



**HAL**  
open science

# Optimization of a gene therapy approach for glycogen storage disease type III using recombinant adeno-associated viral vectors

Antoine Gardin

► **To cite this version:**

Antoine Gardin. Optimization of a gene therapy approach for glycogen storage disease type III using recombinant adeno-associated viral vectors. Biotechnology. Université Paris-Saclay, 2023. English. NNT : 2023UPASL049 . tel-04643375

**HAL Id: tel-04643375**

**<https://theses.hal.science/tel-04643375>**

Submitted on 10 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimization of a gene therapy approach for glycogen storage disease type III using recombinant adeno-associated viral vectors

*Optimisation d'une approche de thérapie génique pour la glycogénose de type  
III basée sur l'utilisation de vecteurs viraux de type adéno-associés*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577, Structure et Dynamique des Systèmes Vivants (SDSV)  
Spécialité de doctorat : Biotechnologies  
Graduate School : Life Science and Health. Référent : Université d'Evry Val d'Essonne

Thèse préparée dans l'unité de recherche Intégrare UMR\_S951 (Université Paris Saclay, Univ Evry, Inserm, Généthon), sous la direction de **Giuseppe RONZITTI** (CR, HDR) et la co-direction de **Fabienne Rajas** (DR).

Thèse soutenue à Paris-Saclay, le 4 juillet 2023,  
par

**Antoine GARDIN**

## Composition du Jury

Membres du jury avec voix délibérative

<b>Emmanuel GONZALES</b> PUPH, Université Paris-Saclay	Président
<b>Caroline LE GUINER</b> CR, Université de Nantes	Rapporteur & Examinatrice
<b>Pasquale PICCOLO</b> Assimilé CR, TIGEM, Naples	Rapporteur & Examineur
<b>Gloria GONZALEZ-ASEGUINOLAZA</b> Assimilée DR, University of Navarra	Examinatrice

**Titre :** Optimisation d'une approche de thérapie génique pour la glycogénose de type III basée sur des vecteurs viraux de type adéno-associés

**Mots clés :** Métabolisme, Thérapie génique, Vecteur AAV recombinant, Glycogénose

**Résumé :**

La glycogénose de type III (GSDIII) est une maladie rare du métabolisme liée à des mutations du gène *AGL* codant l'enzyme débranchante du glycogène (GDE). Elle s'accompagne d'une accumulation de glycogène dans le foie, le cœur et les muscles. La grande taille de l'ADN complémentaire de GDE rend difficile l'utilisation d'une stratégie de thérapie génique basée sur un vecteur adéno-associé recombinant (rAAV) unique.

Dans la première partie de ce travail, nous avons développé une approche utilisant un vecteur rAAV unique en générant une forme tronquée fonctionnelle de GDE, afin de permettre la correction de l'atteinte musculaire et cardiaque dans plusieurs modèles de GSDIII. En utilisant la structure tridimensionnelle de GDE et la position des variants faux-sens décrits dans cette pathologie, nous avons généré plusieurs GDE tronquées conservant une activité enzymatique. Une des GDE ayant une délétion dans le domaine N-terminal,  $\Delta$ Nter2-GDE, conserve une activité équivalente à celle de la GDE de taille complète *in vivo*. Un vecteur rAAV exprimant  $\Delta$ Nter2-GDE, injecté par voie intraveineuse à des souris *Ag1<sup>-/-</sup>*, a permis la correction complète de l'accumulation de glycogène dans le cœur et les muscles squelettiques trois mois après injection et une normalisation de l'histologie et de la force musculaire des souris. Une correction équivalente a aussi été obtenue dans le cœur et les muscles de rats *Ag1<sup>-/-</sup>* injectés avec ce même vecteur. Enfin, l'expression de  $\Delta$ Nter2-GDE dans des cellules musculaires humaines invalidées pour le gène *AGL* permet aussi de corriger l'accumulation de glycogène, sans toxicité spécifique.

Ces résultats démontrent la possibilité de corriger l'atteinte musculaire dans plusieurs modèles de GSDIII en utilisant une GDE tronquée et sont prometteurs pour une éventuelle utilisation clinique de ce vecteur pour traiter des patients atteints de GSDIII.

Dans la seconde partie de ce travail, nous avons optimisé un vecteur rAAV unique afin de cibler le foie et de corriger l'atteinte hépatique des souris *Ag1<sup>-/-</sup>*. Le vecteur développé dans la première partie ne permet pas une correction du foie du fait de l'utilisation d'un promoteur et d'une capsid virale sans ciblage ou expression dans le foie. L'injection d'un vecteur rAAV optimisé pour le ciblage du foie dans des souris *Ag1<sup>-/-</sup>* de 4-5 mois, ayant déjà une fibrose hépatique avancée, permet une correction transitoire des paramètres biochimiques (glycémie, ASAT, ALAT), avec une perte de correction progressive 6 mois après injection. Une injection plus précoce, dans des souris *Ag1<sup>-/-</sup>* de 6 semaines sans fibrose lors de l'injection, permet une correction du phénotype hépatique chez la moitié des animaux, persistant 9 mois après injection. L'autre moitié des animaux présentent néanmoins une perte progressive de correction à partir de 6 mois post-injection, comme chez les souris injectées à un âge plus avancé. Les raisons de ce phénomène sont en cours de caractérisation.

