:200
α-CNGB1a-antibody, rabbit	in lab production	–	1:10,000

	(Huttl et al., 2005)		
Cy TM 5 AffiniPure Goat Anti-Mouse IgG (H+L)	Jackson ImmunoResearch	115-175- 146	1:400
Cy TM 3 AffiniPure Donkey Anti-Rabbit IgG	Jackson ImmunoResearch	711-165- 152	1:400

3.9 CONFOCAL MICROSCOPY

Images of the stained retinas were obtained using a Leica TCS SP8 inverted confocal laser scanning microscope (Leica Microsystems) equipped with a 405 nm diode as well as a 552 nm and a 633 nm optically pumped semiconductor laser appropriate for the excitation of Hoechst 33342, Cy3 and Cy5, respectively. The filter settings were chosen according to the emission spectrum of the respective dye. Images were acquired as z-stacks (1 μ m steps) with a HC PL APO 40x/1.30 Oil CS2 objective (Leica Microsystems) and type F immersion liquid (Leica Microsystems) using the LAS X software (Leica Microsystems). With the same software, z-stacks were condensed into 2D images by applying a maximum intensity projection. The images were processed further with the ImageJ 1.48v software (National Institutes of Health).

Images of transiently transfected living cells were obtained using the Leica TCS SP8 spectral confocal laser scanning microscope (Leica Microsystems) equipped with a 448 nm, a 514 nm and a 552 nm optically pumped semiconductor laser appropriate for the excitation of cerulean, citrine and mCherry, respectively. The filter settings were chosen according to the emission spectrum of the respective fluorophore. Images were acquired with a HCX APO 20x/1.00 W objective (Leica Microsystems). All images were processed with the ImageJ 1.48v software.

3.10 RNA EXTRACTION

For RNA extraction from injected retinas, the mice were euthanized via cervical dislocation. The retinas were collected by placing blunt forceps below the eye, incising the eye ball using a sterile scalpel (Swann-Morton) and gradually moving the forceps upward. Three retinas were pooled per construct and RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer's

instructions. For disruption, the tissue was homogenized by adding 350 μl of RLT buffer (Qiagen, provided with the kit) + 3.5 μl β -mercaptoethanol (β -ME, Sigma Aldrich) and passing it through a 20 G needle fitted onto a sterile syringe at least five times. The remaining steps were executed according to protocol. The RNA was eluted in 30 μl RNase-free H_2O .

For RNA extraction from transiently transfected cells, the RNeasy Mini Kit Plus (Qiagen) was used. For this purpose, the medium was removed from the 6 cm culture plates and the cells were scrapped off using a 16 cm cell scraper (Sarstedt). The cells were collected in 500 μl medium into a 2 ml safe-lock tube (Eppendorf) and centrifuged at 3,000 x g and 4 °C for 10 min. The supernatant was discarded and the pellet was re-suspended in 600 μl RLT Plus buffer (Qiagen, provided with the kit) + 6 μl β -ME (Sigma Aldrich). A steel ball was placed into each tube and the cells were disrupted using the mixer mill MM400 (Retsch) at 30 Hz for 1 min. Afterwards, the balls were removed and the suspension was centrifuged at 21,000 x g and RT for 5 min. The remaining steps were executed according to protocol including an optional step removing the genomic DNA via gDNA eliminator spin columns. The RNA was eluted in 30 μl RNase-free H_2O .

The RNA concentrations were measured using the Nanodrop™ 2000c spectrophotometer (Thermo Fisher Scientific). The RNA was kept on ice until further use or stored at -20 °C for short-term or -80 °C for long-term storage.

3.11 cDNA SYNTHESIS

For cDNA synthesis, the Revert Aid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) was used according to the manufacturer's instructions. Equal amounts of RNA were used per experiment.

cDNA synthesis reaction mix

Total RNA	up to 2 μg
Oligo (dT) ₁₈ primer	1 μl
Random hexamer primer	1 μl
Nuclease-free H_2O	<i>ad</i> 12 μl

The reaction mix was incubated at 65 °C for 5 min to remove secondary structures of the RNA and subsequently chilled to 10 °C. Afterwards, the following components were added.

5x Reaction buffer	4 µl
RiboLock RNase Inhibitor (20 U/µl)	1 µl
dNTP mix (10 mM)	2 µl
RevertAid M-MuLV RT (200 U/µl)	1 µl

The cDNA reaction mix was incubated in a Mastercycler® nexus gradient (Eppendorf) using the following cycling conditions.

cDNA synthesis cycling conditions

25 °C	5 min
42 °C	60 min
70 °C	5 min
10 °C	∞

The cDNA was kept on ice until further use or stored at -20 °C for short-term or -80 °C for long-term storage.

3.12 REVERSE TRANSCRIPTION PCR

Reverse transcription PCR (RT-PCR) was performed using the Herculase II fusion DNA polymerase (Agilent Technologies) or the VWR *Taq* DNA polymerase (VWR) by following the respective protocol.

Herculase II RT-PCR reaction mix

5x Herculase II Reaction Buffer	5 µl
dNTPs (10 mM)	1 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Herculase II DNA polymerase	0.5 µl
cDNA	1 µl
H ₂ O	<i>ad</i> 25 µl

VWR Taq RT-PCR reaction mix

10x Key Buffer	2.5 µl
dNTPs (10 mM)	1 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Taq DNA polymerase (5 U/µl)	0.25 µl
cDNA	1 µl
H ₂ O	ad 25 µl

A list of all RT-PCR primers used in this study can be found in 7.1. The RT-PCR was run in a ProFlex PCR System Cycler (Applied Biosystems, Thermo Fisher Scientific) using the respective cycling conditions described in Table 9 and Table 10.

Table 9 Herculanase II RT-PCR cycling conditions

Step	Number of cycles	Temperature	Duration
Initial Denaturation	1	95 °C	2 min
Denaturation	15 - 35	95 °C	20 s
Annealing		50 °C – 72 °C	30 s
Elongation		72 °C	60 s/kb
Final Elongation	1	72 °C	4 min
Cooling	1	10 °C	15 min

Table 10 VWR Taq polymerase RT-PCR cycling conditions

Step	Number of cycles	Temperature	Duration
Initial Denaturation	1	95 °C	2 min
Denaturation	25 - 35	95 °C	30 s
Annealing		48 °C – 65 °C	40 s
Elongation		72 °C	60 s/kb
Final Elongation	1	72 °C	5 min
Cooling	1	10 °C	15 min

The RT-PCR products were separated on a 0.7-2 % agarose gel and detected with the ChemiDoc™ MP Imaging System (Bio-Rad). The sequence identity of all bands was confirmed by sequencing (Eurofins Genomics).

3.13 BRANCH POINT ANALYSIS

3.13.1 Nested Lariat RT-PCR

RNA was extracted according to the protocol described in 3.10. 10 µg RNA were incubated with RNase R (Lucigen) following the manufacturer's instructions to remove any non-circular RNA. Subsequent cDNA synthesis (Revert Aid First Strand cDNA Synthesis Kit, Life Technologies) was performed as described in 3.11. For this reaction, only the random hexamer primer was used. Next, the lariats were amplified via a nested RT-PCR using the Herculase II fusion DNA polymerase (Agilent Technologies). For the first round of amplification, the reaction mix was prepared as described in 3.12. For the second round of amplification, 5 µl of the first PCR were added to the reaction mix instead of the cDNA. Moreover, a second primer pair binding 25 – 30 bp downstream of the first primer pair was used. The applied cycling conditions are listed in Table 11.

Table 11 Herculase II nested lariat PCR cycling conditions

Step	Number of cycles	Temperature	Duration
Initial Denaturation	1	95 °C	2 min
Denaturation	35	95 °C	20 s
Annealing		65 °C	30 s
Elongation		72 °C	40 s
Final Elongation	1	72 °C	4 min
Cooling	1	10 °C	15 min

The nested RT-PCR products were analyzed on a 2 % agarose gel. Obtained bands were excised and the DNA was extracted using the Qiaquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions.

3.13.2 TOPO Cloning and Analysis of Lariats

To be able to analyze the sequence identity of the single lariats, the products of the nested RT-PCR were subcloned into plasmids. For this purpose, 3'-adenine overhangs were added to the DNA fragments post amplification by incubating one unit of *Taq* Polymerase (VWR) with the PCR reaction for 10 min at 72 °C. Subsequently, the lariats were subcloned into a TOPO vector (TOPO TA Cloning Kit, Thermo Fisher Scientific) according to the manufacturer's instructions. The plasmids were transformed into bacteria and a small-scale plasmid preparation was performed as described in 3.3.2. The resulting plasmids were sequenced (Eurofins Genomics) and the obtained lariat sequences were analyzed by alignment with the investigated intron using the DNAMAN software (Lynnon Biosoft).

3.14 PROTEIN EXTRACTION

For protein extraction, the medium was removed from the 6 or 10 cm culture plates. The cells were scrapped off using a 16 cm cell scraper (Sarstedt) and collected in 500 µl medium into a 2 ml safe-lock tube (Eppendorf). The cell suspension was centrifuged at 3,000 x g and 4 °C for 10 min, the supernatant was discarded and the pellet re-suspended in 150 µl and 250 µl Triton X-100 (TX) lysis buffer for 6 cm and 10 cm plates, respectively. A steel ball was added to each safe-lock tube and the cells were disrupted using the mixer mill MM400 (Retsch) at 15 Hz for 30 s. Afterwards, the tubes containing the balls were rotated end-over-end (VWR™ tube rotator) for 20 min at 4°C. Subsequently, the balls were removed and the lysates were centrifuged at 5,000 x g for 10 min at 4 °C. The protein-containing supernatant was transferred into a new reaction tube and stored at -20 °C.

Triton X-100 lysis buffer*

Triton X-100	2.5 ml
NaCl (5 M)	15 ml
CaCl ₂ (2.5 M)	0.4 ml
H ₂ O	<i>ad</i> 500 ml

*cOmplete™ ULTRA Protease Inhibitor Cocktail tablets (Roche) were added right before use (1 tablet/10 ml)

Total protein concentration was determined using the Bradford assay (Bradford, 1976). 5 µl of the protein lysate were mixed with 95 µl of a 0.15 M NaCl solution and transferred into a PMMA standard disposable cuvette (BRAND). Subsequently, 1 ml of Coomassie blue solution was added, thoroughly mixed by pipetting and incubated at RT for 2 min. Using the BioPhotometer (Eppendorf), the absorption of the solution was measured against a blank control containing 5 µl of the TX lysis buffer. The obtained value represents the total amount of protein contained in 5 µl lysate.

Coomassie blue solution

Coomassie brilliant blue G250	50 mg
95 % ethanol (v/v)	25 ml
85 % phosphoric acid (v/v)	50 ml
H ₂ O	<i>ad</i> 500 ml

3.15 WESTERN BLOTTING

For western blotting, 10 – 35 µg protein were mixed with 6x Laemmli + DTT buffer and incubated at 72 °C for 10 min. Equal amounts of protein were used for each experiment. For SDS-PAGE, polyacrylamide running gels were prepared using the Mini-Protean 3 gel system (Bio-Rad). For the detection of small proteins (up to 35 kDa) a 15 % gel was used, while a 6 % gel was prepared for the detection of larger proteins. After solidification of the running gel, a stacking gel was added on top.

4x Tris-HCl/SDS pH 6.8 buffer*

Tris-HCl	0.5 M
SDS	0.4 % (w/v)

*pH adjusted to 6.8

4x Tris-HCl/SDS pH 8.8 buffer*

Tris-HCl	1.5 M
SDS	0.4 % (w/v)

*pH adjusted to 8.8

6x Laemmli + DTT buffer

4x Tris-HCl/SDS pH 6.8 buffer	7 ml
Glycerol	3 ml
SDS	1 g
Bromophenol blue	1.2 mg
DTT	0.93 g

Materials & Methods

<u>Running gel for SDS-PAGE</u>	<u>6 %</u>	<u>15 %</u>
30 % acrylamide/bis-acrylamide	2.3 ml	5.75 ml
4x Tris-HCl/SDS pH 8.8 buffer	2.8 ml	2.8 ml
H ₂ O	6.2 ml	2.75 ml
APS	22.5 µl	22.5 µl
TEMED*	7.5 µl	7.5 µl

*added right before pouring of gel

<u>Stacking gel for SDS-PAGE</u>	
30 % acrylamide/bis-acrylamide	0.5 ml
4x Tris-HCl/SDS pH 6.8 buffer	0.95 ml
H ₂ O	2.3 ml
APS	18.75 µl
TEMED*	3.75 µl

*added right before pouring of gel

The lysates and a PageRuler™ Plus Prestained Protein Ladder (Thermo Fisher Scientific) were loaded onto the gel and the SDS-PAGE was run at 140 V and RT for 1 h 10 min in 1x electrophoresis buffer.

10x electrophoresis buffer

Tris	30 g
Glycine	144 g
SDS	10 g
H ₂ O	<i>ad 1 l</i>

Subsequently, the separated proteins were transferred onto a PVDF membrane (PEQLAB, VWR) using a Mini Trans-Blot cell (Bio-Rad). Western blotting was performed at 100 V and RT for 1 h 30 min in 1x blotting buffer. Afterwards, the membrane was transferred into a blocking solution containing 5 % non-fat dried milk powder (w/v, AppliChem) in TBST buffer and incubated for 1 h at RT. Next, the membrane was incubated in a solution containing the primary antibody and 1 % non-fat dried milk powder (w/v, AppliChem) in 5 ml TBST buffer for 1 h at RT or overnight at 4 °C. After three subsequent washing steps in TBST buffer for 5 min each, the membrane was incubated in a solution containing the secondary antibody

and 1 % non-fat dried milk powder (w/v, AppliChem) in 5 ml TBST buffer for 1 h at RT. After washing the membrane three more times in TBST buffer for 5 min each and rinsing it with H₂O, the western blot was developed by applying Western Blotting Luminol Reagent (Santa Cruz) according to the manufacturer's instructions. The signal was detected using the ChemiDoc™ MP Imaging System (Bio-Rad). All antibodies used for western blotting can be found in Table 12.

<u>10x blotting buffer</u>		<u>TBST buffer</u>	
Tris	30 g	Tris	1.2 g
Glycine	144 g	NaCl	8.0 g
H ₂ O	ad 1 l	Tween-20	1 ml
		H ₂ O	ad 1 l

Table 12 List of all antibodies used for western blotting

Antibody	Source	Catalog nr.	Dilution
Living Colors A.v. Monoclonal Antibody (JL-8), mouse α -GFP	Clontech, Takara	632380	1:2,000
β -Tubulin (D3U1W) Mouse mAb	Cell Signaling Technology	86298	1:500
CRISPR/Cas9 polyclonal antibody, rabbit	Diagenode	C15310258	1:1,000
goat anti-mouse IgG-HRP	Santa Cruz Biotechnology	sc-2005	1:2,000
goat anti-rabbit IgG-HRP	Santa Cruz Biotechnology	sc-2004	1:2,000

For quantification, the Image Lab software (Bio-Rad) was used. Bands were automatically detected by the software and adjusted manually. The total volume intensity values obtained for the detected bands were transferred into an Excel file (Microsoft). All band intensities were normalized to the intensities obtained for the beta tubulin band for loading control. The relative amount of reconstitution was

calculated with respect to the *cis*-splicing control, which represents 100 % efficiency.

3.16 STATISTICS

The GraphPad Prism software 5 was used for scientific graphing and statistical analysis. The corresponding number of independent samples (n), applied statistical tests and significance levels are indicated for each experiment in the respective figure legend.

4 RESULTS

4.1 INVESTIGATION OF TWO ADRP-CAUSING RHODOPSIN MUTATIONS AND THEIR EFFECT ON MRNA SPLICING

4.1.1 Generation of *De Novo* Splice Acceptor Site in *RHO* Exon 3 by c.620T>G Mutation

Patients carrying the c.620T>G and c.620T>A mutations in *RHO* show a strikingly different disease penetrance. c.620T>G leads to an early onset and severe form of RP, while patients affected by the c.620T>A mutation show a late onset and a mild disease phenotype. These mutations are both located in *RHO* exon 3, which encodes for transmembrane helix domain 5 (H5) and the major part of the second extracellular loop (EL2) (Figure 3A). Both mutations affect the second nucleotide of the triplet encoding the methionine residue at position 207 (Met207), which is involved in binding of 11-*cis*-retinal (Figure 3B). Furthermore, both mutations are classified as missense mutations expected to replace Met207 with a positively charged amino acid, i.e. Met207Arg and Met207Lys. However, recent studies investigating mutations affecting Met207 could only detect moderate impact on 11-*cis*-retinal binding (Aguila et al., 2009, Audo et al., 2010a, Piechnick et al., 2012, Hernandez-Rodriguez et al., 2013) suggesting that the difference in disease course most likely does not result from differential impact on ligand binding. Another plausible explanation for these puzzling disease courses might be that they arise from a differential impact of these two mutations on mRNA splicing.

To address this issue, an *in silico* prediction analysis has been performed. Indeed, c.620T>G was predicted by two independent splicing prediction software tools (NNSplice and Human Splicing Finder, HSF) to generate a novel splice donor or acceptor site (Figure 9A, B). By contrast, c.620T>A was not predicted to alter mRNA splicing. To validate this finding experimentally, HEK293 cells have been transiently transfected with *RHO* minigenes containing all native exons and most native intronic regions of the human rhodopsin gene (Figure 10A, B). When analyzing mRNA splicing via RT-PCR, correctly spliced transcript could only be identified for the wild type (WT) construct and the c.620T>A mutant.

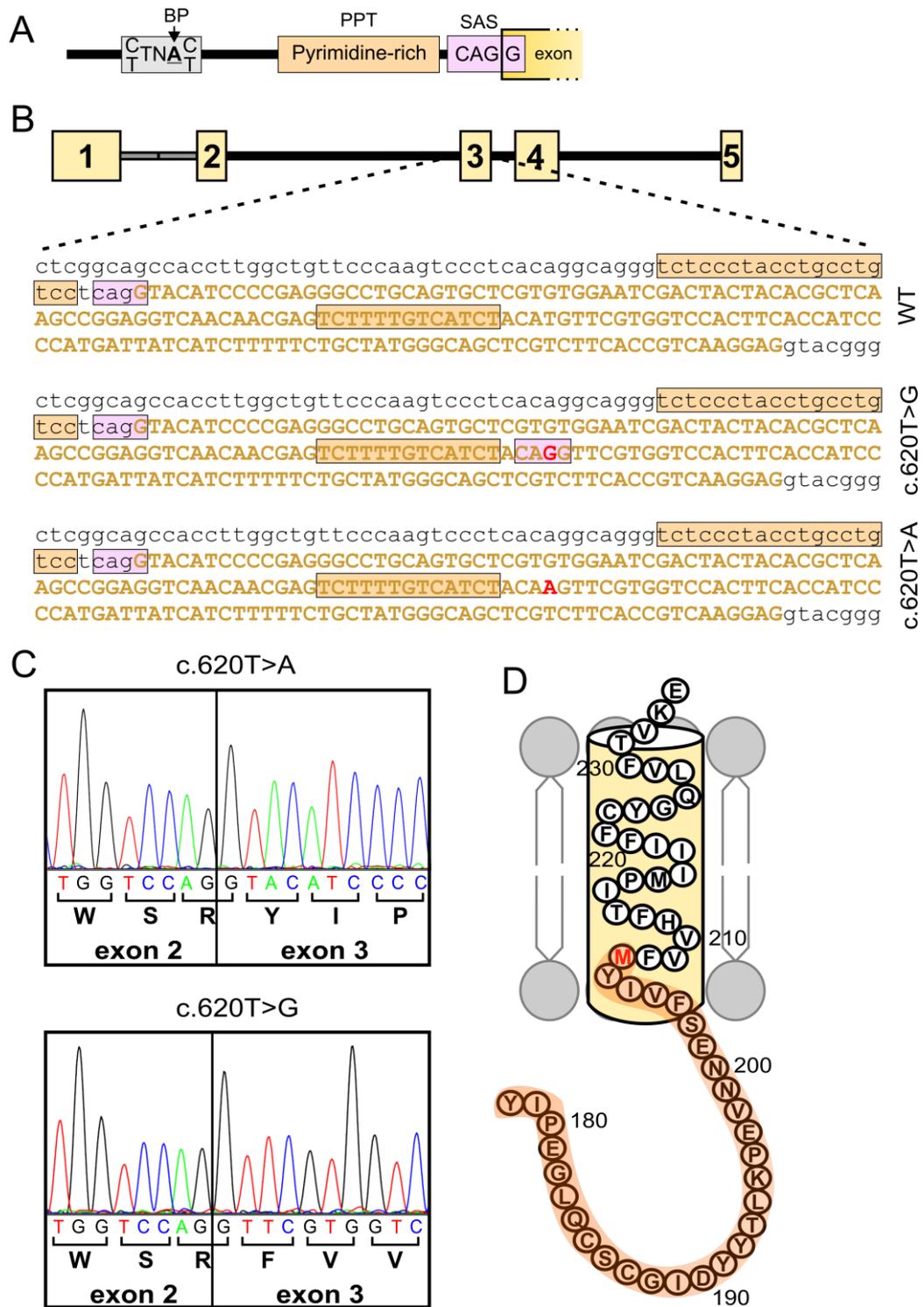


Figure 9 Splice acceptor site elements in WT and mutant RHO exon 3

A, Schematic illustration of all elements necessary to form a mammalian functional unit of a splice acceptor site (SAS). BP, branch point; PPT, polypyrimidine tract. B, Sequences of wild type (WT), c.620T>G and c.620T>A RHO exon 3 (brown uppercase) and partial intronic sequences (black lowercase). The point mutations are highlighted in red. The potential PPTs and SASs matching the consensus sequences are depicted with brown and pink rectangles, respectively. C, Representative electropherogram of c.620T>A and c.620T>G spliced products

showing the exon 2-exon 3 junction. D, Deletion of 90 bp in the c.620T>G mRNA causes an in-frame deletion of the 30 aa forming EL2 and the first 5 aa of H5 (shadowed area).

By contrast, the c.620T>G mutation resulted in two aberrantly spliced mRNAs originating from usage of a novel splice acceptor site (SAS) in exon 3 generated by the point mutation (Figure 10C). This aberrant splicing leads to an in-frame deletion of the 90 bp encoding the main part of EL2 and the first 5 aa of H5 (Figure 9C, D).

4.1.2 Mislocalization of c.620T>G Minigene-Born Protein *in Vitro*

As a next step, the expression pattern of the *RHO* minigene-born protein was investigated in transfected HEK293 cells. In heterologous expression systems, rhodopsin is usually located in the cell membrane (Liu et al., 2013, McKeone et al., 2014). In line with this, a substantial portion of the WT rhodopsin was detected as a small layer surrounding the cells, most likely representing the plasma membrane (Figure 10D). The expression pattern of the c.620T>A mutant protein was not distinguishable from the one observed for the WT rhodopsin. However, the c.620T>G mutant protein was completely absent from the cell membrane and mainly found in the cytosol (Figure 10D).

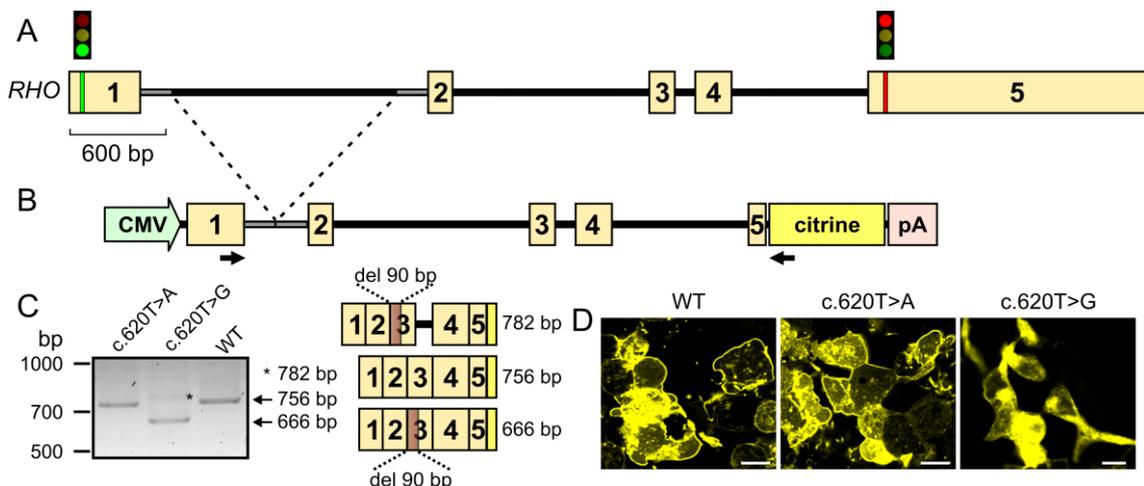


Figure 10 Generation of a *de novo* splice acceptor site in *RHO* exon 3 by c.620T>G

A, Structure of human rhodopsin gene (*RHO*). Exons are depicted as beige boxes. Start and stop codon are marked as traffic lights. B, *RHO* minigene used for *in vitro* studies. The minigene consists of all native exons and most native intronic sequences and is expressed under a cytomegalovirus (CMV) promoter. Citrine was fused to the 3' end of exon 5 for protein detection. To enable packaging into AAV

vectors for subsequent analysis *in vivo*, intron 1 was shortened as indicated. Primers used for RT-PCR (C) are indicated as black arrows. C, RT-PCR from HEK293 cells transiently transfected with the wild type (WT), c.620T>A or c.620T>G minigene. The most prominent bands are indicated by an arrow. An additional faint band representing aberrantly spliced mRNA with an intron 3 retention is highlighted by an asterisk. The sequence identity determined by sequencing is depicted in the right panel. D, Confocal live images of transfected HEK293 cells. Scale bar, 10 μ m.

In the RT-PCR assay, no correctly spliced *RHO* mRNA could be detected for the c.620T>G mutant. Nevertheless, this experiment is not sensitive enough to exclude that some residual protein originating from correctly spliced transcript below the detection limit might be present. This issue is particularly important, as a gain-of-function or dominant-negative effect of the resulting Met207Arg protein could be contributing to the severe disease phenotype. The live imaging experiment shown in Figure 10D is in principle expected to be sensitive enough for the detection of proteins expressed at low levels. However, this experimental setup does not allow for discrimination between protein expressed from aberrantly or from correctly spliced mRNA.

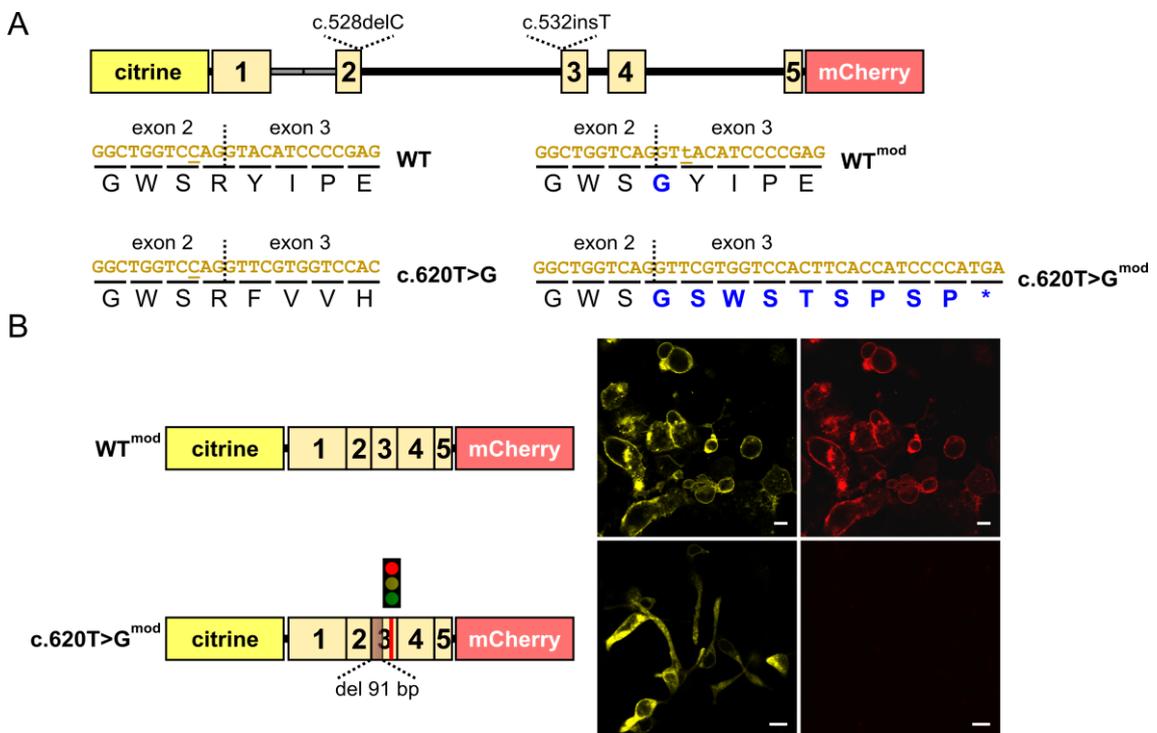


Figure 11 Live imaging dual fluorescence splicing detection assay

A, Modified *RHO* minigenes (WT^{mod} and c.620T>G^{mod}) containing a citrine sequence 5' of exon 1 as an expression control and a mCherry sequence 3' of exon 5 as a splicing indicator. Two mutations (c.528delC and c.532insT) were introduced into

the minigenes resulting in an aa exchange (Arg177Gly) upon correct splicing or a 1 bp frameshift and a premature stop codon in exon 3 upon aberrant splicing at the c.620T>G SAS. Changed amino acids are shown in blue. B, Left panel, Schematic depiction of differentially spliced mRNAs expressed from WT^{mod} or c.620T>G^{mod} minigenes. Right panel, Confocal live images of HEK293 cells transfected with the modified minigenes. Aberrant splicing results in an absence of mCherry fluorescence. Scale bar, 10 μ m.

To enable a fluorescence-based discrimination of proteins expressed from these two splice isoforms, I developed a live imaging dual fluorescence splicing detection assay (Figure 11). Using this assay, a robust signal originating from protein expressed from aberrantly spliced c.620T>G transcript could be detected. However, there was no detectable signal corresponding to minigene-born Met207Arg protein expressed from correctly spliced transcript. These data are indicating that the novel SAS generated by the c.620T>G mutation is used exclusively and that the Met207Arg protein most likely does not naturally exist.

4.1.3 Missplicing and Mislocalization of the c.620T>G Mutant in Mouse Photoreceptors

mRNA splicing can vary in a cell type-specific manner (Arias et al., 2010). Therefore, the results obtained from HEK293 cells might not reflect the native situation in rod photoreceptors. To investigate this issue, *RHO* WT or mutant minigenes were delivered subretinally into mouse retinas via recombinant adeno-associated virus (rAAV)-mediated gene transfer and expressed in rod photoreceptors using a murine rhodopsin promoter (Figure 12A). The RT-PCR results obtained from these retinas were very similar to those from the *in vitro* experiments shown in Figure 10C. However, the faint band corresponding to aberrantly spliced mRNA with an intron 3 retention found in HEK293 cells (cf. Figure 10C) was not detected *in vivo* (Figure 12B). Next, the WT and mutant protein expression was investigated in transduced murine rod photoreceptors. As expected, minigene-born WT rhodopsin was found in the rod outer segments. The expression of c.620T>A mutant protein was virtually indistinguishable from the WT one suggesting no substantial effects of this mutant on protein localization.

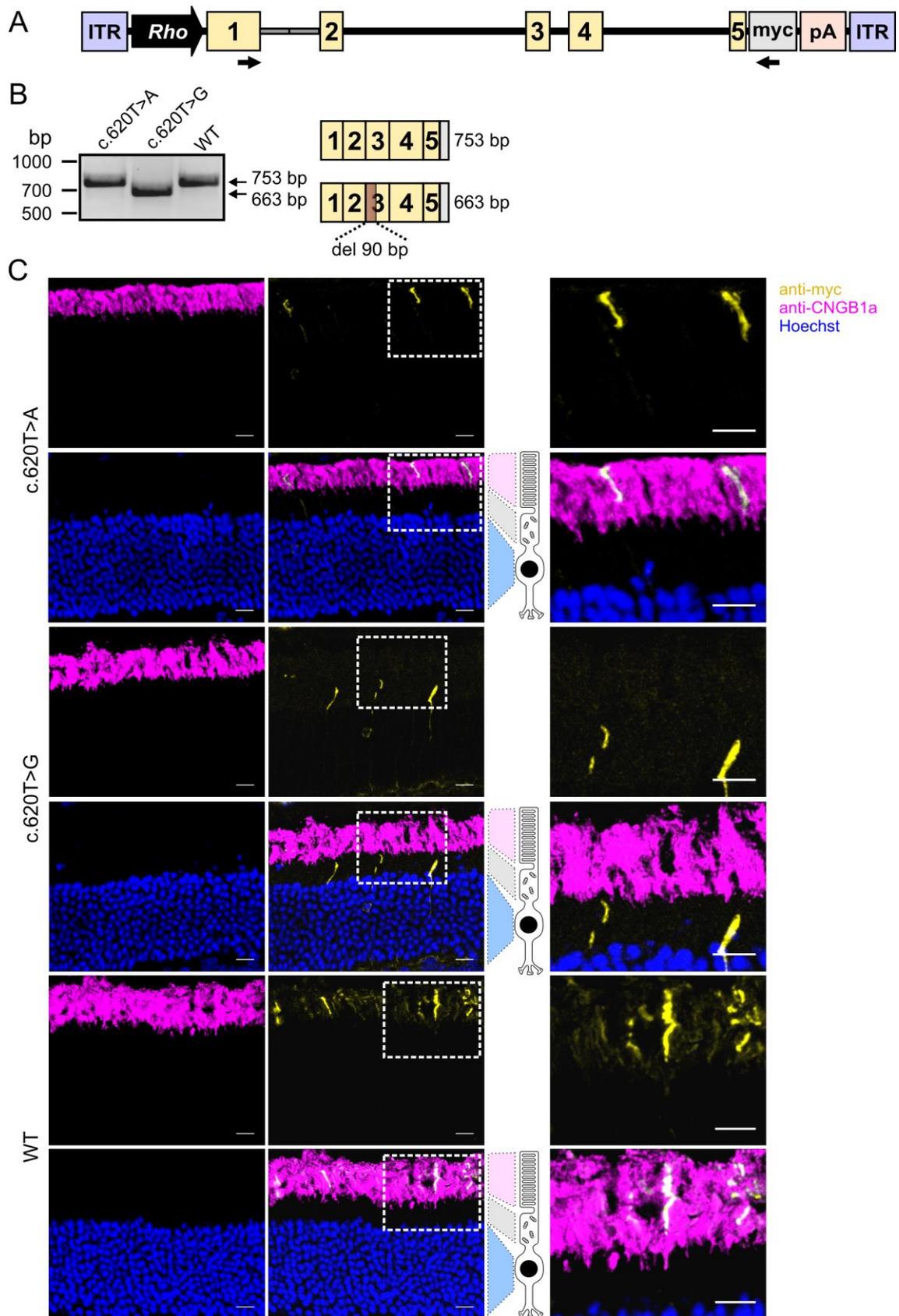


Figure 12 *RHO* minigene expression in mouse rod photoreceptors

A, Structure of *RHO* minigenes used for *in vivo* expression. Gene expression was driven by a murine rhodopsin promoter (*Rho*) and a myc-tag was used for protein

detection. The minigene was utilized for simultaneous analysis of mRNA splicing and protein expression. Primers used for RT-PCR are indicated as black arrows. ITR, inverted terminal repeats. pA, polyadenylation signal. B, Left panel, RT-PCR from retinas isolated four weeks post-injection from C57Bl/6J mice. Right panel, sequence identity of the detected bands determined by sequencing. C, Confocal images of immunostained retinal slices harvested four weeks post-injection. Minigene-born protein was detected using an anti-myc antibody. The anti-CNGB1a antibody was used as a rod outer segment marker. Cell nuclei were stained with Hoechst. Dashed rectangles mark the areas magnified in the right panel. Scale bar, 10 μm .

However, the c.620T>G mutant protein was not detectable in the outer segments but was mainly localized in the inner segments of the photoreceptors with some weak signal in the cell soma (Figure 12C). Collectively, these findings demonstrate that c.620T>G, in contrast to c.620T>A, is a splicing mutation which leads to a truncated and mislocalized protein *in vitro* and *in vivo*.

4.1.4 Identification of Core Element Underlying the c.620T>G SAS Strength

Both, the *in vitro* as well as the *in vivo* experiments revealed that the c.620T>G SAS was used exclusively, which indicates that this splice site might be particularly strong. SASs have an intricate structure as they consist of several sequence elements (c.f. Figure 9A). Because of this complexity, the exact parameters defining the strength and functionality of SASs remain elusive so far rendering a reliable prediction of SAS efficiency rather difficult. To identify the sequence components defining the strength of the c.620T>G SAS, its performance was analyzed in an independent gene environment. For this purpose, the c.620T>G SAS including flanking elements from *RHO* exon 3 has been introduced into exon 3 of the human ribosomal protein S27 (*RPS27*) minigene (Figure 13A). The native *RPS27* exon 3 SAS was predicted to be distinctly stronger than the one of the c.620T>G SAS. Therefore, an insertion of different c.620T>G SAS sequences into this exon enables a more stringent analysis of their splice site strength in comparison to the native *RPS27* exon 3 SAS. Moreover, the native introns and coding exons of *RPS27* comprise only 1330 bp. This convenient size allowed for an expression of the entire gene from common expression vectors and the mRNA splicing analysis could thus be conducted in a largely native gene setting.

With this experimental design, a 26 bp long core sequence could be identified, which was virtually exclusively used as SAS in exon 3 of *RPS27* (Figure 13B). This core element, hereafter referred to as SAS_620, contains the SAS, a PPT and a 7 bp sequence element upstream of the PPT. Intriguingly, further deletions or insertions 5' or 3' of the SAS_620 reduced the splicing efficiency suggesting the presence and a complex interplay of different splice enhancer and silencer sequences in this region.

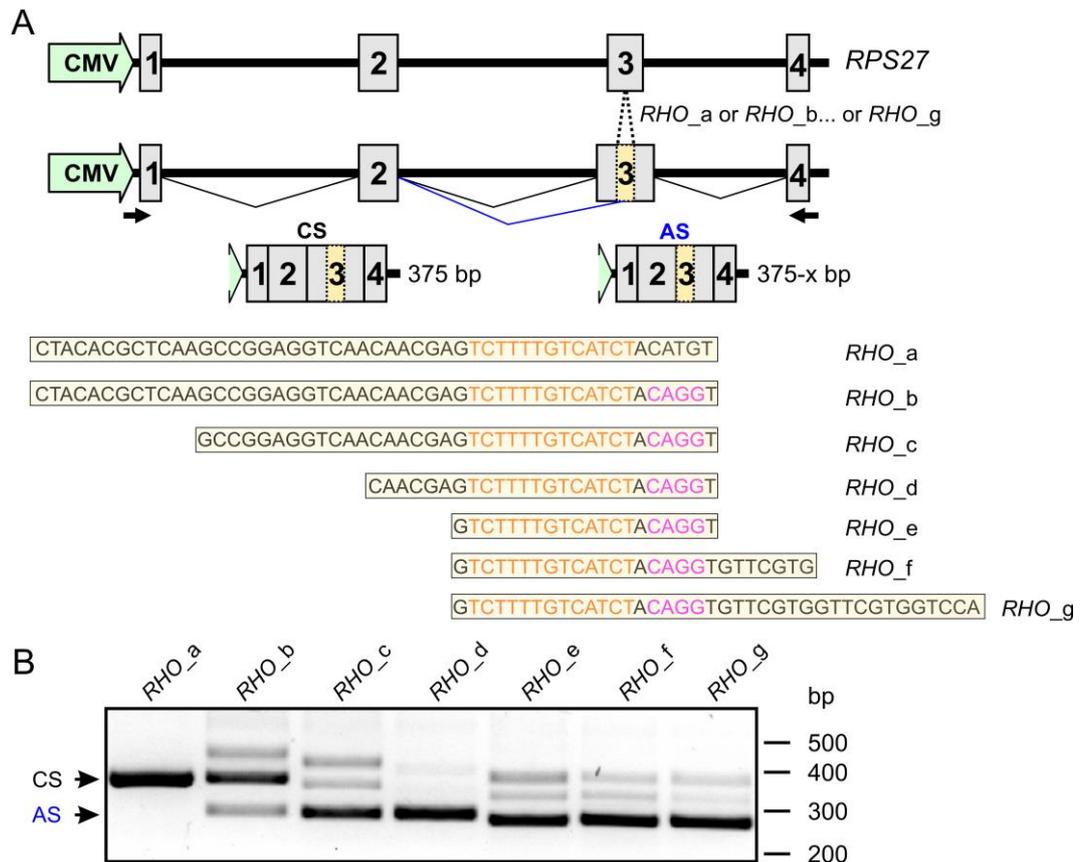


Figure 13 Splicing of c.620T>G SAS sequences in the *RPS27* minigene

A, Upper panel, Structure of the *RPS27* minigene containing all native exons and introns. The c.620T>G SAS including flanking sequences of different lengths (*RHO*_a – *RHO*_g) were introduced into the middle part of *RPS27* exon 3. Primers used for RT-PCR are shown as black arrows. Lower panel, Individual *RHO* sequences with the PPT depicted in orange and the SAS sequence depicted in pink. B, RT-PCR from HEK293 cells transiently transfected with the respective *RPS27* minigene. CS, correct splicing at the native exon 3 SAS; AS, aberrant splicing at the introduced c.620T>G SAS. The top band is a hybrid double strand comprising the correctly and the aberrantly spliced product.

4.1.5 Independence of SAS_620 Strength from Gene Environment

For potential usage in biotechnological applications, the strength and functionality of a SAS should be independent of the gene environment. To test whether SAS_620 fulfills this important requirement, the core element was introduced into four additional minigenes, i.e. the hemoglobin subunit theta 1 gene (*HBQ1*), the S100 calcium binding protein A12 gene (*S100A12*), the clarin-1 gene (*CLRN1*), and the cyclic nucleotide gated channel beta 1 gene (*CNGB1*) (Figure 14A).

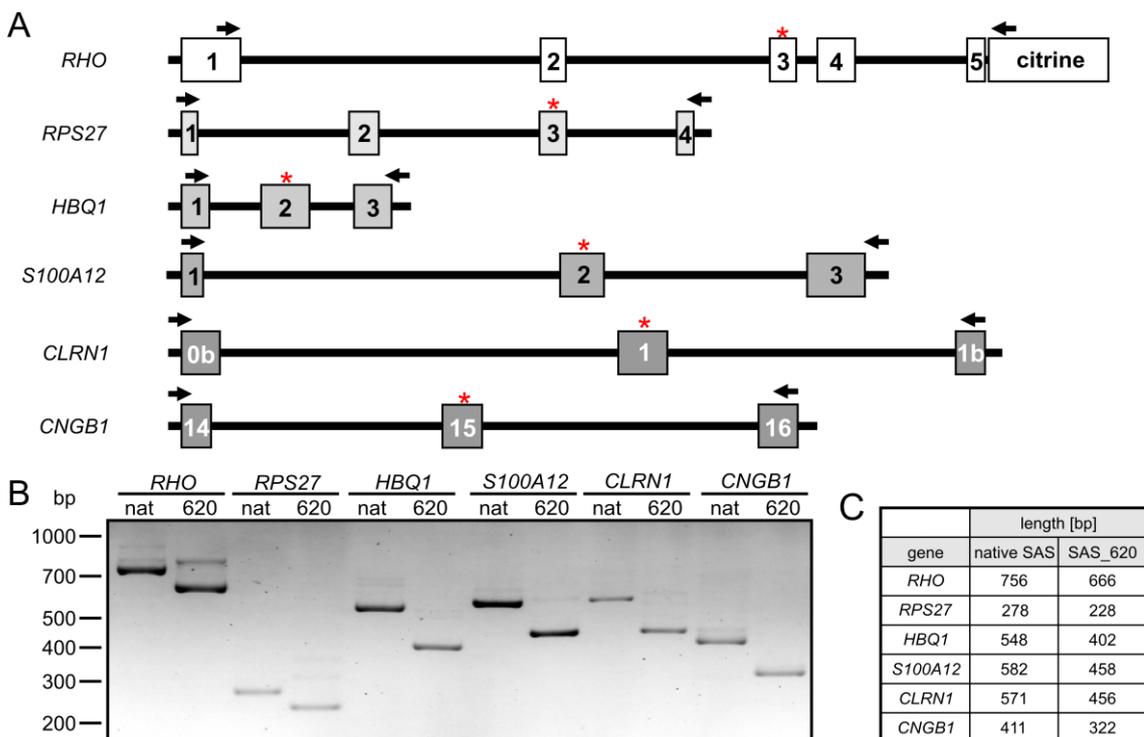


Figure 14 Splicing of SAS_620 in independent gene environments

A, The core element SAS_620 was introduced into the middle exons of minigenes originating from five unrelated genes (*RPS27*, *HBQ1*, *S100A12*, *CLRN1* and *CNGB1*) and splicing was analyzed via RT-PCR using the indicated primer pairs (black arrows). The position of the SAS_620 is marked with a red asterisk. B, RT-PCR of HEK293 cells transiently transfected with the respective minigene. nat, minigene containing the native splice site only. 620, minigene containing the inserted SAS_620 in addition to the native splice site. C, Table indicating the length of the spliced transcripts for usage of the native SAS vs. SAS_620.