Une poursuite de l'optimisation reste donc nécessaire afin d'améliorer la correction du phénotype hépatique et surtout afin de pouvoir corriger la totalité des atteintes de la GSDIII, hépatique, musculaire et cardiaque, en utilisant un seul vecteur rAAV.

**Title :** Optimization of a gene therapy approach for glycogen storage disease type III using recombinant adeno-associated viral vectors

**Keywords :** Metabolism, Gene Transfer, recombinant AAV vector, Glycogen Storage Disease

**Abstract :**

Glycogen storage disease type III (GSDIII) is a rare inborn error of metabolism affecting liver, skeletal muscle, and heart, due to mutations of the *AGL* gene encoding for the glycogen debranching enzyme (GDE). The 4.6 kb GDE cDNA represents the major technical challenge toward the development of a single recombinant adeno-associated virus (rAAV)-derived vector strategy and alternative approaches using bacterial GDE orthologs or dual vectors have been proposed.

In the first part of this thesis, we developed a single rAAV vector strategy expressing a truncated GDE protein in order to correct the heart and muscle impairment in multiple GSDIII models. Using molecular modelling and the existing information on *AGL* missense variants, we generated multiple truncated GDEs retaining activity. Among them, the N-terminal-truncated mutant  $\Delta$ Nter2-GDE had a similar efficacy *in vivo* compared to the full-size enzyme. A rAAV vector expressing  $\Delta$ Nter2-GDE allowed complete glycogen clearance in the heart and muscle of *Agf<sup>-/-</sup>* mice, three months after intravenous injection, as well as normalization of histology features and restoration of muscle strength. Similarly, glycogen accumulation and histological features were corrected in a recently generated *Agf<sup>-/-</sup>* rat model. Finally,  $\Delta$ Nter2-GDE corrected glycogen accumulation in an *in vitro* human skeletal muscle cellular model of GSDIII, without specific toxicity. In conclusion, our results demonstrate the ability of a functional truncated GDE transgene to correct the muscle and heart phenotype in multiple models of GSDIII, supporting its clinical translation to GSDIII patients.

In the second part of this thesis, we optimized a single vector rAAV approach to correct the liver phenotype, which could not be achieved with our  $\Delta$ Nter2-GDE expressing rAAV vector because of the use of a promoter and a capsid both with liver-detargeting features. We designed an rAAV with a capsid with high liver tropism expressing the full-length GDE under the control of a liver promoter. Intravenous administration of high dose of this vector to 4-5-month old *Agf<sup>-/-</sup>* could only achieve transient correction of the glycemia and liver enzymes 2 months after injection, with progressive loss of correction during the follow-up. Because, at this age, *Agf<sup>-/-</sup>* mice already display significant liver fibrosis, we then administered the rAAV vector intravenously to younger 6-week-old *Agf<sup>-/-</sup>* mice. Glycemia normalized quickly after injection. Long-term correction of the liver phenotype, up to 9 months post-injection, was only achieved in half of the injected mice, with normalization of both glycogen content and histologic features of the disease. In the other half of the cohort, a progressive loss of correction was observed from 6 months post-injection, similar to *Agf<sup>-/-</sup>* mice injected at an older age. The mechanism behind the progressive loss of correction observed is unclear and currently under investigations.

Further optimization is required to improve liver targeting and GDE expression, especially with the use of the  $\Delta$ Nter2-GDE mutant, as well as to achieve the correction of both the liver and muscle diseases.

## **ACKNOWLEDGEMENTS**

I am grateful to Pr Emmanuel Gonzales for having accepted to chair the jury during my PhD defense. Thank you also for your help these last years and for your support to obtain the Inserm PhD fellowship.

I am grateful to Dr Caroline Le Guiner for having accepted to be rapporteur for my PhD defense, considering your expertise in gene therapy for muscle disorders.

I am grateful to Dr Pasquale Piccolo for having accepted to be rapporteur for my PhD defense, considering your expertise in gene therapy for liver diseases.

I am grateful to Dr Gloria Gonzalez-Aseguinolaza for having accepted to participate to my PhD defense, considering your expertise in gene therapy for liver diseases and inborn errors of metabolism.

I am grateful to Dr Giuseppe Ronzitti for having accepted me in the lab, as well as for the supervision of this PhD. Thank you for your kindness and for everything that you taught me. You encouraged me to present in congresses, to write a review and even to revise a gene therapy paper even if I would spontaneously not have felt qualified to do so ! I have learned a lot on research and science in general, during these 3 years, as well as in many other non-scientific aspects. I am glad of this PhD and I hope that we will be able to pursue collaborations in the future.

I am grateful to Dr Fabienne Rajas for having accepted to supervise my PhD along with Giuseppe. Thank you for your kindness, for what you taught me especially on liver metabolism and glycogen storage diseases. I am very happy of this PhD and I hope that we will be able to pursue collaborations in the future.