The strengths of the respective splice sites competing with SAS_620, i.e. the native splice sites of the exons, in which SAS_620 was inserted, were predicted using different splicing prediction software tools (Table 13). The strengths of these native SASs varied markedly from gene to gene, but some variation was also present within

Results

the strengths predicted for the same splice site depending on the used software. Nevertheless, SAS_620 was recognized by the splicing machinery in all cases and appears to be used exclusively independent of the gene environment or the predicted splice site strength (Figure 14B, C).

Table 13 Predicted SAS strengths using different prediction software tools

SAS strength ranges from 0 to 100 for NNSplice and HSF with 100 indicating highest possible strength. ASSP calculates scores using position specific score matrices, where higher values indicate higher scores.

SAS	NNSplice	HSF	ASSP
<i>RHO_E3</i>	95	90	9.20
<i>RHO_SAS_620</i>	91	92	11.48
<i>RPS27_E3</i>	99	96	12.71
<i>HBQ1_E2</i>	92	95	9.44
<i>S100A12_E2</i>	97	86	9.57
<i>CLRN1_E1</i>	86	90	9.40
<i>CNGB1_E15</i>	77	85	3.96

To test whether the strength of SAS_620 is depending on the cell type, the respective minigenes were also transiently expressed in mouse embryonic fibroblasts (MEFs). Again, SAS_620 is preferred over the native SAS and appears to be used almost exclusively in all cases (Figure 15). Collectively, these findings suggest that SAS_620 is an exceptionally strong and efficient splicing element.

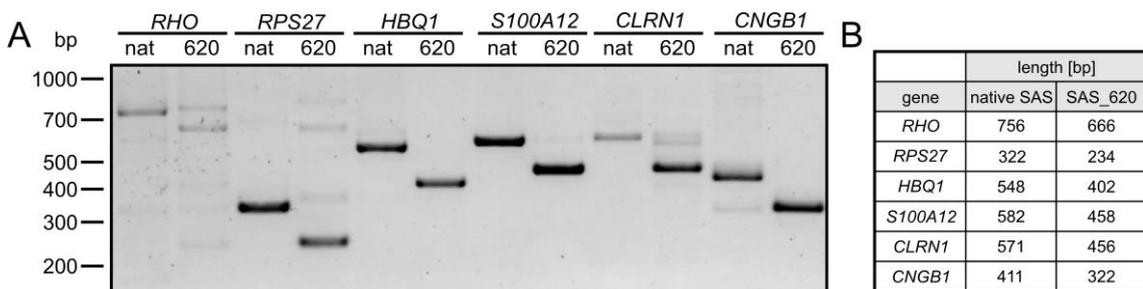


Figure 15 Splicing of SAS_620 in mouse embryonic fibroblasts

A, The minigenes containing the SAS_620 (*RPS27*, *HBQ1*, *S100A12*, *CLRN1* and *CNGB1*, cf. Figure 14) were transiently expressed in mouse embryonic fibroblasts and splicing was analyzed via RT-PCR. nat, minigene containing the native splice site only. 620, minigene containing the inserted SAS_620. B, Table indicating the length of the single spliced transcripts resulting from the usage of the native SAS vs. SAS_620.

In the experiments conducted so far, the SAS_620 was always tested in an exonic environment. However, it is conceivable that the native splice site upstream of SAS_620 might be recognized first and that its presence or the presence of other exon-defining splice elements could be contributing to the remarkable performance of SAS_620. This possibility would limit a broad usage of SAS_620 for biotechnological applications. Therefore, the functionality of SAS_620 was tested in an intronic region, where no exon-defining elements or upstream splice sites are expected to exist. For this purpose, the core element was introduced into the WT *RHO* minigene upstream of the native exon 3 SAS. Importantly, the SAS_620 was also exclusively utilized by the splicing machinery in this intronic context (Figure 16).

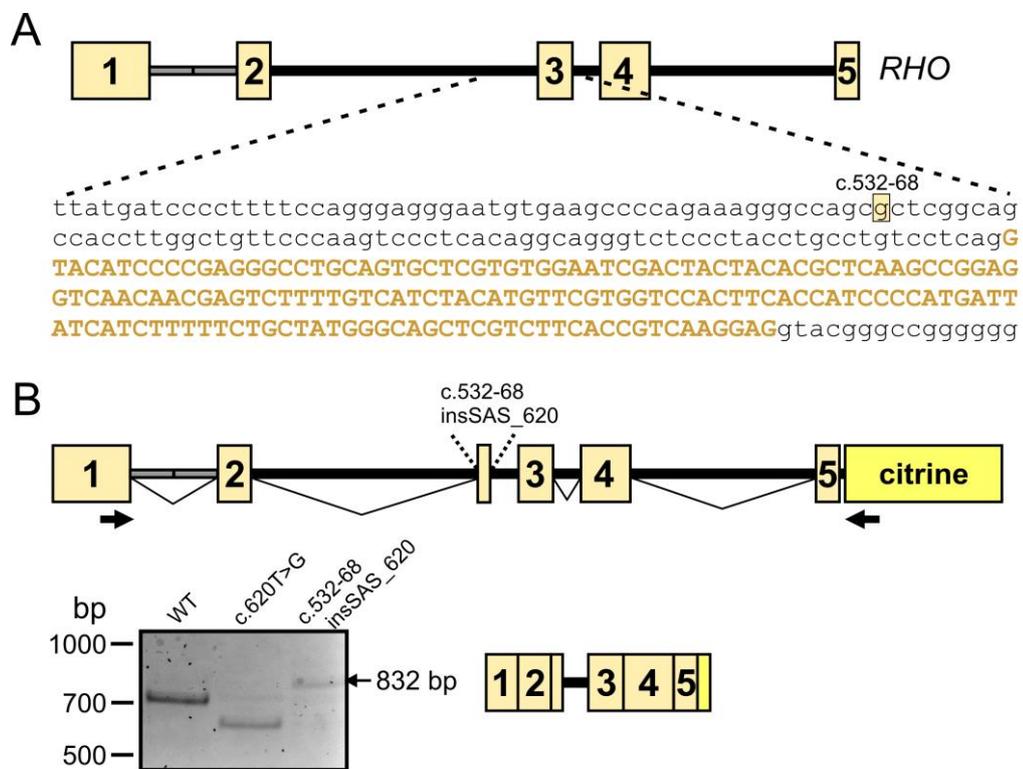


Figure 16 Introduction of SAS_620 into *RHO* intron 2

A, SAS_620 was introduced into the WT *RHO* minigene upstream of the native exon 3 SAS at the indicated position (c.532-68). B, Upper panel, HEK293 cells were transiently transfected with the WT, the c.620T>G and the c.532-68insSAS_620 minigene. Lower panel, Splicing was analyzed via RT-PCR using the indicated primers. The sequence identity of the splice product for the c. 532-68insSAS_620 minigene is schematically depicted on the right.

4.1.6 High Flexibility of the c.620T>G SAS in Branch Point Choice

As shown in Figure 13A, the SAS₆₂₀ consists of a SAS, a PPT and an additional 7 bp sequence upstream, the deletion of which is impairing SAS₆₂₀ recognition. Therefore, it seems reasonable to assume that this 7 bp sequence might contain a very potent branch point potentially explaining the universal and efficient recognition of SAS₆₂₀ regardless of the gene environment. However, there were no potent branch points predicted within this sequence. Instead, other sequences positioned up to 40 bp upstream of the c.620T>G SAS were predicted to serve as branch points (Figure 17). Nevertheless, when simultaneously mutating all predicted branch point nucleotides and all potential branch point adenines contained in the 7 bp sequence upstream of the PPT, splicing could not be altered for the c.620T>G mutant (Figure 17B). This finding indicates that the branch point(s) are either located elsewhere or that the c.620T>G SAS is highly flexible in branch point choice.

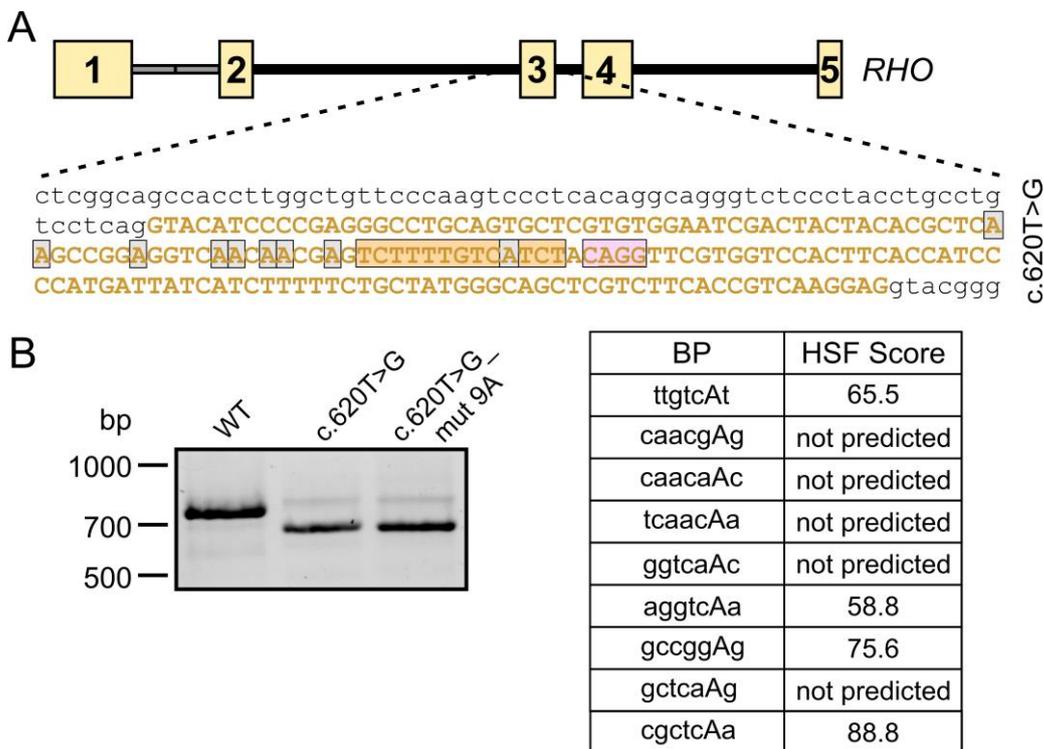


Figure 17 Mutation of potential branch points in RHO c.620T>G exon 3

A, Sequence of the RHO c.620T>G minigene. Potential branch point nucleotides within the 7 bp sequence upstream of the PPT or predicted by the HSF software are marked with gray boxes. B, Left panel, RT-PCR analysis of the transiently expressed RHO WT minigene and two variants of the c.620T>G minigene. In the c.620T>G_{mut9A} minigene, all potential branch point adenines located up to 40 bp

upstream of the c.620T>G SAS have been replaced by guanines or cytosines. Right panel, HSF scores for all predicted branch points or for potential branch point sequences present in the 7 bp sequence of SAS_620. Branch point nucleotides are shown as capital letters. BP, branch point.

To identify the branch point(s) utilized for splicing at the c.620T>G SAS more directly, a nested lariat RT-PCR was performed using HEK293 cells transiently expressing the *RHO* c.620T>G minigene (Figure 18A). HEK293 cells transfected with the *RHO* WT minigene served as a reference. When performing the lariat RT-PCR, one band was obtained for the WT and for the mutant minigene, respectively, both differing in their size. Both bands appeared somewhat diffuse suggesting that they contained lariats of varying size. To identify the single sequences contained in these diffuse bands, the lariat RT-PCR products were subcloned into a TOPO vector and the obtained clones were analyzed individually via sequencing. When investigating the *RHO* WT lariats, two major branch points (used in 42 % and 33 % of cases) and three minor branch points (8 % of cases each) were identified (Figure 18B). The two major branch points closely resembled the consensus sequence and yielded high prediction scores (Table 14). All *RHO* WT branch points were found 57 – 184 bp upstream of the intron-exon junction. As more than 90 % of human branch points are found within 50 bp upstream of the SAS sequence, mRNA splicing of *RHO* exon 3 seems to be somewhat unconventional (Corvelo et al., 2010).

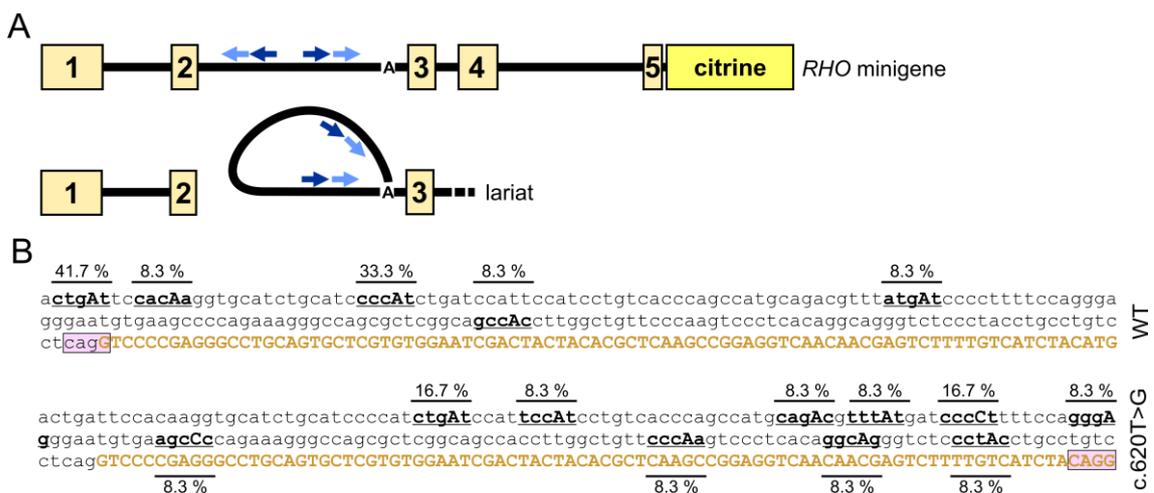


Figure 18 Identification of branch points via nested lariat RT-PCR

A, Upper panel, *RHO* minigene expressed for nested lariat RT-PCR. The used primer pairs are depicted as light and dark blue arrows. Lower panel, Minigene after lariat formation via branching at the indicated branch point nucleotide. B, Partial sequences of the *RHO* WT and c.620T>G exon 3 (brown uppercase) and intron 2

(black lowercase). The respective SAS is indicated via a pink box. Identified branch point sequences are depicted underlined and bold with the branch point nucleotide shown as a capital letter. The frequency of the identified branch point is shown above or below the respective sequence.

Table 14 Detected branch points and corresponding prediction scores

<i>RHO</i> WT minigene			
Branch point	Relative position	HSF score	Frequency of usage
cactgAt	-184 bp	84.3	41.6 %
tccacAa	-177 bp	76.8	8.3 %
tccccAt	-158 bp	85.6	33.3 %
ttatgAt	-113 bp	60.6	8.3 %
cagccAc	-57 bp	58.4	8.3 %
<i>RHO</i> c.620T>G minigene			
atcgAt	-245 bp	83.6	16.7 %
attccAt	-236 bp	55.6	8.3 %
tgCagAc	-213 bp	73.9	8.3 %
cgtttAt	-208 bp	55.9	8.3 %
atcccCt	-200 bp	not predicted	16.7 %
cagggAg	-189 bp	not predicted	8.3 %
gaagCc	-174 bp	not predicted	8.3 %
ttcccAa	-132 bp	79.5	8.3 %
caggcAg	-118 bp	not predicted	8.3 %
tcctAc	-107 bp	77.6	8.3 %

Nevertheless, the branch point profile obtained for *RHO* c.620T>G was remarkably different when compared to the WT minigene. First, various branch points have been identified for the c.620T>G mutant. However, no major branch point(s) could be detected and none of them were identical with the ones obtained for the WT minigene. Second, the branch points were located even further upstream from the used SAS, i.e. 107 – 245 bp corresponding to 21 – 159 bp upstream of the native exon 3 SAS. Third, almost half of the detected branch points hardly resembled the consensus sequence and accordingly were not predicted by the Human Splicing Finder (HSF) splice prediction tool. All in all, these data suggest that the strength of SAS_620 does not originate from the presence of a contained, very potent branch point but might partially be caused by a high flexibility in branch point choice. This

could explain its unusually efficient performance and renders it a very attractive tool for biotechnological applications requiring efficient splicing.

4.2 DEVELOPMENT OF A NOVEL DUAL VECTOR TECHNOLOGY USING THE SAS_620 SPLICING ELEMENT

4.2.1 Reconstitution of a Fluorescent Reporter Using mRNA *Trans*-Splicing

For reconstitution of full-length coding sequences (CDSs) in the framework of dual vector technologies, DNA-based and protein-based reconstitution strategies have been employed so far. However, these approaches harbor limitations due to their low and unpredictable reconstitution efficiencies or their restricted applicability and potential safety issues. Therefore, a novel strategy aiming at reconstitution of CDSs at the mRNA level via mRNA *trans*-splicing has been explored in this study. For this purpose, I developed a fluorescent reporter-based assay, for which the CDS of the cyan fluorescent protein “cerulean” was split into two halves at c.154 (Figure 19A). Both parts were equipped with a splice donor (SDS) or acceptor site and a complementary binding domain (BD) and were expressed separately under the control of a CMV promoter. After transcription, base-pairing of the BDs brings the contained splicing elements into close proximity, which facilitates splicing in *trans* and results in a full-length (FL) mature mRNA (Figure 19B). As the SAS_620 should be a very suitable splicing element for such an application, it was used on the 3' vector to enable efficient *trans*-splicing. Furthermore, a SDS yielding the highest possible prediction scores was used for the subsequent experiments. Due to their low complexity, SDSs are well characterized and the prediction of their strength is well-matched on their experimental performance. For initial evaluation, a 100 bp sequence originating from the human rhodopsin intron 2 was chosen as a BD.

HEK293 cells were transiently co-transfected with each CDS half and the presence of cerulean fluorescence was evaluated via confocal live cell imaging. A construct containing both CDS halves in *cis* with an intervening artificial intron equipped with the same splicing elements was used as *cis*-splicing reference control (*cis*-ctrl). When the two halves were transfected separately, no fluorescence could be detected. By contrast, when co-transfecting both constructs, some fluorescent cells could be observed indicating successful *trans*-splicing and reconstitution of the FL

cerulean coding sequence (Figure 19C). As a next step, different BDs were tested in order to optimize the *trans*-splicing efficiency. For this purpose, sequences differing in their length and their position in the 5' and the 3' region of *RHO* intron 2 were used (Figure 20A). Moreover, artificial BDs were created by fusing sequences originating from the 5' and the 3' end of the intron. A list of all BD sequences tested is displayed in section 7.2.

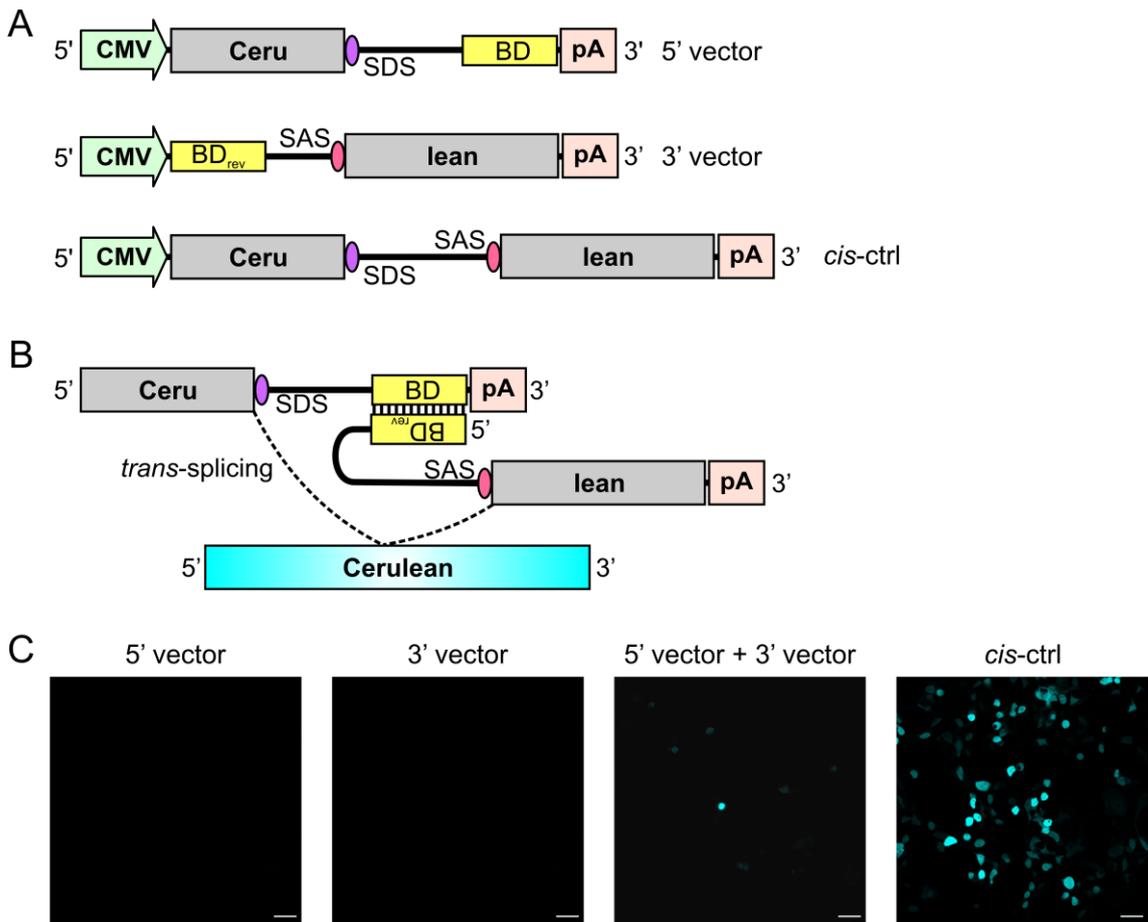


Figure 19 Schematic illustration of the reporter assay used for evaluation of gene reconstitution via mRNA trans-splicing

A, Reporter genes used for evaluation of the mRNA *trans*-splicing dual vector approach. The coding sequence of the fluorophore “cerulean” was split into two halves and equipped with all elements necessary for *trans*-splicing, i.e. complementary binding domains (BD and BD_{rev}, rev: reverse complement) and a splice donor (SDS) or acceptor site (SAS). A *cis*-spliced cerulean construct (*cis*-ctrl) containing an artificial intron served as a positive control. B, pre-mRNA from 5' vector- and 3' vector-constructs. After successful *trans*-splicing of both molecules, the FL cerulean mRNA is reconstituted. C, Confocal live imaging of HEK293 cells transiently transfected with the respective reporter gene(s). The presence of cerulean fluorescence indicates successful *cis*- or *trans*-splicing. Scale bar, 50 μ m.

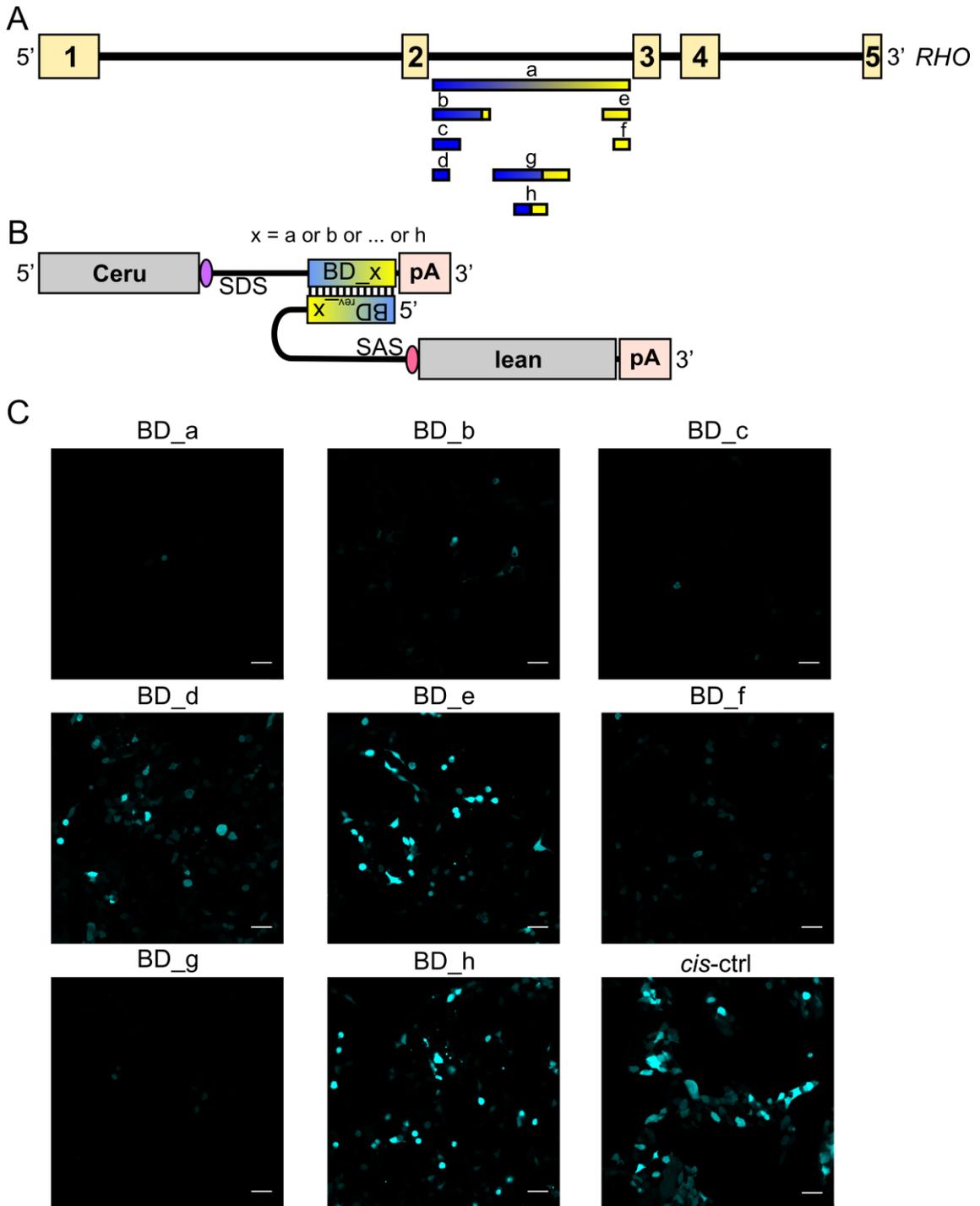


Figure 20 Binding domain optimization for increase in reconstitution efficiency

A, Schematic illustration of *RHO* intron 2 sequences tested as BDs (a – h). B, BD-mediated binding of both pre-mRNA molecules before *trans*-splicing. C, Confocal live cell imaging of HEK293 cells co-transfected with the 5' vector- and the 3' vector-constructs containing the respective binding domains (BD_a – BD_h). *cis-ctrl*, *cis*-splicing control. Scale bar, 50 μ m.

When co-transfecting HEK293 cells with constructs containing the respective binding domains, a great variance in the amount of cerulean fluorescence could be observed indicating great variability in the corresponding *trans*-splicing efficiencies (Figure 20B, C). Consequently, the binding domain sequence seems to be an important element for efficient reconstitution. Next, the precision of the *trans*-splicing process was analyzed via RT-PCR. With this approach, no unspecific bands, which might arise from usage of alternative splice sites present in the expression cassette, could be detected (Figure 21A, B). Accordingly, sequencing revealed that all constructs were spliced exactly at the splicing elements as intended (Figure 21C). Collectively, these findings suggest that *trans*-splicing is a seamless and precise process reliably restoring the split reading frame.

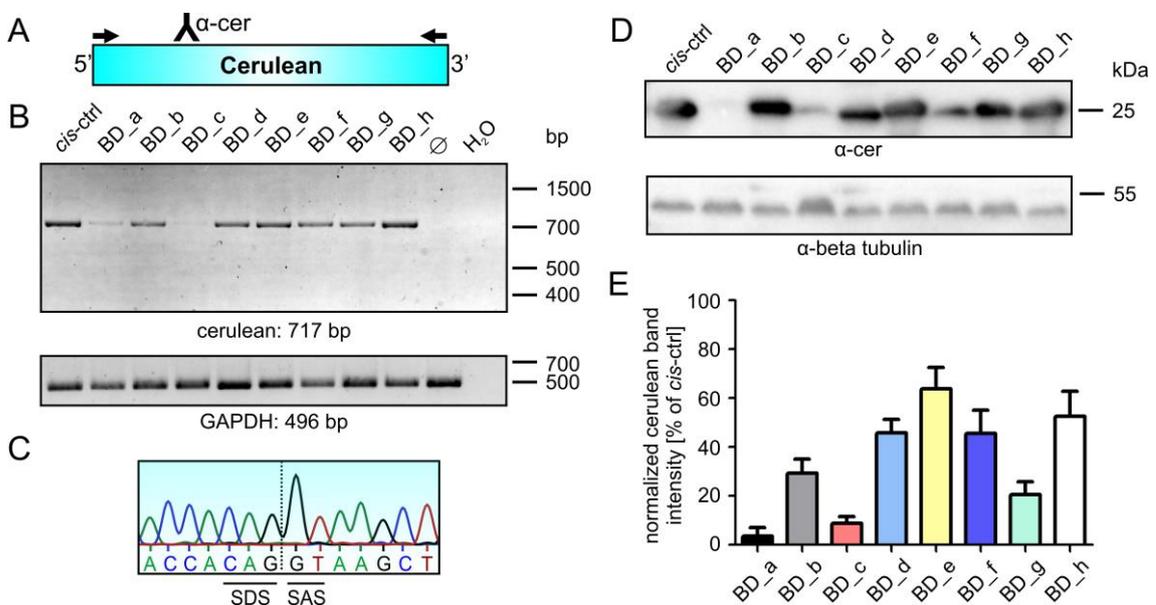


Figure 21 Evaluation of *trans*-splicing precision and reconstitution efficiency

A, FL cerulean mRNA was detected via RT-PCR using the indicated primer pair (black arrows). Cerulean protein was detected by western blotting using an anti-GFP antibody capable of detecting cerulean (α -cer). B, RT-PCR from HEK293 cells co-transfected with the respective 5' and 3' vector containing BD_a – BD_h. \emptyset , untransfected cells. *GAPDH* was amplified as a loading control. C, Representative sequencing result of the reconstituted cerulean. D, Western blot of protein lysates from co-transfections using the single BD combinations (BD_a – BD_h). Beta tubulin served as a loading control. E, Reconstitution efficiencies were determined via a ratiometric analysis of the cerulean protein band intensities in relation to the *cis*-ctrl ($n = 3 - 8$). All protein bands were normalized to beta tubulin prior to quantification. Reconstitution efficiency values for each BD can be found in 7.3.

For quantification of the reconstitution efficiencies, the amount of cerulean protein was determined via western blotting. When compared to the *cis*-ctrl, reconstitution efficiencies of $62.3 \% \pm 7.9 \%$ could be obtained (Figure 21D, E). The BDs yielding the highest efficiencies are both composed of approximately 100 bp. One of them (BD_h) is an artificially created fusion domain consisting of a 5' and a 3' intronic part, whereas the second one (BD_e) is representing a region in the 3' part of *RHO* intron 2. The calculated reconstitution efficiencies are exceeding the best *in vitro* efficiencies reported for DNA-based reconstitution so far (Carvalho et al., 2017) and are comparable to the *in vitro* efficiencies obtained for split intein based reconstitution at the protein level (Chew et al., 2016, Tornabene et al., 2019)

4.2.2 Impact of Splice Acceptor Site vs. Binding Domain on Reconstitution Efficiency

As shown in 4.2.1, the BD length and sequence represent key determinants for the reconstitution efficiency. The BDs most likely influence the tight binding and, potentially, the folding of the mRNAs, but they are not expected to directly facilitate the efficiency or precision of the subsequent splicing process. As mentioned above, SDSs are well characterized and the prediction of their strength is well-matched on their experimental performance. Consequently, there is no apparent need for optimization of this splice site in the framework of the split fluorophore reconstitution assay. By contrast, due to their complexity, the strength of the SASs cannot be reliably predicted (cf. 4.1.4 and 4.1.5). The results shown in 4.1 suggest that SAS_620 is an exceptionally strong splice acceptor site. Given that the splice site strength can impact on the reconstitution efficiency of the split fluorophore assay, SAS_620 should lead to higher values when compared to other SASs. To analyze this, the reconstitution efficiencies in presence of SAS_620 or two other SASs, i.e. the native SAS of *RHO* exon 3 (S3, cf. Figure 14) and a hybrid SAS (S2) which was created by replacing the PPT of SAS_620 with the PPT of the native *RHO* exon 3 SAS, were compared. Moreover, each SAS was combined with each of three different binding domains: a strong (BD_e, B1) and a weak (BD_c, B3) binding domain originating from this study and a published BD sequence (PTM1, B2) obtained from *RHO* intron 1. The last sequence was shown to yield high efficiencies for the repair of mutated *RHO* transcripts via spliceosome-mediated mRNA *trans*-splicing (Berger et al.,

Results

2015). All combinations were analyzed via confocal live cell imaging, RT-PCR and western blotting (Figure 22).

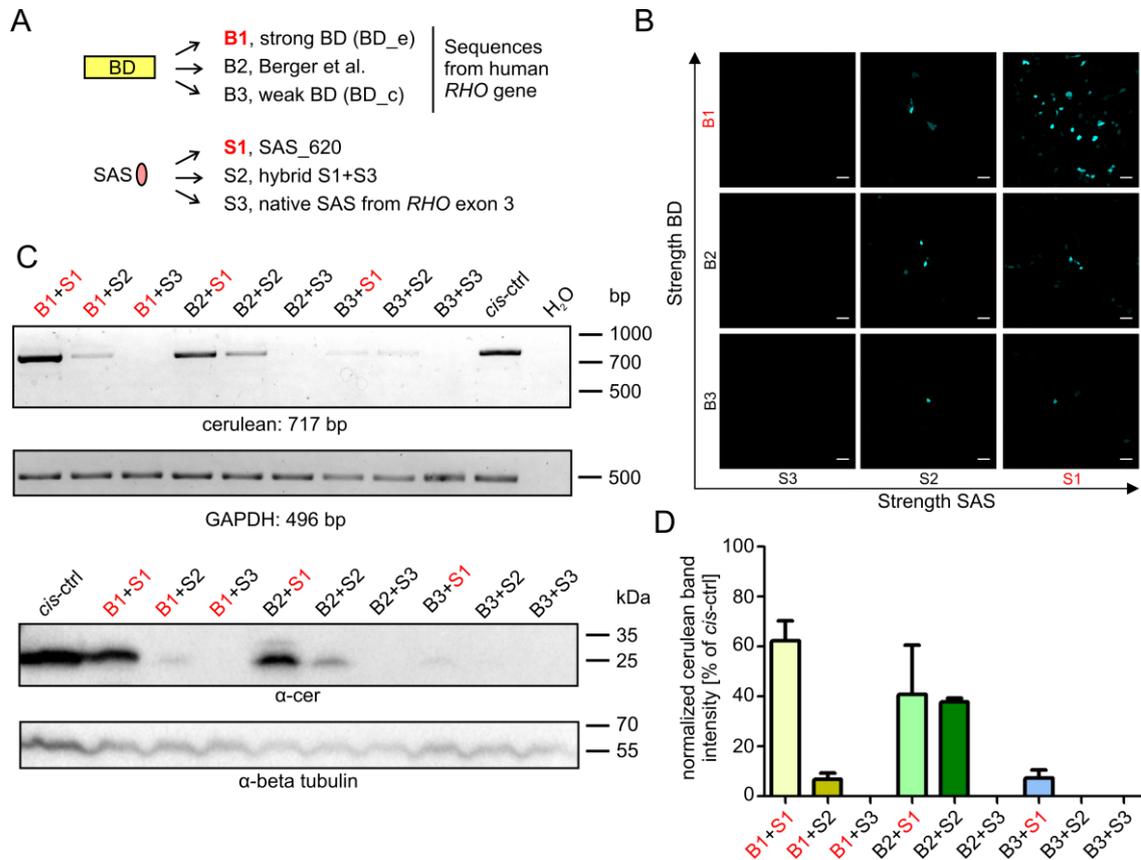


Figure 22 Impact of SAS vs. BD on reconstitution efficiency

A, BDs and SASs tested in this experiment. All BD sequences originate from the human *RHO* gene. BDs and SASs expected to yield high efficiencies are highlighted in red. B, Confocal live images of HEK293 cells transfected with constructs containing different combinations of the three BDs and the three SASs. The strength of the respective BD and SAS is indicated. Scale bar, 50 μ m. C, Upper panel, RT-PCR of the different BD and SAS combinations. Lower panel, Western blot of the different BD and SAS combinations. *GAPDH* and beta tubulin served as loading controls. D, Quantification of the reconstitution efficiency by the ratiometric analysis of the cerulean protein bands in relation to the *cis*-ctrl (n = 3 – 8). All protein bands were normalized to beta tubulin prior to quantification. Reconstitution efficiency values for each combination can be found in 7.3.

This experiment resulted in several important findings. First, it revealed that both, the BD and the SAS are key components determining reconstitution efficiencies. Second, the most potent BD identified in this study (BD_e, B1) outperforms the published *RHO* binding domain (B2). Third, the SAS strength appears to influence the reconstitution efficiencies even more potently than the BD, as the weakest BD combined with the strongest SAS (B3+S1) still results in detectable *trans*-splicing,

whereas the strongest BD combined with the weakest SAS (B1+S3) does not lead to a detectable reconstitution of the CDS.

4.2.3 Reconstitution via mRNA *Trans*-Splicing in Different Cell Types

Splicing can vary between different cell types, as it is a highly tissue-specific process (Arias et al., 2010). For a broad application of mRNA *trans*-splicing dual vectors, high reconstitution efficiencies need to be achieved irrespective of the target tissue. To address the cell type dependency of the split fluorophore reconstitution assay, the best BDs (BD_e and BD_h) in combination with the best SAS (S1) were tested for their reconstitution efficiency in 661W cells and in MEFs. As controls, a weak BD (BD_g) combined with a strong SAS (S1) and a strong BD (BD_e) combined with a weak SAS (S3) were used.

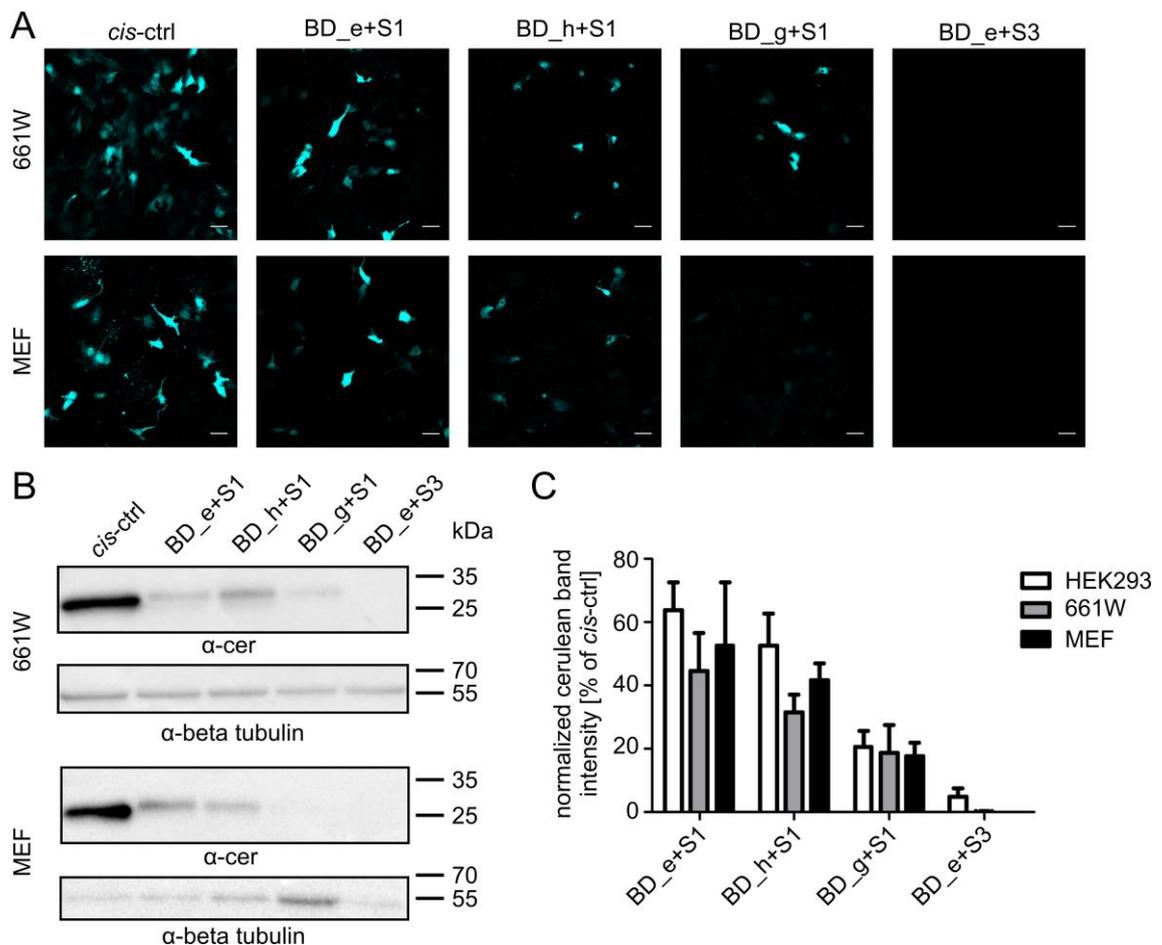


Figure 23 Reconstitution via mRNA *trans*-splicing in different cell types

A, Confocal live cell images of 661W cells and MEFs transiently transfected with constructs containing different BDs (BD_e, BD_h or BD_g) and SASs (S1 or S3). Scale bar, 50 μm. B, Western blots of protein lysates obtained from the transfected cells.

C, Quantification of the reconstitution efficiency based on the ratiometric analysis of the western blot band intensities ($n = 3 - 8$). For statistical testing, one-way ANOVA was applied. Reconstitution efficiency values for each combination can be found in 7.3.

Successful reconstitution could be observed via confocal live cell imaging and via western blotting for both cell types (Figure 23A, B). BD_e and BD_h in combination with S1 yielded high reconstitution efficiencies, i.e. $44.7 \% \pm 12.0 \%$ and $31.5 \% \pm 6.6 \%$ in 661W cells and $52.6 \% \pm 20.0 \%$ and $41.7 \% \pm 5.4 \%$ in MEFs (Figure 23C). No significant difference in *trans*-splicing efficiencies could be detected when comparing all three cell types. This suggests that the mRNA *trans*-splicing approach is cell type independent.

4.2.4 Investigation of mRNA *Trans*-Splicing-Based rAAV Dual Vectors *in Vivo*

The most potent application of the mRNA *trans*-splicing-based assay evaluated in the previous sections is the reconstitution of large genes in the framework of dual rAAV vectors. Consequently, the mRNA *trans*-splicing approach was tested in the mouse retina using rAAVs. For this purpose, a slightly modified version of the split fluorophore assay was used. To control the rAAV vector-born expression in the cells transduced with the single viruses, both dual rAAV vector cassettes were equipped with a fluorophore sequence, i.e. a citrine at the 5' end of the CDS for the 5' vector and a mCherry at the 3' end of the CDS for the 3' vector. One of the BDs yielding the highest reconstitution efficiencies *in vitro*, i.e. BD_h, was used *in vivo* (Figure 24A). In this experimental setup, cerulean fluorescence should be present in cells expressing citrine as well as mCherry. The titer-matched viruses were injected subretinally on P21 into WT C57Bl6/J mice. After harvesting the retinas two weeks post-injection, a solid fluorophore expression could be detected in the RPE (Figure 24B). Moreover, cerulean fluorescence could be observed in all areas where citrine and mCherry were expressed indicating successful mRNA *trans*-splicing in cells co-transduced with both AAVs. However, citrine and cerulean possess partially overlapping excitation and emission spectra. To exclude the possibility of the cerulean fluorescence being an artefact caused by e.g. bleed-through of citrine or mCherry, both fluorophores were bleached selectively in a small area of the RPE by exciting the fluorophores with a high intensity 514 nm laser. With this procedure,

the citrine and mCherry fluorescence could be removed completely (Figure 24C). Nevertheless, the cerulean fluorescence remained unchanged indicating that it originates solely from *trans*-spliced cerulean mRNA. This experiment provides proof of principle for the usability of mRNA *trans*-splicing for reconstitution of genes expressed from AAVs *in vivo*.

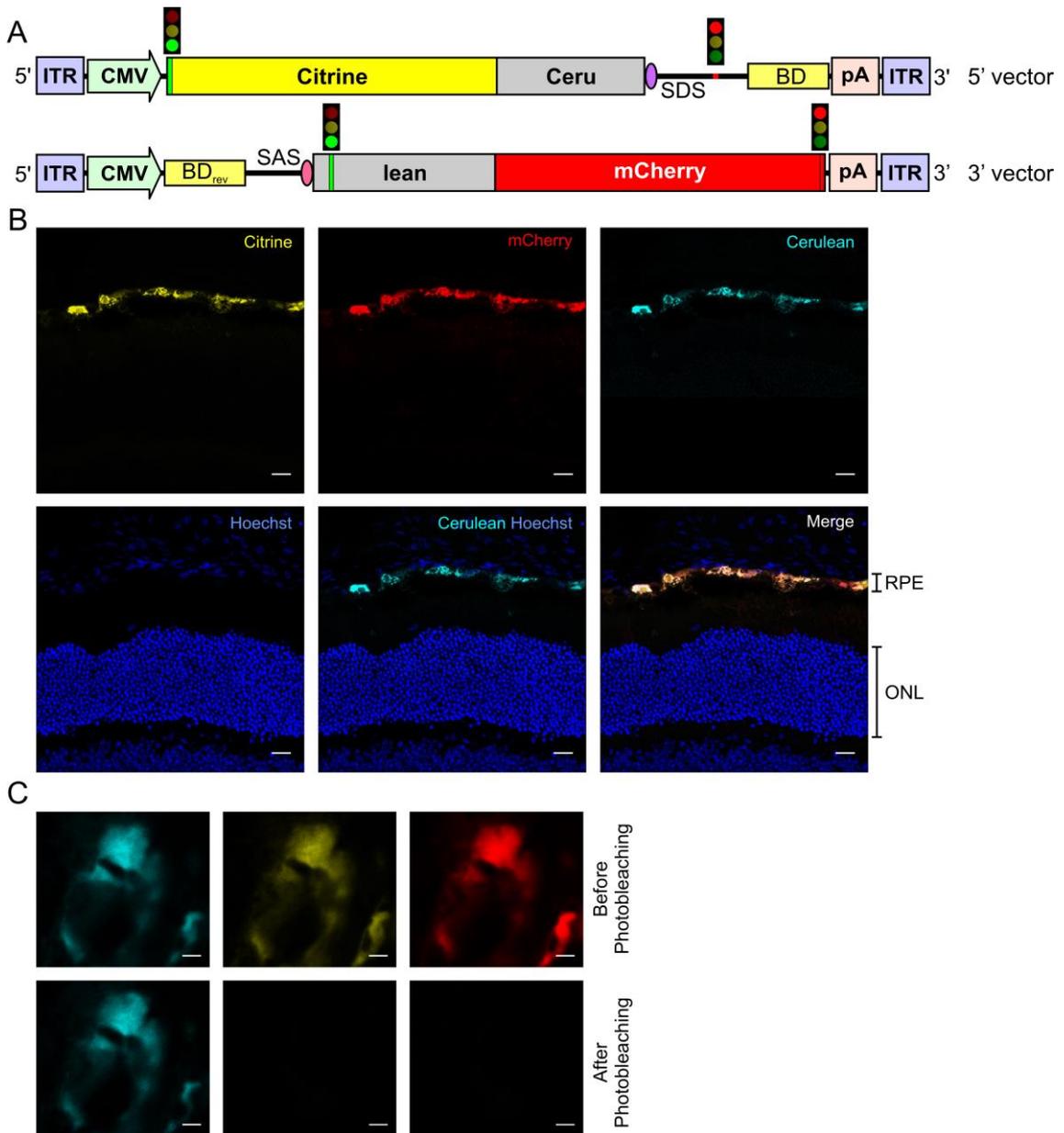


Figure 24 mRNA trans-splicing rAAV dual vector approach in vivo

A, 5' vector- and 3' vector-constructs used for *in vivo* expression. As an expression control, a citrine and a mCherry sequence was fused 5' of the cerulean 5' CDS and 3' of the cerulean 3' CDS, respectively. Start and stop codons are indicated by traffic lights. ITR, inverted terminal repeats. B, Representative confocal images of retinal sections harvested two weeks post-injection expressing constructs containing BD_h. Fluorophore expression is present in the retinal pigment epithelium (RPE). ONL,

outer nuclear layer. Scale bar, 20 μm . C, Confocal images of RPE cells before (upper panel) and after (lower panel) selective photobleaching of citrine and mCherry fluorophores using a 514 nm laser. Scale bar, 2 μm .

4.2.5 Identification of Potent Binding Domain Suitable for Human Gene Therapy

So far, all BD sequences were obtained from human intronic regions. When used for human gene therapy, they thus could potentially also bind to endogenous mRNAs and induce *trans*-splicing with these transcripts. It is therefore necessary to find identify BDs which do not contain sequences homologous to the human genome. For this purpose, a random 100 bp sequence was taken from the bacterial *lacZ* gene and modified via random insertions, deletions and substitutions to obtain four sequences without any homology to the human genome (Figure 25A). When co-transfecting HEK293 cells with the 5' vector- and 3' vector-constructs containing the respective binding domains, very high cerulean reconstitution efficiencies of 78.3 % \pm 2.1 % could be observed for one of the BDs (BD_j, Figure 25B – D). BD_j was thus more efficient than the best performing human BDs shown in Figure 21 and was therefore used for preliminary experiments evaluating reconstitution of large genes via mRNA *trans*-splicing.

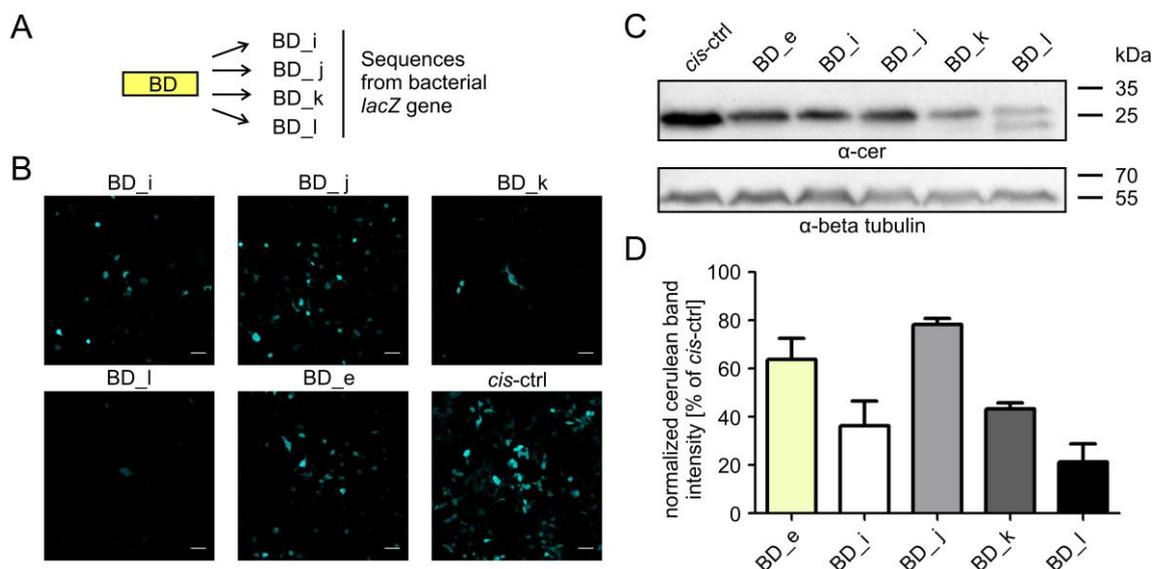


Figure 25 Identification of suitable BDs originating from *lacZ* gene

A, BDs taken from the bacterial *lacZ* gene and modified to possess no detectable homology to the human genome. B, Confocal live images of HEK293 cells transiently co-transfected with constructs containing the BDs shown in A. Scale bar, 50 μm . C,

Western blot obtained from transfected HEK293 cell lysates. Quantification of reconstitution efficiency based on the ratiometric analysis of the western blot band intensity ($n = 3 - 8$). BD_e efficiency served as a measure for the highest reconstitution obtained so far. Reconstitution efficiency values for each BD can be found in 7.3.

4.2.6 Proof of Principle for Reconstitution of Large Genes via mRNA *Trans*-Splicing

So far, only the cerulean split reporter assay has been used for evaluation of mRNA *trans*-splicing. A suitable candidate for investigating mRNA *trans*-splicing on a large gene is a recently developed novel tool for gene therapy: the transcriptional activator *SpCas9*-VPR, a catalytically inactive nuclease fused to the transcriptional activator domains VP64-p65-Rta (VPR). Due to its large size (5.8 kb), it needs to be delivered via dual vectors for *in vivo* applications.

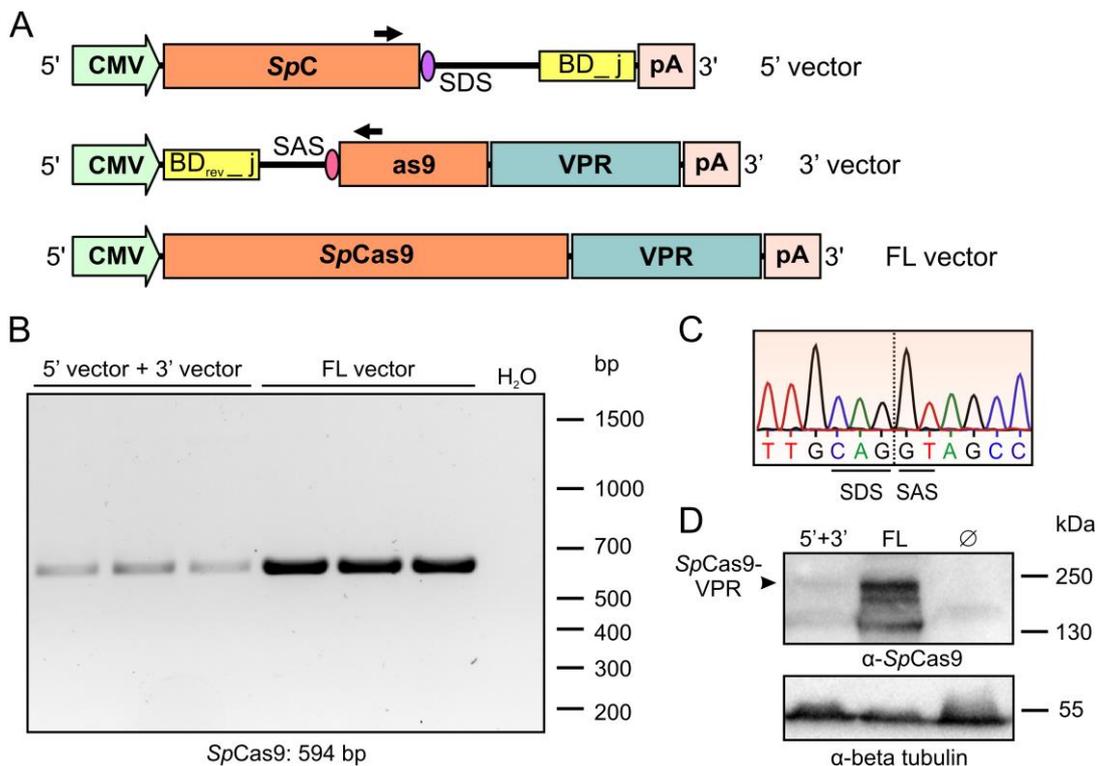


Figure 26 Reconstitution of *SpCas9*-VPR

A, Constructs designed for reconstitution of *SpCas9*-VPR. The full-length (FL) construct was used as a positive control. A junction-spanning primer pair was used for RT-PCR (black arrows). B, RT-PCR from three independent transfections of HEK293 cells with the respective construct(s). C, Sequencing result of the reconstituted *SpCas9*-VPR product. D, Western blot of *SpCas9*-VPR. Ø, untransfected cells.

The CDS of *SpCas9-VPR* was split into two halves at c.2185 and the two CDS halves were equipped with BD_j and the appropriate splicing elements, i.e. the SDS or the SAS (Figure 26A). A FL *SpCas9-VPR* construct was used as a positive control. The RT-PCR from HEK293 cells co-transfected with the split constructs revealed that *SpCas9-VPR* mRNA was reconstituted successfully and that no unwanted side-products were generated (Figure 26B). Sequencing of the PCR product confirmed an accurate restoration of the reading frame (Figure 26C). Moreover, FL *SpCas9-VPR* protein could be detected via western blotting (Figure 26D). The reconstitution efficiencies of *SpCas9-VPR* were $13.2 \% \pm 0.9 \%$. This provides a proof of principle for the applicability of the mRNA *trans*-splicing approach for reconstitution of large genes.

5 DISCUSSION

5.1 IMPLICATIONS OF C.620T>G MUTATION CREATING A SPLICE ACCEPTOR SITE WITH UNIQUE PROPERTIES

In this study, strong evidence is provided that the deep exonic adRP mutation c.620T>G is a splicing mutation which has been misclassified as a missense mutation since its discovery. For the development of adequate treatment strategies, it is important to understand the molecular mechanism underlying a disease-linked mutation. Exonic point mutations are generally expected to affect key amino acid residues and are therefore categorized as missense mutations. However, as already proposed in previous studies (Becirovic et al., 2016b, Soukarieh et al., 2016), the findings of this study suggest that an underestimated number of such mutations might instead impact on mRNA splicing. This conclusion is of particular importance, as mRNA splicing might be a so far misjudged molecular factor markedly influencing the pathogenicity of various disease-causing mutations. Moreover, this finding shows that the effect of exonic mutations at the mRNA level should be investigated before their final classification as missense mutation. Such an examination can help to avoid the elaborate analysis of mutant proteins which do not naturally exist. Additionally, such investigations can help to potentially clarify other puzzling differences in disease courses caused by similar mutations at the molecular level. Lastly, this finding could improve the understanding of genotype-phenotype correlations thereby potentially refining disease course predictions in affected patients in the framework of genetic counselling.

The results obtained in this study indicate that the adRP phenotype severity in patients carrying mutations in the coding region of Met207 and in surrounding areas is correlating with the number of affected amino acids. Patients affected by the c.620T>A mutation, in which only one amino acid, i.e. Met207, is altered, exhibit a mild, late-onset adRP phenotype (Audo et al., 2010a, Audo et al., 2010b). By contrast, the c.614-622del mutation causing the deletion of three aa leads to an intermediate phenotype (Maubaret et al., 2012) and the c.620T>G mutation resulting in a deletion of 30 aa correlates with a severe, early-onset disease variant. However, the exact molecular mechanism underlying the severity of the c.620T>G mutation remains to

be elucidated. One possible mechanism already shown for other severe rhodopsin mutants, e.g. for P23H, is that c.620T>G might result in protein misfolding, aggregation and gain-of-function effects (Athanasίου et al., 2018).

Met207 has been found to be an important aa residue in the ligand binding pocket of GPCRs (Moitra et al., 2012). Therefore, it has been assumed that the main reason for the puzzling adRP phenotype difference associated with the c.620T>G and the c.620T>A mutations originates from their differential impact on binding of the ligand, i.e. 11-*cis*-retinal, in rhodopsin. However, a recent study failed to provide conclusive evidence supporting this hypothesis (Audo et al., 2010a). Here, I show that a completely different molecular mechanism is underlying the differential disease courses of these mutations. Moreover, it was demonstrated that the Met207Arg protein does not naturally exist rendering an investigation of the impact of this aa exchange on ligand binding obsolete. This finding again highlights the importance of examining the impact of disease-linked mutations on mRNA splicing.

Another interesting finding of this study is the fact that the c.620T>G mutation is creating an exceptionally strong SAS, as it is exclusively used independent of the gene environments and cell types tested herein. Furthermore, I could identify a 26 bp sequence, termed SAS_620, as the core element responsible for this strength. This sequence could be a suitable tool for biotechnological applications requiring strong splicing elements like spliceosome-mediated mRNA *trans*-splicing (SMaRT) (Berger et al., 2016), mirtrons (Curtis et al., 2012) or gene trapping (Stanford et al., 2001). The exact mechanism defining the extraordinary strength of SAS_620 remains to be investigated. One possible explanation is the presence of efficient intronic splice enhancer (ISE) motifs. A recent study focused on characterizing such human ISE elements and defined six distinct groups (A – F) (Wang et al., 2012). Indeed, two of these ISE motives are present in the first 7 bp of the SAS_620, i.e. AACG (group F) and CGAG (group D). Another interesting characteristic of the c.620T>G SAS is its flexibility in branch point choice. The underlying molecular mechanism and sequence elements facilitating this process have yet to be elucidated. Generally, SAS strength can be predicted using different prediction software tools. Nevertheless, little is known about how accurately these predictions can be matched with experimental data. Here, no consistent correlation between predicted splice site

strength and experimental findings could be identified, indicating that current algorithms cannot reliably predict SAS performance yet. Based on the results of this study, it is tempting to speculate that the flexibility in branch point choice is another property influencing SAS strength and therefore needs to be considered for such predictions.

Collectively, these data identify mRNA splicing as an important disease mechanism most likely responsible for the severe adRP phenotype in patients harboring the c.620T>G mutation. Therefore, it seems reasonable to test other mutations in other genes for a potential impact on mRNA splicing. Finally, the SAS_620 could be a useful tool for biotechnological applications requiring strong splice elements.

5.2 EVALUATION OF NOVEL MRNA *TRANS*-SPLICING DUAL VECTOR APPROACH

In the second part of this study, a novel dual vector approach based on reconstitution via mRNA *trans*-splicing has been developed. After subtracting additional expression elements like promoters, pAs, all elements necessary for *trans*-splicing and flanking ITRs the total packaging capacity of the mRNA *trans*-splicing dual vectors amounts to 7.7 kb, which substantially expands the cargo capacity of single AAVs. Furthermore, this novel strategy is tackling the major limitations of the existing dual vector strategies, which are based on reconstitution at the DNA or at the protein level. Genomic reconstitution is mainly limited due to its low efficiencies (Duan et al., 2001, Trapani et al., 2014). The mRNA *trans*-splicing dual vectors tested herein were found to yield efficiencies of up to 78 % for reconstitution of a reporter gene *in vitro*. This value is by far exceeding the efficiencies obtained for genomic *in vitro* reconstitution of a reporter gene via the overlapping, genomic “*trans*-splicing” or hybrid approach, which range between 10 % and 18 % (Duan et al., 2001, Carvalho et al., 2017). Surprisingly, the reconstitution efficiencies were even higher than the ones reported for the well-performing split intein strategy, which yields about 55 % reconstitution of a reporter gene *in vitro* (Tornabene et al., 2019). However, despite providing some initial proof for the applicability of the mRNA *trans*-splicing strategy *in vivo*, reconstitution efficiencies within the retina or other target tissues remain to be investigated.

One of the major limitations of the split intein approach is its limited applicability, as split sites need to contain suitable aa and lie outside of structural domains to avoid improper folding (Tornabene et al., 2019). Therefore, this strategy requires extensive optimization steps for every targeted protein, which can be a costly endeavor. By contrast, the mRNA *trans*-splicing approach only requires the presence of an AAGGT or a CAGGT motif, which can usually be found abundantly, as they are naturally occurring at exon-exon junctions in the CDS. Another limitation of the split intein vectors is the continuous expression of the bacterial intein protein, which could potentially raise safety concerns. By contrast, all bacterial sequences contained in the mRNA *trans*-splicing vectors are removed during splicing and should consequently not lead to the expression of any bacterial protein.

The mRNA *trans*-splicing strategy is potentially also suitable for triple vector reconstitution, which would expand the packaging capacity even further. As efficient co-transduction of retinal cells by three AAVs has already been shown (Maddalena et al., 2018), this would open up new treatment options, e.g. for patients with mutations in genes requiring more than two vectors like *CDH23*. In this context, *trans*-splicing in the correct order and orientation could be ensured by using different BDs for the two *trans*-splicing events. However, in these triple vector-based approaches the overall reconstitution efficiency is expected to be reduced and needs to be thoroughly evaluated.

In 5.1, I postulated that the SAS_620 could be a suitable tool for biotechnological applications. By successfully employing the SAS_620 for mRNA *trans*-splicing-based reconstitution of genes, I provided experimental proof supporting this hypothesis. Moreover, the capability of the SAS_620 to facilitate high reconstitution efficiencies was compared to the performance of the native *RHO* exon 3 SAS. In this setting, SAS_620 vastly outperformed the native SAS, as the latter was not able to generate detectable amounts of reconstituted protein.

Despite the numerous advantages of the mRNA *trans*-splicing dual vector technology listed above, this strategy also harbors potential limitations. For instance, in contrast to the classical dual vectors used for reconstitution at the genomic level, the second expression cassette requires the presence of a promoter, which is further restricting the overall packaging capacity. In addition, classical dual

vectors used for genomic reconstitution also lack the regulatory pA sequence in the 5' vector. This implies a reduced risk of expressing truncated proteins from the split CDS contained in this vector. For the mRNA *trans*-splicing approach, both mRNAs could potentially be translated into truncated proteins. However, translation from mRNA originating from the 3' vector is less likely, as this transcript usually does not contain a suitable start codon. Moreover, the likelihood for translation of the mRNA originating from the 5' vector into truncated protein can be decreased by deleting the pA. Nevertheless, it remains to be evaluated, whether this modification is reducing the reconstitution efficiency. Furthermore, potential adverse effects of the expression of such truncated proteins in the target tissue need to be investigated. However, most diseases suitable for gene supplementation therapy follow a recessive mode of inheritance. In these cases, healthy individuals heterozygous for the respective mutations constantly express the non-functional proteins. For these diseases, a presence of truncated proteins is therefore not expected to have detrimental effects in the affected tissue.

In this study, I show that mRNA *trans*-splicing is a highly efficient and precise process suitable for reconstitution of genes *in vitro*. Moreover, when using AAV vectors for gene delivery, I demonstrated that this technology is applicable *in vivo*. Finally, I provide a proof of principle for reconstitution of the *SpCas9*-VPR module *in vitro*. However, the *SpCas9*-VPR reconstitution efficiency was reduced when compared to those obtained for the reporter gene. One possible limiting factor contributing to this finding might be the bacterial origin of this gene. In fact, the *SpCas9*-VPR mRNA might not resemble human transcripts closely enough to be recognized efficiently by the splicing machinery.

A possibility to increase the amount of *trans*-spliced transcript could be the introduction of small introns into the CDS to mimic the structure of native eukaryotic pre-mRNAs, which in turn might recruit the splicing machinery into proximity of the splicing elements used for *trans*-splicing. Other options for increasing the amount of reconstituted protein are the introduction of intronic splice enhancer elements or the addition of more potent regulatory elements like the woodchuck hepatitis virus posttranscriptional regulatory element. However, the effectiveness of these options remains to be tested. Moreover, the reconstitution

efficiencies for large genes need to be investigated *in vivo*. Very similar to the hybrid dual AAV vectors, the mRNA *trans*-splicing AAV vectors contain a binding domain and splice elements. Therefore, one can assume that these vectors might be capable of reconstitution at the genomic level in addition to the mRNA level, which could further increase the amount of reconstituted protein *in vivo*.

Collectively, in this study, I demonstrate that the dual mRNA *trans*-splicing vector strategy is an efficient and precise approach for reconstitution of genes *in vitro* and *in vivo*. By combining an efficient reconstitution with a safe and universally applicable vector design, this strategy offers many advantages over the commonly used dual vector technologies and therefore holds the potential to further expand treatment options for patients with inherited diseases.

6 REFERENCES

- ABRAMOWICZ, A. & GOS, M. 2018. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet*, 59, 253-268.
- AGUILA, M., TOLEDO, D., MORILLO, M., DOMINGUEZ, M., VAZ, B., ALVAREZ, R., DE LERA, A. R. & GARRIGA, P. 2009. Structural coupling of 11-cis-7-methyl-retinal and amino acids at the ligand binding pocket of rhodopsin. *Photochem Photobiol*, 85, 485-93.
- AHUJA, S., HORNAK, V., YAN, E. C., SYRETT, N., GONCALVES, J. A., HIRSHFELD, A., ZILIOX, M., SAKMAR, T. P., SHEVES, M., REEVES, P. J., SMITH, S. O. & EILERS, M. 2009. Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. *Nat Struct Mol Biol*, 16, 168-75.
- AL-UBAIDI, M. R., FONT, R. L., QUIAMBAO, A. B., KEENER, M. J., LIOU, G. I., OVERBEEK, P. A. & BAEHR, W. 1992. Bilateral retinal and brain tumors in transgenic mice expressing simian virus 40 large T antigen under control of the human interphotoreceptor retinoid-binding protein promoter. *J Cell Biol*, 119, 1681-7.
- ALI, R. R., REICHEL, M. B., THRASHER, A. J., LEVINSKY, R. J., KINNON, C., KANUGA, N., HUNT, D. M. & BHATTACHARYA, S. S. 1996. Gene transfer into the mouse retina mediated by an adeno-associated viral vector. *Hum Mol Genet*, 5, 591-4.
- ALLOCCA, M., DORIA, M., PETRILLO, M., COLELLA, P., GARCIA-HOYOS, M., GIBBS, D., KIM, S. R., MAGUIRE, A., REX, T. S., DI VICINO, U., CUTILLO, L., SPARROW, J. R., WILLIAMS, D. S., BENNETT, J. & AURICCHIO, A. 2008. Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest*, 118, 1955-64.
- ARIAS, M. A., KE, S. & CHASIN, L. A. 2010. Splicing by cell type. *Nat Biotechnol*, 28, 686-7.
- ARSHAVSKY, V. Y., LAMB, T. D. & PUGH, E. N., JR. 2002. G proteins and phototransduction. *Annu Rev Physiol*, 64, 153-87.
- ATHANASIOU, D., AGUILA, M., BELLINGHAM, J., LI, W., MCCULLEY, C., REEVES, P. J. & CHEETHAM, M. E. 2018. The molecular and cellular basis of rhodopsin retinitis pigmentosa reveals potential strategies for therapy. *Prog Retin Eye Res*, 62, 1-23.
- AUDO, I., FRIEDRICH, A., MOHAND-SAID, S., LANCELOT, M. E., ANTONIO, A., MOSKOVA-DOUMANOVA, V., POCH, O., BHATTACHARYA, S., SAHEL, J. A. & ZEITZ, C. 2010a. An unusual retinal phenotype associated with a novel mutation in RHO. *Arch Ophthalmol*, 128, 1036-45.
- AUDO, I., MANES, G., MOHAND-SAID, S., FRIEDRICH, A., LANCELOT, M. E., ANTONIO, A., MOSKOVA-DOUMANOVA, V., POCH, O., ZANLONGHI, X., HAMEL, C. P., SAHEL, J. A., BHATTACHARYA, S. S. & ZEITZ, C. 2010b. Spectrum of rhodopsin mutations in French autosomal dominant rod-cone dystrophy patients. *Invest Ophthalmol Vis Sci*, 51, 3687-700.
- AYUSO, C., GARCIA-SANDOVAL, B., NAJERA, C., VALVERDE, D., CARBALLO, M. & ANTINOLO, G. 1995. Retinitis pigmentosa in Spain. The Spanish Multicentric and Multidisciplinary Group for Research into Retinitis Pigmentosa. *Clin Genet*, 48, 120-2.
- BACCUS, S. A., OLVECKZY, B. P., MANU, M. & MEISTER, M. 2008. A retinal circuit that computes object motion. *J Neurosci*, 28, 6807-17.

- BECIROVIC, E., BOHM, S., NGUYEN, O. N., RIEDMAYR, L. M., HAMMELMANN, V., SCHON, C., BUTZ, E. S., WAHL-SCHOTT, C., BIEL, M. & MICHALAKIS, S. 2016a. AAV Vectors for FRET-Based Analysis of Protein-Protein Interactions in Photoreceptor Outer Segments. *Front Neurosci*, 10, 356.
- BECIROVIC, E., BOHM, S., NGUYEN, O. N., RIEDMAYR, L. M., KOCH, M. A., SCHULZE, E., KOHL, S., BORSCH, O., SANTOS-FERREIRA, T., ADER, M., MICHALAKIS, S. & BIEL, M. 2016b. In Vivo Analysis of Disease-Associated Point Mutations Unveils Profound Differences in mRNA Splicing of Peripherin-2 in Rod and Cone Photoreceptors. *PLoS Genet*, 12, e1005811.
- BERGER, A., LORAIN, S., JOSEPHINE, C., DESROSIERS, M., PECCATE, C., VOIT, T., GARCIA, L., SAHEL, J. A. & BEMELMANS, A. P. 2015. Repair of rhodopsin mRNA by spliceosome-mediated RNA trans-splicing: a new approach for autosomal dominant retinitis pigmentosa. *Mol Ther*, 23, 918-930.
- BERGER, A., MAIRE, S., GAILLARD, M. C., SAHEL, J. A., HANTRAYE, P. & BEMELMANS, A. P. 2016. mRNA trans-splicing in gene therapy for genetic diseases. *Wiley Interdiscip Rev RNA*, 7, 487-98.
- BERGET, S. M., MOORE, C. & SHARP, P. A. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A*, 74, 3171-5.
- BOHM, S., RIEDMAYR, L. M., NGUYEN, O. N. P., GIESSL, A., LIEBSCHER, T., BUTZ, E. S., SCHON, C., MICHALAKIS, S., WAHL-SCHOTT, C., BIEL, M. & BECIROVIC, E. 2017. Peripherin-2 and Rom-1 have opposing effects on rod outer segment targeting of retinitis pigmentosa-linked peripherin-2 mutants. *Sci Rep*, 7, 2321.
- BOUGHMAN, J. A., VERNON, M. & SHAVER, K. A. 1983. Usher syndrome: definition and estimate of prevalence from two high-risk populations. *J Chronic Dis*, 36, 595-603.
- BOYE, S. E., BOYE, S. L., LEWIN, A. S. & HAUSWIRTH, W. W. 2013. A comprehensive review of retinal gene therapy. *Mol Ther*, 21, 509-19.
- BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 72, 248-54.
- BRAVO-GIL, N., GONZALEZ-DEL POZO, M., MARTIN-SANCHEZ, M., MENDEZ-VIDAL, C., RODRIGUEZ-DE LA RUA, E., BORREGO, S. & ANTINOLO, G. 2017. Unravelling the genetic basis of simplex Retinitis Pigmentosa cases. *Sci Rep*, 7, 41937.
- BUNDEY, S. & CREWS, S. J. 1984. A study of retinitis pigmentosa in the City of Birmingham. I Prevalence. *J Med Genet*, 21, 417-20.
- BYUN, J., LAN, N., LONG, M. & SULLENGER, B. A. 2003. Efficient and specific repair of sickle beta-globin RNA by trans-splicing ribozymes. *RNA*, 9, 1254-63.
- CARTEGNI, L., CHEW, S. L. & KRAINER, A. R. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet*, 3, 285-98.
- CARTER-DAWSON, L. D. & LAVAIL, M. M. 1979. Rods and cones in the mouse retina. I. Structural analysis using light and electron microscopy. *J Comp Neurol*, 188, 245-62.
- CARVALHO, L. S., TURUNEN, H. T., WASSMER, S. J., LUNA-VELEZ, M. V., XIAO, R., BENNETT, J. & VANDENBERGHE, L. H. 2017. Evaluating Efficiencies of Dual AAV Approaches for Retinal Targeting. *Front Neurosci*, 11, 503.
- CHAVEZ, A., SCHEIMAN, J., VORA, S., PRUITT, B. W., TUTTLE, M., E, P. R. I., LIN, S., KIANI, S., GUZMAN, C. D., WIEGAND, D. J., TER-OVANESYAN, D., BRAFF, J. L., DAVIDSOHN, N., HOUSDEN, B. E., PERRIMON, N., WEISS, R., AACH, J., COLLINS, J.

- J. & CHURCH, G. M. 2015. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*, 12, 326-8.
- CHEN, Z. Y., YANT, S. R., HE, C. Y., MEUSE, L., SHEN, S. & KAY, M. A. 2001. Linear DNAs concatemerize in vivo and result in sustained transgene expression in mouse liver. *Mol Ther*, 3, 403-10.
- CHEW, W. L., TABEBORDBAR, M., CHENG, J. K., MALI, P., WU, E. Y., NG, A. H., ZHU, K., WAGERS, A. J. & CHURCH, G. M. 2016. A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods*, 13, 868-74.
- CHOI, V. W., SAMULSKI, R. J. & MCCARTY, D. M. 2005. Effects of adeno-associated virus DNA hairpin structure on recombination. *J Virol*, 79, 6801-7.
- CHUANG, T. J., CHEN, Y. J., CHEN, C. Y., MAI, T. L., WANG, Y. D., YE, C. S., YANG, M. Y., HSIAO, Y. T., CHANG, T. H., KUO, T. C., CHO, H. H., SHEN, C. N., KUO, H. C., LU, M. Y., CHEN, Y. H., HSIEH, S. C. & CHIANG, T. W. 2018. Integrative transcriptome sequencing reveals extensive alternative trans-splicing and cis-backsplicing in human cells. *Nucleic Acids Res*, 46, 3671-3691.
- COLELLA, P., TRAPANI, I., CESI, G., SOMMELLA, A., MANFREDI, A., PUPPO, A., IODICE, C., ROSSI, S., SIMONELLI, F., GIUNTI, M., BACCI, M. L. & AURICCHIO, A. 2014. Efficient gene delivery to the cone-enriched pig retina by dual AAV vectors. *Gene Ther*, 21, 450-6.
- CORVELO, A., HALLEGGGER, M., SMITH, C. W. & EYRAS, E. 2010. Genome-wide association between branch point properties and alternative splicing. *PLoS Comput Biol*, 6, e1001016.
- CURCIO, C. A., SLOAN, K. R., KALINA, R. E. & HENDRICKSON, A. E. 1990. Human photoreceptor topography. *J Comp Neurol*, 292, 497-523.
- CURTIS, H. J., SIBLEY, C. R. & WOOD, M. J. 2012. Mirtrons, an emerging class of atypical miRNA. *Wiley Interdiscip Rev RNA*, 3, 617-32.
- DONG, B., NAKAI, H. & XIAO, W. 2010. Characterization of genome integrity for oversized recombinant AAV vector. *Mol Ther*, 18, 87-92.
- DOOLEY, S. J., MCDUGALD, D. S., FISHER, K. J., BENNICELLI, J. L., MITCHELL, L. G. & BENNETT, J. 2018. Spliceosome-Mediated Pre-mRNA trans-Splicing Can Repair CEP290 mRNA. *Mol Ther Nucleic Acids*, 12, 294-308.
- DRYJA, T. P., MCGEE, T. L., REICHEL, E., HAHN, L. B., COWLEY, G. S., YANDELL, D. W., SANDBERG, M. A. & BERSON, E. L. 1990. A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. *Nature*, 343, 364-6.
- DUAN, D., SHARMA, P., YANG, J., YUE, Y., DUDUS, L., ZHANG, Y., FISHER, K. J. & ENGELHARDT, J. F. 1998. Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J Virol*, 72, 8568-77.
- DUAN, D., YUE, Y. & ENGELHARDT, J. F. 2001. Expanding AAV packaging capacity with trans-splicing or overlapping vectors: a quantitative comparison. *Mol Ther*, 4, 383-91.
- DUNBAR, C. E., HIGH, K. A., JOUNG, J. K., KOHN, D. B., OZAWA, K. & SADELAIN, M. 2018. Gene therapy comes of age. *Science*, 359.
- DYKA, F. M., BOYE, S. L., CHIODO, V. A., HAUSWIRTH, W. W. & BOYE, S. E. 2014. Dual adeno-associated virus vectors result in efficient in vitro and in vivo expression of an oversized gene, MYO7A. *Hum Gene Ther Methods*, 25, 166-77.
- EULER, T., HAVERKAMP, S., SCHUBERT, T. & BADEN, T. 2014. Retinal bipolar cells: elementary building blocks of vision. *Nat Rev Neurosci*, 15, 507-19.

- FARRAR, G. J., FINDLAY, J. B., KUMAR-SINGH, R., KENNA, P., HUMPHRIES, M. M., SHARPE, E. & HUMPHRIES, P. 1992. Autosomal dominant retinitis pigmentosa: a novel mutation in the rhodopsin gene in the original 3q linked family. *Hum Mol Genet*, 1, 769-71.
- FISHEL, M. L., VASKO, M. R. & KELLEY, M. R. 2007. DNA repair in neurons: so if they don't divide what's to repair? *Mutat Res*, 614, 24-36.
- FRIEDMANN, T. & ROBLIN, R. 1972. Gene therapy for human genetic disease? *Science*, 175, 949-55.
- FUJINAMI, K., ZERNANT, J., CHANA, R. K., WRIGHT, G. A., TSUNODA, K., OZAWA, Y., TSUBOTA, K., WEBSTER, A. R., MOORE, A. T., ALLIKMETS, R. & MICHAELIDES, M. 2013. ABCA4 gene screening by next-generation sequencing in a British cohort. *Invest Ophthalmol Vis Sci*, 54, 6662-74.
- GAO, K., MASUDA, A., MATSUURA, T. & OHNO, K. 2008. Human branch point consensus sequence is γ UnAy. *Nucleic Acids Res*, 36, 2257-67.
- GHOSH, A., YUE, Y., LAI, Y. & DUAN, D. 2008. A hybrid vector system expands adeno-associated viral vector packaging capacity in a transgene-independent manner. *Mol Ther*, 16, 124-30.
- GIBSON, D. G., YOUNG, L., CHUANG, R. Y., VENTER, J. C., HUTCHISON, C. A., 3RD & SMITH, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, 6, 343-5.
- GOLDSTEIN, J. A. & MCNALLY, E. M. 2010. Mechanisms of muscle weakness in muscular dystrophy. *J Gen Physiol*, 136, 29-34.
- GRONDAHL, J. 1987. Estimation of prognosis and prevalence of retinitis pigmentosa and Usher syndrome in Norway. *Clin Genet*, 31, 255-64.
- HAGEMANN, I. S., NIKIFOROVICH, G. V. & BARANSKI, T. J. 2006. Comparison of the retinitis pigmentosa mutations in rhodopsin with a functional map of the C5a receptor. *Vision Res*, 46, 4519-31.
- HAGINS, W. A., PENN, R. D. & YOSHIKAMI, S. 1970. Dark current and photocurrent in retinal rods. *Biophys J*, 10, 380-412.
- HALBERT, C. L., ALLEN, J. M. & MILLER, A. D. 2002. Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nat Biotechnol*, 20, 697-701.
- HARTONG, D. T., BERSON, E. L. & DRYJA, T. P. 2006. Retinitis pigmentosa. *Lancet*, 368, 1795-809.
- HASTINGS, K. E. 2005. SL trans-splicing: easy come or easy go? *Trends Genet*, 21, 240-7.
- HERNANDEZ-RODRIGUEZ, E. W., MONTERO-ALEJO, A. L., LOPEZ, R., SANCHEZ-GARCIA, E., MONTERO-CABRERA, L. A. & DE LA VEGA, J. M. 2013. Electron density deformations provide new insights into the spectral shift of rhodopsins. *J Comput Chem*, 34, 2460-71.
- HERNANDEZ-RODRIGUEZ, E. W., SANCHEZ-GARCIA, E., CRESPO-OTERO, R., MONTERO-ALEJO, A. L., MONTERO, L. A. & THIEL, W. 2012. Understanding rhodopsin mutations linked to the retinitis pigmentosa disease: a QM/MM and DFT/MRCI study. *J Phys Chem B*, 116, 1060-76.
- HERRMANN, R., HEFLIN, S. J., HAMMOND, T., LEE, B., WANG, J., GAINETDINOV, R. R., CARON, M. G., EGGERS, E. D., FRISHMAN, L. J., MCCALL, M. A. & ARSHAVSKY, V. Y. 2011. Rod vision is controlled by dopamine-dependent sensitization of rod bipolar cells by GABA. *Neuron*, 72, 101-10.

- HIRANO, M. & NODA, T. 2004. Genomic organization of the mouse Msh4 gene producing bicistronic, chimeric and antisense mRNA. *Gene*, 342, 165-77.
- HIRSCH, M. L., AGBANDJE-MCKENNA, M. & SAMULSKI, R. J. 2010. Little vector, big gene transduction: fragmented genome reassembly of adeno-associated virus. *Mol Ther*, 18, 6-8.
- HOON, M., OKAWA, H., DELLA SANTINA, L. & WONG, R. O. 2014. Functional architecture of the retina: development and disease. *Prog Retin Eye Res*, 42, 44-84.
- HORIUCHI, T., GINIGER, E. & AIGAKI, T. 2003. Alternative trans-splicing of constant and variable exons of a Drosophila axon guidance gene, lola. *Genes Dev*, 17, 2496-501.
- HOSSAIN, M. A. & JOHNSON, T. L. 2014. Using yeast genetics to study splicing mechanisms. *Methods Mol Biol*, 1126, 285-98.
- HUTTL, S., MICHALAKIS, S., SEELIGER, M., LUO, D. G., ACAR, N., GEIGER, H., HUDL, K., MADER, R., HAVERKAMP, S., MOSER, M., PFEIFER, A., GERSTNER, A., YAU, K. W. & BIEL, M. 2005. Impaired channel targeting and retinal degeneration in mice lacking the cyclic nucleotide-gated channel subunit CNGB1. *J Neurosci*, 25, 130-8.
- IGARASHI, T., MIYAKE, K., ASAKAWA, N., MIYAKE, N., SHIMADA, T. & TAKAHASHI, H. 2013. Direct comparison of administration routes for AAV8-mediated ocular gene therapy. *Curr Eye Res*, 38, 569-77.
- JAT, P. S., CEPKO, C. L., MULLIGAN, R. C. & SHARP, P. A. 1986. Recombinant retroviruses encoding simian virus 40 large T antigen and polyomavirus large and middle T antigens. *Mol Cell Biol*, 6, 1204-17.
- JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337, 816-21.
- KERSEY, P. J., ALLEN, J. E., ALLOT, A., BARBA, M., BODDU, S., BOLT, B. J., CARVALHO-SILVA, D., CHRISTENSEN, M., DAVIS, P., GRABMUELLER, C., KUMAR, N., LIU, Z., MAUREL, T., MOORE, B., MCDOWALL, M. D., MAHESWARI, U., NAAMATI, G., NEWMAN, V., ONG, C. K., BOLSER, D. M., DE SILVA, N., HOWE, K. L., LANGRIDGE, N., MASLEN, G., STAINES, D. M. & YATES, A. 2018. Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. *Nucleic Acids Research*, 46, D802-D808.
- KESSLER, P. D., PODSAKOFF, G. M., CHEN, X., MCQUISTON, S. A., COLOSI, P. C., MATELIS, L. A., KURTZMAN, G. J. & BYRNE, B. J. 1996. Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein. *Proc Natl Acad Sci U S A*, 93, 14082-7.
- KHAN, A. O., BECIROVIC, E., BETZ, C., NEUHAUS, C., ALTMULLER, J., MARIA RIEDMAYR, L., MOTAMENY, S., NURNBERG, G., NURNBERG, P. & BOLZ, H. J. 2017. A deep intronic CLRN1 (USH3A) founder mutation generates an aberrant exon and underlies severe Usher syndrome on the Arabian Peninsula. *Sci Rep*, 7, 1411.
- KIMATA, N., POPE, A., EILERS, M., OPEFI, C. A., ZILIOX, M., HIRSHFELD, A., ZAITSEVA, E., VOGEL, R., SHEVES, M., REEVES, P. J. & SMITH, S. O. 2016. Retinal orientation and interactions in rhodopsin reveal a two-stage trigger mechanism for activation. *Nat Commun*, 7, 12683.
- KISER, P. D., GOLCZAK, M., MAEDA, A. & PALCZEWSKI, K. 2012. Key enzymes of the retinoid (visual) cycle in vertebrate retina. *Biochim Biophys Acta*, 1821, 137-51.

- KOCH, S., SOTHILINGAM, V., GARCIA GARRIDO, M., TANIMOTO, N., BECIROVIC, E., KOCH, F., SEIDE, C., BECK, S. C., SEELIGER, M. W., BIEL, M., MUHLFRIEDEL, R. & MICHALAKIS, S. 2012. Gene therapy restores vision and delays degeneration in the CNGB1(-/-) mouse model of retinitis pigmentosa. *Hum Mol Genet*, 21, 4486-96.
- KOTTERMAN, M. A., CHALBERG, T. W. & SCHAFFER, D. V. 2015. Viral Vectors for Gene Therapy: Translational and Clinical Outlook. *Annu Rev Biomed Eng*, 17, 63-89.
- KRAUSE, M. & HIRSH, D. 1987. A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell*, 49, 753-61.
- KUMAR, M., KELLER, B., MAKALOU, N. & SUTTON, R. E. 2001. Systematic determination of the packaging limit of lentiviral vectors. *Hum Gene Ther*, 12, 1893-905.
- LAI, Y., YUE, Y., LIU, M., GHOSH, A., ENGELHARDT, J. F., CHAMBERLAIN, J. S. & DUAN, D. 2005. Efficient in vivo gene expression by trans-splicing adeno-associated viral vectors. *Nat Biotechnol*, 23, 1435-9.
- LEI, Q., LI, C., ZUO, Z., HUANG, C., CHENG, H. & ZHOU, R. 2016. Evolutionary Insights into RNA trans-Splicing in Vertebrates. *Genome Biol Evol*, 8, 562-77.
- LI, H., WANG, J., MOR, G. & SKLAR, J. 2008a. A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. *Science*, 321, 1357-61.
- LI, J., SUN, W., WANG, B., XIAO, X. & LIU, X. Q. 2008b. Protein trans-splicing as a means for viral vector-mediated in vivo gene therapy. *Hum Gene Ther*, 19, 958-64.
- LIM, L. P. & BURGE, C. B. 2001. A computational analysis of sequence features involved in recognition of short introns. *Proc Natl Acad Sci U S A*, 98, 11193-8.
- LINO, C. A., HARPER, J. C., CARNEY, J. P. & TIMLIN, J. A. 2018. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv*, 25, 1234-1257.
- LIPINSKI, D. M., BARNARD, A. R., CHARBEL ISSA, P., SINGH, M. S., DE SILVA, S. R., TRABALZA, A., ELEFThERIADOU, I., ELLISON, S. M., MAZARAKIS, N. D. & MACLAREN, R. E. 2014. Vesicular stomatitis virus glycoprotein- and Venezuelan equine encephalitis virus-derived glycoprotein-pseudotyped lentivirus vectors differentially transduce corneal endothelium, trabecular meshwork, and human photoreceptors. *Hum Gene Ther*, 25, 50-62.
- LIU, M. Y., LIU, J., MEHROTRA, D., LIU, Y., GUO, Y., BALDERA-AGUAYO, P. A., MOONEY, V. L., NOUR, A. M. & YAN, E. C. 2013. Thermal stability of rhodopsin and progression of retinitis pigmentosa: comparison of S186W and D190N rhodopsin mutants. *J Biol Chem*, 288, 17698-712.
- LOPES, V. S., BOYE, S. E., LOUIE, C. M., BOYE, S., DYKA, F., CHIODO, V., FOFO, H., HAUSWIRTH, W. W. & WILLIAMS, D. S. 2013. Retinal gene therapy with a large MYO7A cDNA using adeno-associated virus. *Gene Ther*, 20, 824-33.
- LUO, D. G., XUE, T. & YAU, K. W. 2008. How vision begins: an odyssey. *Proc Natl Acad Sci U S A*, 105, 9855-62.
- MADDALENA, A., TORNABENE, P., TIBERI, P., MINOPOLI, R., MANFREDI, A., MUTARELLI, M., ROSSI, S., SIMONELLI, F., NAGGERT, J. K., CACCHIARELLI, D. & AURICCHIO, A. 2018. Triple Vectors Expand AAV Transfer Capacity in the Retina. *Mol Ther*, 26, 524-541.
- MALI, P., YANG, L., ESVELT, K. M., AACH, J., GUELL, M., DICARLO, J. E., NORVILLE, J. E. & CHURCH, G. M. 2013. RNA-guided human genome engineering via Cas9. *Science*, 339, 823-6.

- MANN, R., MULLIGAN, R. C. & BALTIMORE, D. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*, 33, 153-9.
- MASLAND, R. H. 1996. Processing and encoding of visual information in the retina. *Curr Opin Neurobiol*, 6, 467-74.
- MASLAND, R. H. 2001. The fundamental plan of the retina. *Nat Neurosci*, 4, 877-86.
- MASLAND, R. H. 2012. The neuronal organization of the retina. *Neuron*, 76, 266-80.
- MAUBARET, C., KOSMAOGLU, M., LOW, S., CHAKAROVA, C. F., BIDOT, S., THAUVIN-ROBINET, C., ROBSON, A. G., WASEEM, N., CHEETHAM, M. E. & BHATTACHARYA, S. S. 2012. Functional characterization of a novel c.614-622del rhodopsin mutation in a French pedigree with retinitis pigmentosa. *Mol Vis*, 18, 581-7.
- MCCLEMENTS, M. E., BARNARD, A. R., SINGH, M. S., CHARBEL ISSA, P., JIANG, Z., RADU, R. A. & MACLAREN, R. E. 2019. An AAV Dual Vector Strategy Ameliorates the Stargardt Phenotype in Adult *Abca4*(-/-) Mice. *Hum Gene Ther*, 30, 590-600.
- MCCLEMENTS, M. E. & MACLAREN, R. E. 2017. Adeno-associated Virus (AAV) Dual Vector Strategies for Gene Therapy Encoding Large Transgenes. *Yale J Biol Med*, 90, 611-623.
- MCKEONE, R., WIKSTROM, M., KIEL, C. & RAKOCZY, P. E. 2014. Assessing the correlation between mutant rhodopsin stability and the severity of retinitis pigmentosa. *Mol Vis*, 20, 183-99.
- MENDELL, J. R., AL-ZAIDY, S., SHELL, R., ARNOLD, W. D., RODINO-KLAPAC, L. R., PRIOR, T. W., LOWES, L., ALFANO, L., BERRY, K., CHURCH, K., KISSEL, J. T., NAGENDRAN, S., L'ITALIEN, J., SPROULE, D. M., WELLS, C., CARDENAS, J. A., HEITZER, M. D., KASPAR, A., CORCORAN, S., BRAUN, L., LIKHTE, S., MIRANDA, C., MEYER, K., FOUST, K. D., BURGHESE, A. H. M. & KASPAR, B. K. 2017. Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N Engl J Med*, 377, 1713-1722.
- MICHALAKIS, S., BECIROVIC, E. & BIEL, M. 2018. Retinal Cyclic Nucleotide-Gated Channels: From Pathophysiology to Therapy. *Int J Mol Sci*, 19.
- MICHALAKIS, S., MUHLFRIEDEL, R., TANIMOTO, N., KRISHNAMOORTHY, V., KOCH, S., FISCHER, M. D., BECIROVIC, E., BAI, L., HUBER, G., BECK, S. C., FAHL, E., BUNING, H., PAQUET-DURAND, F., ZONG, X., GOLLISCH, T., BIEL, M. & SEELIGER, M. W. 2010. Restoration of cone vision in the *CNGA3*^{-/-} mouse model of congenital complete lack of cone photoreceptor function. *Mol Ther*, 18, 2057-63.
- MIHELEC, M., PEARSON, R. A., ROBBIE, S. J., BUCH, P. K., AZAM, S. A., BAINBRIDGE, J. W., SMITH, A. J. & ALI, R. R. 2011. Long-term preservation of cones and improvement in visual function following gene therapy in a mouse model of leber congenital amaurosis caused by guanylate cyclase-1 deficiency. *Hum Gene Ther*, 22, 1179-90.
- MOITRA, S., TIRUPULA, K. C., KLEIN-SEETHARAMAN, J. & LANGMEAD, C. J. 2012. A minimal ligand binding pocket within a network of correlated mutations identified by multiple sequence and structural analysis of G protein coupled receptors. *BMC Biophys*, 5, 13.
- MONTINI, E., CESANA, D., SCHMIDT, M., SANVITO, F., BARTHOLOMAE, C. C., RANZANI, M., BENEDICENTI, F., SERGI, L. S., AMBROSI, A., PONZONI, M., DOGLIONI, C., DI SERIO, C., VON KALLE, C. & NALDINI, L. 2009. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest*, 119, 964-75.

- MÜLLER, H. 1854. Über die entoptische Wahrnehmung der Netzhautgefäße, insbesondere als Beweismittel für die Lichtperzeption durch die nach hinten gelegenen Netzhautelemente. *Verh. Phys. Med. Ges. Würzburg*, 5.
- NATKUNARAJAH, M., TRITTIBACH, P., MCINTOSH, J., DURAN, Y., BARKER, S. E., SMITH, A. J., NATHWANI, A. C. & ALI, R. R. 2008. Assessment of ocular transduction using single-stranded and self-complementary recombinant adeno-associated virus serotype 2/8. *Gene Ther*, 15, 463-7.
- NATUREBIOTECHNOLOGYNEWS. 2018. *FDA approves hereditary blindness gene therapy* [Online]. Available: <https://www.nature.com/articles/nbt0118-6a> [Accessed 03.12.2019].
- NI, M. J., HU, Z. H., LIU, Q., LIU, M. F., LU, M. H., ZHANG, J. S., ZHANG, L. & ZHANG, Y. L. 2011. Identification and characterization of a novel non-coding RNA involved in sperm maturation. *PLoS One*, 6, e26053.
- NIBLOCK, M. & GALLO, J. M. 2012. Tau alternative splicing in familial and sporadic tauopathies. *Biochem Soc Trans*, 40, 677-80.
- NIELSEN, K. B., SORESEN, S., CARTEGNI, L., CORYDON, T. J., DOKTOR, T. K., SCHROEDER, L. D., REINERT, L. S., ELPELEG, O., KRAINER, A. R., GREGERSEN, N., KJEMS, J. & ANDRESEN, B. S. 2007. Seemingly neutral polymorphic variants may confer immunity to splicing-inactivating mutations: a synonymous SNP in exon 5 of MCAD protects from deleterious mutations in a flanking exonic splicing enhancer. *Am J Hum Genet*, 80, 416-32.
- NIKONOV, S. S., KHOLODENKO, R., LEM, J. & PUGH, E. N., JR. 2006. Physiological features of the S- and M-cone photoreceptors of wild-type mice from single-cell recordings. *J Gen Physiol*, 127, 359-74.
- NOVIKOVA, O., TOPILINA, N. & BELFORT, M. 2014. Enigmatic distribution, evolution, and function of inteins. *J Biol Chem*, 289, 14490-7.
- OCHAKOVSKI, G. A., BARTZ-SCHMIDT, K. U. & FISCHER, M. D. 2017. Retinal Gene Therapy: Surgical Vector Delivery in the Translation to Clinical Trials. *Front Neurosci*, 11, 174.
- OLDENBURG, J. & EL-MAARRI, O. 2006. New insight into the molecular basis of hemophilia A. *Int J Hematol*, 83, 96-102.
- PADRON-GARCIA, J. A., CRESPO-OTERO, R., HERNANDEZ-RODRIGUEZ, E. W., GARRIGA, P., MONTERO, L. A. & GARCIA-PINEIRO, J. C. 2004. Patterns of retinal light absorption related to retinitis pigmentosa mutants from in silico model structures of rhodopsin. *Proteins*, 57, 392-9.
- PALCZEWSKI, K., KUMASAKA, T., HORI, T., BEHNKE, C. A., MOTOSHIMA, H., FOX, B. A., LE TRONG, I., TELLER, D. C., OKADA, T., STENKAMP, R. E., YAMAMOTO, M. & MIYANO, M. 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, 289, 739-45.
- PALFI, A., CHADDERTON, N., MCKEE, A. G., BLANCO FERNANDEZ, A., HUMPHRIES, P., KENNA, P. F. & FARRAR, G. J. 2012. Efficacy of codelivery of dual AAV2/5 vectors in the murine retina and hippocampus. *Hum Gene Ther*, 23, 847-58.
- PENG, K. W., PHAM, L., YE, H., ZUFFEREY, R., TRONO, D., COSSET, F. L. & RUSSELL, S. J. 2001. Organ distribution of gene expression after intravenous infusion of targeted and untargeted lentiviral vectors. *Gene Ther*, 8, 1456-63.
- PETERSEN-JONES, S. M., OCCELLI, L. M., WINKLER, P. A., LEE, W., SPARROW, J. R., TSUKIKAWA, M., BOYE, S. L., CHIODO, V., CAPASSO, J. E., BECIROVIC, E., SCHON, C., SEELIGER, M. W., LEVIN, A. V., MICHALAKIS, S., HAUSWIRTH, W. W. & TSANG, S. C. 2011. Gene therapy for inherited blindness: a review. *Hum Gene Ther*, 22, 1111-22.

- S. H. 2018. Patients and animal models of CNGBeta1-deficient retinitis pigmentosa support gene augmentation approach. *J Clin Invest*, 128, 190-206.
- PETRS-SILVA, H., DINCULESCU, A., LI, Q., DENG, W. T., PANG, J. J., MIN, S. H., CHIODO, V., NEELEY, A. W., GOVINDASAMY, L., BENNETT, A., AGBANDJE-MCKENNA, M., ZHONG, L., LI, B., JAYANDHARAN, G. R., SRIVASTAVA, A., LEWIN, A. S. & HAUSWIRTH, W. W. 2011. Novel properties of tyrosine-mutant AAV2 vectors in the mouse retina. *Mol Ther*, 19, 293-301.
- PETRS-SILVA, H., DINCULESCU, A., LI, Q., MIN, S. H., CHIODO, V., PANG, J. J., ZHONG, L., ZOLOTUKHIN, S., SRIVASTAVA, A., LEWIN, A. S. & HAUSWIRTH, W. W. 2009. High-efficiency transduction of the mouse retina by tyrosine-mutant AAV serotype vectors. *Mol Ther*, 17, 463-71.
- PIECHNICK, R., RITTER, E., HILDEBRAND, P. W., ERNST, O. P., SCHEERER, P., HOFMANN, K. P. & HECK, M. 2012. Effect of channel mutations on the uptake and release of the retinal ligand in opsin. *Proc Natl Acad Sci U S A*, 109, 5247-52.
- PUECH, B., KOSTRUBIEC, B., HACHE, J. C. & FRANCOIS, P. 1991. [Epidemiology and prevalence of hereditary retinal dystrophies in the Northern France]. *J Fr Ophthalmol*, 14, 153-64.
- QUAZI, F., LENEVICH, S. & MOLDAI, R. S. 2012. ABCA4 is an N-retinylidene-phosphatidylethanolamine and phosphatidylethanolamine importer. *Nat Commun*, 3, 925.
- RAKOCZY, E. P., KIEL, C., MCKEONE, R., STRICHER, F. & SERRANO, L. 2011. Analysis of disease-linked rhodopsin mutations based on structure, function, and protein stability calculations. *J Mol Biol*, 405, 584-606.
- RANGARAJAN, S., WALSH, L., LESTER, W., PERRY, D., MADAN, B., LAFFAN, M., YU, H., VETTERMANN, C., PIERCE, G. F., WONG, W. Y. & PASI, K. J. 2017. AAV5-Factor VIII Gene Transfer in Severe Hemophilia A. *N Engl J Med*, 377, 2519-2530.
- REICH, S. J., AURICCHIO, A., HILDINGER, M., GLOVER, E., MAGUIRE, A. M., WILSON, J. M. & BENNETT, J. 2003. Efficient trans-splicing in the retina expands the utility of adeno-associated virus as a vector for gene therapy. *Hum Gene Ther*, 14, 37-44.
- RETNET 2019. The Retinal Information Network. Data services and software for identifying genes and mutations causing retinal degeneration. *Stephen P. Daiger, PhD, Administrator, The Univ. of Texas Health Science Center at Houston*, <https://sph.uth.edu/retnet/home.htm>, viewed on 25 June 2019.
- RIBEIL, J. A., HACEIN-BEY-ABINA, S., PAYEN, E., MAGNANI, A., SEMERARO, M., MAGRIN, E., CACCAVELLI, L., NEVEN, B., BOURGET, P., EL NEMER, W., BARTOLUCCI, P., WEBER, L., PUY, H., MERITET, J. F., GREVENT, D., BEUZARD, Y., CHRETIEN, S., LEFEBVRE, T., ROSS, R. W., NEGRE, O., VERES, G., SANDLER, L., SONI, S., DE MONTALEMBERT, M., BLANCHE, S., LEBOULCH, P. & CAVAZZANA, M. 2017. Gene Therapy in a Patient with Sickle Cell Disease. *N Engl J Med*, 376, 848-855.
- RIEDMAYR, L. M., BÖHM, S., MICHALAKIS, S. & BECIROVIC, E. 2018. Construction and Cloning of Minigenes for in vivo Analysis of Potential Splice Mutations. *Bio-protocol*, 8, e2760.
- RIVIERE, I., BROSE, K. & MULLIGAN, R. C. 1995. Effects of retroviral vector design on expression of human adenosine deaminase in murine bone marrow transplant recipients engrafted with genetically modified cells. *Proc Natl Acad Sci U S A*, 92, 6733-7.
- RYU, S. M., KOO, T., KIM, K., LIM, K., BAEK, G., KIM, S. T., KIM, H. S., KIM, D. E., LEE, H., CHUNG, E. & KIM, J. S. 2018. Adenine base editing in mouse embryos and an

- adult mouse model of Duchenne muscular dystrophy. *Nat Biotechnol*, 36, 536-539.
- SAMPATH, A. P. & RIEKE, F. 2004. Selective transmission of single photon responses by saturation at the rod-to-rod bipolar synapse. *Neuron*, 41, 431-43.
- SCHNEPP, B. C., JENSEN, R. L., CHEN, C. L., JOHNSON, P. R. & CLARK, K. R. 2005. Characterization of adeno-associated virus genomes isolated from human tissues. *J Virol*, 79, 14793-803.
- SCHULTZE, M. 1866. Zur Anatomie und Physiologie der Retina. In *Arch. Mikrosk. Anat. Entwicklungsmech*, 165–286.
- SCOTTI, M. M. & SWANSON, M. S. 2016. RNA mis-splicing in disease. *Nat Rev Genet*, 17, 19-32.
- SHAO, W., ZHAO, Q. Y., WANG, X. Y., XU, X. Y., TANG, Q., LI, M., LI, X. & XU, Y. Z. 2012. Alternative splicing and trans-splicing events revealed by analysis of the Bombyx mori transcriptome. *RNA*, 18, 1395-407.
- SHI, Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol*, 18, 655-670.
- SMITH, A. J., BAINBRIDGE, J. W. & ALI, R. R. 2012. Gene supplementation therapy for recessive forms of inherited retinal dystrophies. *Gene Ther*, 19, 154-61.
- SOUKARIEH, O., GAILDRAT, P., HAMIEH, M., DROUET, A., BAERT-DESURMONT, S., FREBOURG, T., TOSI, M. & MARTINS, A. 2016. Exonic Splicing Mutations Are More Prevalent than Currently Estimated and Can Be Predicted by Using In Silico Tools. *PLoS Genet*, 12, e1005756.
- STANFORD, W. L., COHN, J. B. & CORDES, S. P. 2001. Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet*, 2, 756-68.
- STEPANKIW, N., RAGHAVAN, M., FOGARTY, E. A., GRIMSON, A. & PLEISS, J. A. 2015. Widespread alternative and aberrant splicing revealed by lariat sequencing. *Nucleic Acids Res*, 43, 8488-501.
- STEVENS, A. J., BROWN, Z. Z., SHAH, N. H., SEKAR, G., COWBURN, D. & MUIR, T. W. 2016. Design of a Split Intein with Exceptional Protein Splicing Activity. *J Am Chem Soc*, 138, 2162-5.
- STEVENS, A. J., SEKAR, G., SHAH, N. H., MOSTAFAVI, A. Z., COWBURN, D. & MUIR, T. W. 2017. A promiscuous split intein with expanded protein engineering applications. *Proc Natl Acad Sci U S A*, 114, 8538-8543.
- TAN, E., DING, X. Q., SAADI, A., AGARWAL, N., NAASH, M. I. & AL-UBAIDI, M. R. 2004. Expression of cone-photoreceptor-specific antigens in a cell line derived from retinal tumors in transgenic mice. *Invest Ophthalmol Vis Sci*, 45, 764-8.
- TANNA, P., STRAUSS, R. W., FUJINAMI, K. & MICHAELIDES, M. 2017. Stargardt disease: clinical features, molecular genetics, animal models and therapeutic options. *Br J Ophthalmol*, 101, 25-30.
- TORNABENE, P., TRAPANI, I., MINOPOLI, R., CENTRULO, M., LUPO, M., DE SIMONE, S., TIBERI, P., DELL'AQUILA, F., MARROCCO, E., IODICE, C., IULIANO, A., GESUALDO, C., ROSSI, S., GIAQUINTO, L., ALBERT, S., HOYNG, C. B., POLISHCHUK, E., CREMERS, F. P. M., SURACE, E. M., SIMONELLI, F., DE MATTEIS, M. A., POLISHCHUK, R. & AURICCHIO, A. 2019. Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina. *Sci Transl Med*, 11.
- TRAPANI, I. & AURICCHIO, A. 2018. Seeing the Light after 25 Years of Retinal Gene Therapy. *Trends Mol Med*, 24, 669-681.

- TRAPANI, I., COLELLA, P., SOMMELLA, A., IODICE, C., CESI, G., DE SIMONE, S., MARROCCO, E., ROSSI, S., GIUNTI, M., PALFI, A., FARRAR, G. J., POLISHCHUK, R. & AURICCHIO, A. 2014. Effective delivery of large genes to the retina by dual AAV vectors. *EMBO Mol Med*, 6, 194-211.
- TRAPANI, I., TORIELLO, E., DE SIMONE, S., COLELLA, P., IODICE, C., POLISHCHUK, E. V., SOMMELLA, A., COLECCHI, L., ROSSI, S., SIMONELLI, F., GIUNTI, M., BACCI, M. L., POLISHCHUK, R. S. & AURICCHIO, A. 2015. Improved dual AAV vectors with reduced expression of truncated proteins are safe and effective in the retina of a mouse model of Stargardt disease. *Hum Mol Genet*, 24, 6811-25.
- VAN DER PLOEG, L. H., LIU, A. Y., MICHELS, P. A., DE LANGE, T., BORST, P., MAJUMDER, H. K., WEBER, H., VEENEMAN, G. H. & VAN BOOM, J. 1982. RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes. *Nucleic Acids Res*, 10, 3591-604.
- WANG, Y., MA, M., XIAO, X. & WANG, Z. 2012. Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules. *Nat Struct Mol Biol*, 19, 1044-52.
- WANG, Z., XIAO, X., VAN NOSTRAND, E. & BURGE, C. B. 2006. General and specific functions of exonic splicing silencers in splicing control. *Mol Cell*, 23, 61-70.
- WASSLE, H. & BOYCOTT, B. B. 1991. Functional architecture of the mammalian retina. *Physiol Rev*, 71, 447-80.
- WEBVISION 2011. Simple Anatomy of the Retina by Helga Kolb. Retrieved from <http://webvision.med.utah.edu/>, 16 July 2019.
- WILLETT, K. & BENNETT, J. 2013. Immunology of AAV-Mediated Gene Transfer in the Eye. *Front Immunol*, 4, 261.
- WU, C. S., YU, C. Y., CHUANG, C. Y., HSIAO, M., KAO, C. F., KUO, H. C. & CHUANG, T. J. 2014. Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency. *Genome Res*, 24, 25-36.
- WU, H., HU, Z. & LIU, X. Q. 1998. Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci U S A*, 95, 9226-31.
- WU, J., ZHAO, W., ZHONG, L., HAN, Z., LI, B., MA, W., WEIGEL-KELLEY, K. A., WARRINGTON, K. H. & SRIVASTAVA, A. 2007. Self-complementary recombinant adeno-associated viral vectors: packaging capacity and the role of rep proteins in vector purity. *Hum Gene Ther*, 18, 171-82.
- WU, Z., YANG, H. & COLOSI, P. 2010. Effect of genome size on AAV vector packaging. *Mol Ther*, 18, 80-6.
- XIAO, X., LI, J. & SAMULSKI, R. J. 1998. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J Virol*, 72, 2224-32.
- XU, J. 2005. Preparation, culture, and immortalization of mouse embryonic fibroblasts. *Curr Protoc Mol Biol*, Chapter 28, Unit 28 1.
- YANG, R. B., FOSTER, D. C., GARBERS, D. L. & FULLE, H. J. 1995. Two membrane forms of guanylyl cyclase found in the eye. *Proc Natl Acad Sci U S A*, 92, 602-6.
- YEO, N. C., CHAVEZ, A., LANCE-BYRNE, A., CHAN, Y., MENN, D., MILANOVA, D., KUO, C. C., GUO, X., SHARMA, S., TUNG, A., CECCHI, R. J., TUTTLE, M., PRADHAN, S., LIM, E. T., DAVIDSOHN, N., EBRAHIMKHANI, M. R., COLLINS, J. J., LEWIS, N. E., KIANI, S. & CHURCH, G. M. 2018. An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat Methods*, 15, 611-616.

References

- YOUNG, R. W. & BOK, D. 1969. Participation of the retinal pigment epithelium in the rod outer segment renewal process. *J Cell Biol*, 42, 392-403.
- ZHANG, M. Q. 1998. Statistical features of human exons and their flanking regions. *Hum Mol Genet*, 7, 919-32.
- ZHONG, L., LI, B., MAH, C. S., GOVINDASAMY, L., AGBANDJE-MCKENNA, M., COOPER, M., HERZOG, R. W., ZOLOTUKHIN, I., WARRINGTON, K. H., JR., WEIGEL-VAN AKEN, K. A., HOBBS, J. A., ZOLOTUKHIN, S., MUZYCZKA, N. & SRIVASTAVA, A. 2008. Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci U S A*, 105, 7827-32.
- ZHOU, X. E., MELCHER, K. & XU, H. E. 2012. Structure and activation of rhodopsin. *Acta Pharmacol Sin*, 33, 291-9.

7 APPENDIX

7.1 LIST OF RT-PCR PRIMERS

Primer	Sequence
<i>RHO</i> exon 1 forward	5' GCTTCTTTGCCACCCTGGGCG 3'
Citrine reverse	5' TCCTCGCCCTTGCTCACCAT 3'
Myc-tag reverse	5' CCTCTTCTGAGATGAGTTTTTTG 3'
pcDNA3.1 forward	5' ACTATAGGGAGACCCAAGCTG 3'
pcDNA3.1 reverse	5' AGGCTGATCAGCGGGTTTAA 3'
<i>RPS27</i> forward	5'CGACCTACGCACACGAGAAC 3'
<i>RPS27</i> reverse	5' CGACCTACGCACACGAGAAC 3'
<i>HBQ1</i> forward	5' TGGAAGAAGCTGGGCAGCAA 3'
<i>S100A12</i> forward	5' AAACCTTCCTTGGCTCAGTG 3'
<i>CLRN1</i> reverse	5' TTGAGCCTGGTGCCTGGTA 3'
<i>CNGB1</i> forward	5' CCAACTCAGAGAGCTGCTCCG 3'
Lariat 1 forward	5' GGCATAGCCCTAGCCTCAGC 3'
Lariat 1 reverse	5' TTCTGCCCTACACCCTACC 3'
Lariat 2.1 forward	5' GCGACACTGATTCCACAAGGT 3'
Lariat 2.1 reverse	5' CTGGAGGACCCTACAAAGAGC 3'
Lariat 2.2 forward	5' ATGCTGGGGCTGGGCATG 3'
Lariat 2.2 reverse	5' TTTGGGCCAGGAGACATAACAAGG 3'
Cerulean forward	5' ATGGTGAGCAAGGGCGAGG 3'
Cerulean reverse	5' CTTGTACAGCTCGTCCATGCC 3'
<i>SpCas9</i> forward	5' AGAACGCTTGAAAACCTTACGCT 3'
<i>SpCas9</i> reverse	5' TTGATGTCCAGTTCCTGATCC 3'

7.2 LIST OF BINDING DOMAIN SEQUENCES

Binding Domain	Sequence (Length)
a	5'GTTAACCCTCCAGTCAGGACTCAAACCCAGTAGTGTCTGGTTCCAGGC ACTGACCTGCTAGCTGTATGTCTCCTGGCCCAAATGCCACTCAGGGTAG GGGTGTAGGGCAGAAGAAGAAACAGACTCTAATGTTGCTACAAGGGCTG GTCCCATCTCCTGAGCCCCATGTCAAACAGAATCCAAGACATCCCAACCC TTCACCTTGGCTGTGGAATTCCCCCTAATCCTCAACTAAGCTAGGCGCAA ATTCCAATCCTCTTTGGTCTAGTACCCCGGGGGCAGCCCCCTCTAACCTT GGGCTCAGCAGCAGGGGAGGGGTACCCACACCTTCCTAGTGCAGGTG GCCATATTGTGGCCCCTTGGAACTGGGTCCCCTCAGCCTCTAGGCGATT GTCTCCTAATGGGGCTGAGATGAGACACGATATCAGTGGGGACAGTGGT TTGGACAATAGGACTGGTGACTCTGGTCCCAGAGGCCTCATGTCCCTCT GTCTCCAGAAAATTCCCCTCCTCACTTCCCTTTCCGGTACCTCCTCAGTC TTGCTAGGGTCCATTTCTTACCCCTTGTGAATTTGAGCCCACCCCTGG ACTTTTTCCCCTCTTCTCCAATCTGGCCTAGTTCTATCCTTCTAGACTG GAAGCAGAGCCGCTGGACGCTCTGGGTTTCCCTGAGGCATTTAAATCCGT CCACTGTCACCAATATCAGGAACCATTGCCACGTCCTAATGACGTGCGCT GGAAGCGGGCCCCTCTAGTTTCCAGAAGCTGCACAAAGATCCCTTAGAT ACTCTGTGTGTCCATCTTTGGCCTGGAAAATACTCTCACCCCTGGGGCTAG GAAGACCTCGGTTGCCCGGGCTGTACAAACTTCCCTCAAATGCAGAGCCTG AGGGCTCTCCCCACCTCCTCACCAACCCTCTGCGTGGCATAGCCCTAGCC TCAGCGGGCAGTGGATGCTGGTCTAGAGGCTGGGCATGCAGGGAGAGGC TGGGTGGTGTCTGTTAACGCAGCCACCAACAATGAAGCGACACTG ATTCCACAAGGTGCATCTGCATCCCCTCGGGCCCTGATCCATTCCATCC TGTCACCCAGCCATGCAGACGTTTATGATCCCCTTTTCCAGGGAGGGAAT GTGAAGCCCCAGAAAGGGCCAGCGCCAGGTGGAATTCGCTAGCTCGGCA GCCACCTTGGCTGTTCCCAAGTCCCTCACAGGCAGGG 3' (1227 bp)
b	5'GTTAACCCTCCAGTCAGGACTCAAACCCAGTAGTGTCTGGTTCCAGGC ACTGACCTGCTAGCTGTATGTCTCCTGGCCCAAATGCCACTCAGGGTAG GGGTGTAGGGCAGAAGAAGAAACAGACTCTAATGTTGCTACAAGGGCTG GTCCCATCTCCTGAGCCCCATGTCAAACAGAATCCAAGACATCCCAACCC TTCACCTTGGCTGTGGAATTCGCTAGCTCGGCAGCCACCTTGGCTGTTCC CAAGTCCCTCACAGGCAGGG 3' (267 bp)
c	5'TCTGCCCTACACCCCTACCCTGAGTGGGCATTTGGGCCAGGAGACATA CAAGGTCAGTGCCTGGAACCAGACACTACTGGGTTTGAGTCCTGACTGG AGG 3' (100bp)
d	5'CCTCCAGTCAGGACTCAAACCCAGTAGTGTCTGGTTCCAGGCCTGAC CTG 3' (51 bp)

Appendix

e	5'GTTCCAGGGAGGGAATGTGAAGCCCCAGAAAGGGCCAGCGCCAGGTGG AATTCGCTAGCTCGGCAGCCACCTTGGCTGTTCCCAAGTCCCTCACAGGC AGGG 3' (102 bp)
f	5'GTTAACCCTCCAGTCAGGACTCAAACCCAGTAGTGTCTGGTTCCAGGC ACTGACCTGCTAGCTCGGCAGCCA 3' (72 bp)
g	5'GTTAACCCTCCAGTCAGGACTCAAACCCAGTAGTGTCTGGTTCCAGGC ACTGACCTGCTAGCTGTATGTCTCCTGGCCCAAATGCCACTCAGGGTAG GGGTGTAGGGCAGAAGAAGAAACAGACTCTAATGTTGCTACAAGGGCTG GTCCCATCTCCTGAGCCCCATGTCAAACAGAATCCAAGACATCCCAACCC TTCACCTTGGCTGTGGAATTTCCCCTAATCCTCAACTAAGCTAGGCGCAA ATTCCAATCCTCTTTGGTCTAGTACCCCGGGGGCAGCCCCCTTAACCTT GGGCCTCAGCAGCAGGGGAGGGGTACCCACACCTTCCTAGTCCAGGGAG GGAATGTGAAGCCCCAGAAAGGGCCAGCGCCAGGTGGAATTCGCTAGCT CGGCAGCCACCTTGGCTGTTCCCAAGTCCCTCACAGGCAGGG 3' (438 bp)
h	5'GTTAACCCTCCAGTCAGGACTCAAACCCAGTAGTGTCTGGTTCCAGGC ACTGACCTGCTAGCTCGGCAGCCACCTTGGCTGTTCCCAAGTCCCTCACA GGCAGGG 3' (105 bp)
B2	5'CACCATTTCATGGTGATAGCCGGGCTGCTGTTTGTGCAGGGCTGGCACT GAACACTGCCTTGATCTTATTTGGAGCAATATGCGCTTGTCTAATTTCA CAGCAAGAAAAGTGAAGCTGAGGCTCAAAGAAGTCAAGCGCCCTGCTGGG GCG 3' (149 bp)
i	5'CATCTGACCACCAGCGAATTTTTGCATCCAGCTGTTTAATCAGCGTTG GTTTAACCGCCAGTCAGGCTTTCTTTCAAAGATGTGATCGCCTTGCAGCA CT 3' (100 bp)
j	5'CATCTGACCACCTGCGAATTTTTGCATCCTGCTGTTTAATCTGCGTTG GTTTAACCGCCTGTCTGGCTTTTTTTCACTGATGTGATCGCCTTGTGCA CT 3' (100 bp)
k	5'CCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCCAGCTGG GTAACGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTCAAAGATGTGG ATT 3' (100 bp)
l	5'CCATCCCGCATCTGACCACCTGCGAATGGATTTTTGCATCCTGCTGG GTAACGTTGGCAATTTAACCGCCTGTCTGGCTTTCTTTCAATGATGTGG ATT 3' (100 bp)

7.3 RECONSTITUTION EFFICIENCIES

HEK293			
Co-transfected Constructs	Reconstitution efficiency (mean \pm SEM)	n	Depicted in
BD_a + SAS_620	3.6 % \pm 3.5 %	3	Figure 21
BD_b + SAS_620	29.2 % \pm 6.0 %	3	Figure 21
BD_c (B3) + SAS_620	7.6 % \pm 2.3 %	4	Figure 21
BD_d + SAS_620	45.8 % \pm 5.5 %	3	Figure 21
BD_e (B1) + SAS_620	62.3 % \pm 7.9 %	8	Figure 21
BD_f + SAS_620	45.4 % \pm 9.5 %	4	Figure 21
BD_g + SAS_620	20.6 % \pm 5.1 %	3	Figure 21
BD_h + SAS_620	52.6 % \pm 10.1 %	6	Figure 21
B_e (B1) + S2	4.4 % \pm 2.5 %	3	Figure 22
B_e (B1) + S3	0 % \pm 0 %	3	Figure 22
B2 + SAS_620 (S1)	40.7 % \pm 19.9 %	4	Figure 22
B2 + S2	35.6 % \pm 1.5 %	3	Figure 22
B2 + S3	0 % \pm 0 %	3	Figure 22
B_c (B3) + S2	0 % \pm 0 %	3	Figure 22
B_c (B3) + S3	0 % \pm 0 %	3	Figure 22
BD_i + SAS_620	36.4 % \pm 8.8 %	3	Figure 25
BD_j + SAS_620	78.3 % \pm 2.1 %	3	Figure 25
BD_k + SAS_620	43.4 % \pm 2.2 %	3	Figure 25
BD_l + SAS_620	21.2 % \pm 6.5 %	3	Figure 25
661W			
BD_e + SAS_620 (S1)	44.7 % \pm 12.0 %	4	Figure 23
BD_h + SAS_620 (S1)	31.5 % \pm 6.6 %	4	Figure 23
BD_g + SAS_620 (S1)	18.8 % \pm 8.7 %	4	Figure 23
BD_e + S3	0.2 % \pm 0.2 %	4	Figure 23
MEF			
BD_e + SAS_620 (S1)	52.6 % \pm 20.0 %	3	Figure 23
BD_h + SAS_620 (S1)	41.7 % \pm 5.4 %	3	Figure 23
BD_g + SAS_620 (S1)	17.8 % \pm 4.3 %	3	Figure 23

Appendix

BD_e + S3	0 % ± 0 %	3	Figure 23
-----------	-----------	---	-----------

LIST OF PUBLICATIONS

Riedmayr LM, Böhm S, Biel M and Becirovic E (2019). "Enigmatic rhodopsin mutation creates an exceptionally strong splice acceptor site." *Human Molecular Genetics*. Accepted for publication.

Panagiotopoulos A, Pavlou M, Böhm S, Biel M, **Riedmayr LM**[#] and Becirovic E[#] (2019). "Antisense oligonucleotide- and CRISPR-Cas9-mediated rescue of mRNA splicing for a deep intronic *CLRN1* mutation." Under Review. #: corresponding author

Böhm S, Splith V, **Riedmayr LM**, Rötzer RD, Gasparoni G, Nordström KJV, Wagner J, Hinrichsmeyer KS, Walter J, Wahl-Schott C, Fenske S, Biel M, Michalakis S and Becirovic E (2019). "A gene therapy for inherited blindness using dCas9-VPR-mediated transcriptional activation". Manuscript in preparation.

Riedmayr, LM (2020). "SMaRT for Therapeutic Purposes." *Methods Mol Biol*, 2079, 219-232.

Riedmayr LM¹, Böhm S¹, Michalakis S and Becirovic E (2018). "Construction and Cloning of Minigenes for *in vivo* Analysis of Potential Splice Mutations." *Bio-protocol* 8(5): e2760. doi: 10.21769/BioProtoc.2760. ¹: equal contribution

Böhm S¹, **Riedmayr LM**¹, Nguyen ONP, Giessler A, Liebscher T, Butz ES, Schön C, Michalakis S, Wahl Schott C, Biel M and Becirovic E (2017). "Peripherin-2 and Rom-1 have opposing effects on rod outer segment targeting of retinitis pigmentosa-linked peripherin-2 mutants." *Sci Rep* 7(1): 2321. doi: 10.1038/s41598-017-02514-5. ¹: equal contribution

Khan AO, Becirovic E, Betz C, Neuhaus C, Altmüller J, **Riedmayr LM**, Motameny S, Nürnberg G, Nürnberg P and Bolz HJ (2017). "A deep intronic *CLRN1* (*USH3A*) founder mutation generates an aberrant exon and underlies severe Usher syndrome on the Arabian Peninsula." *Sci Rep* 7(1): 1411. doi: 10.1038/s41598-017-01577-8.

Becirovic E, Böhm S, Nguyen ONP, **Riedmayr LM**, Hammelmann V, Schön C, Butz ES, Wahl-Schott C, Biel M and Michalakis, S. (2016). "AAV Vectors for FRET-Based Analysis of Protein-Protein Interactions in Photoreceptor Outer Segments." *Front Neurosci*, 10, 356. doi: 10.3389/fnins.2016.00356.

Becirovic E, Böhm S, Nguyen ONP, **Riedmayr LM**, Koch MA, Schulze E, Kohl S, Borsch O, Santos-Ferreira T, Ader M, Michalakis S and Biel M (2016). "In Vivo Analysis of Disease-Associated Point Mutations Unveils Profound Differences in mRNA Splicing of Peripherin-2 in Rod and Cone Photoreceptors." *PLoS Genet* 12(1): e1005811.

ACKNOWLEDGEMENTS

First of all, I want to thank my supervisor Elvir Becirovic for his relentless support and incredible scientific ideas. Elvir, thank you very much for being the greatest mentor and for always supporting me in any possible way. Moreover, I am very grateful to Prof. Biel for giving me the opportunity to be a part of his amazing lab and for supporting me as a member of my Thesis advisory committee (TAC) with his keen scientific insights. Another big thank you goes to my other two TAC members, Prof. Kiebler and Dr. Godinho, who accompanied me on the demanding way through my PhD.

Next, I want to thank the GSN for making all of this possible. Thank you for admitting me to the Fast-track program, which taught me the world and surrounded me with incredible people and the brightest minds. Thank you for supporting my career as well as my personal development by financing various workshops and conferences.

Moreover, I want to thank Berit Noack for her continuous and unbreakable cloning efforts. Another thank you goes to all of my students, especially to Sanas, for sharing the heavy workload with me.

I particularly want to thank all of my colleagues for making the lab a place that I loved going to every single day. I am especially thankful for finding such amazing friends in many of them. Sybille, I don't know what I would have done without you! Thank you so much for always being there for me and being an amazing shoulder to cry on for any scientific (or personal) break-downs along the way. Vicky, I admire your strength and your infinite scientific knowledge! Thank you for always sharing your wisdom with me and for making the many late night work days bearable. Thanks to Anna, Johanna and Constanze for many relaxing evenings with great conversations and great food. Another thank you goes to Ella and Klara for completing our room and making it an entertaining place to be and to Verena, Marc, Konsti and René for the fun joint lunch breaks!

Zuletzt möchte ich mich bei meiner Familie und bei meinem Mann bedanken. Lieber Riedmayr-Clan, vielen, vielen Dank für die endlose und unermüdliche Unterstützung, die ich von euch schon mein ganzes Leben lang erhalte! Danke, dass

Acknowledgements

ihr mir beigebracht habt, dass man alles erreichen kann, wenn man nur daran glaubt! Ihr seid die beste Familie, die man sich nur wünschen kann!

Felix, danke, dass du alles bist, was ich im Leben brauche – mein Fels in der Brandung und meine rettende Stütze! Danke, dass du Tag und Nacht für mich da bist, dass du an meiner Seite die steinigsten Wege beschreitest, dass du meine wildesten Träume unterstützt und, dass du immer und unermüdlich an mich glaubst. Ohne dich würde es diese Arbeit nicht geben. Deswegen widme ich Sie hiermit dir. Ich liebe dich!

AFFIDAVIT

Eidesstattliche Versicherung/Affidavit

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation „IDENTIFICATION OF A UNIQUE SPLICE ACCEPTOR SITE ENABLES NOVEL BIOTECHNOLOGICAL APPLICATION“ selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation “IDENTIFICATION OF A UNIQUE SPLICE ACCEPTOR SITE ENABLES NOVEL BIOTECHNOLOGICAL APPLICATION” is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München/Munich, 31.07.2019

Lisa Riedmayr

DECLARATION OF AUTHOR CONTRIBUTIONS

Subretinal injections were performed by Dr. Sybille Böhm (Biel group, LMU Munich). Sanas Mir-Bashiri (Biel group, LMU Munich) contributed to achieving the necessary sample numbers for the cerulean western blotting experiments.

Munich, 31.07.2019

Supervisor (Dr. Elvir Becirovic)

Lisa Riedmayr