I would also like to thank Généthon, as well as the national research agency (Inserm), for their support and for the doctoral fellowship that has allowed me to pursue this PhD project.

Je remercie toute l'équipe de Généthon que j'ai pu côtoyer durant ces trois années passées à Généthon, et principalement tous les membres de l'équipe IMF qui m'ont si bien accueilli !

Merci plus spécifiquement à toi, Jérémy, pour ton accueil à mon arrivée dans l'équipe et pour ton aide et ton soutien tout au long de la thèse. J'étais assez stressé de débarquer en Novembre 2020 dans un monde complètement nouveau pour moi, fait de clonages, de souris, de *Wire Hang* et de dosages de glycogène. J'ai réussi à m'intégrer facilement et je reste intimement persuadé que c'est en grande partie grâce à toi et à Sévérine ! Merci pour ta patience quand tu as dû tout me réexpliquer depuis le début, jusqu'à l'utilisation des pipettes ! Merci pour ta gentillesse, ton écoute et pour nos discussions qui ont certainement permis au projet d'en arriver là où il est aujourd'hui.

Merci à mes trois super co-doctorantes, Lindsay, Mallaury et Justine, pour votre gentillesse et votre soutien tout au long de la thèse, ainsi que pour ces nombreux débriefs coincés dans le RER D ou dans les embouteillages ! Je garde un très bon souvenir de notre voyage à Athènes l'année dernière.

Merci beaucoup Sévérine pour ton soutien chaleureux (et bruyant) !

Merci bien sûr aussi aux autres membres de la « team GSDIII » : Valle surtout, mais aussi Youssef, Ali et Louis !

Merci enfin aux autres membres de l'équipe IMF, pour votre aide au quotidien durant la thèse et pour toutes les discussions enrichissantes qu'on a pu avoir : Bérangère, Christian, Coralie, Pierre-Romain, Fanny, Tiziana, Novella, Hanadi, Pauline V., Pauline F., David, Louisa, Ante, Amandine, Audrey, Sara, Danella, Duohao, Emilie, Gianni, Louis B., Marie-Thenia, Luca, Paul, Sophie et Sylvain.

Merci aussi aux personnes des autres équipes de Généthon, incluant notamment Juliette L., Jérôme D. et les membres de l'équipe de l'Histo (Emilie, Fanny et Christophe) pour leur aide. Merci beaucoup à nos collaborateurs d'I-STEM, Lucille Rossiaud, Lucile Hoch et Xavier Nissan !

Merci enfin à ma famille et à mes amis pour leur soutien sans faille durant ces trois années, notamment ma sœur Lulu, mon frère Paul et mes amis My-Van, Romain et Laura. Et, bien sûr, merci à toi, Marica, pour ton aide et ton soutien précieux, au quotidien. Je pense que tu es devenue, un peu contre ton gré, une spécialiste des approches de thérapie génique ciblant la GSDIII !

## **TABLE OF CONTENTS**

<b>ABBREVIATIONS</b>	8.
<b>LIST OF TABLES AND FIGURES</b>	10.
<b>STRUCTURE OF THE MANUSCRIPT</b>	13.
<b>INTRODUCTION</b>	14.
I. <u>GENE THERAPY: GENERAL CONCEPTS</u>	15.
1. <i>EX VIVO</i> VS. <i>IN VIVO</i> GENE THERAPY	16.
2. THREE MAIN STRATEGIES OF <i>IN VIVO</i> GENE THERAPY	16.
3. TYPES OF VECTORS	18.
II. <u>GENE THERAPY AND rAAV VECTORS</u>	20.
1. WILD TYPE AAVs	20.
2. RECOMBINANT AAV VECTORS (rAAV)	23.
3. CAPSID DISCOVERY AND ENGINEERING	24.
4. MAIN TECHNICAL LIMITATIONS	27.
5. IMMUNE RESPONSE-ASSOCIATED TOXICITIES	30.
6. TUMORIGENESIS	32.
III. <u>LIVER-TARGETED GENE THERAPY WITH A FOCUS ON rAAV</u>	33.
1. ADVANTAGES OF LIVER TARGETING	33.
2. LIVER GENE THERAPY STRATEGIES	35.
3. SPECIFIC ISSUES ASSOCIATED WITH LIVER TARGETING	39.
IV. <u>MUSCLE-TARGETED GENE THERAPY BASED ON rAAV</u>	45.
1. UNDERLYING MUSCLE DISEASES	45.
2. ISSUES ASSOCIATED WITH MUSCLE TARGETING USING rAAV	48.
V. <u>GLYCOGEN METABOLISM AND GLYCOGEN STORAGE DISEASE TYPE III</u>	49.
1. GLYCOGEN METABOLISM	49.
2. CLINICAL ASPECTS OF GLYCOGEN DISEASE TYPE III (GSDIII)	56.
3. PHYSIOPATHOLOGY OF GSDIII	58.
4. GENE THERAPY APPROACHES IN ANIMAL MODELS OF GSDIII	62.
<b>OBJECTIVES</b>	64.
<b>PART I: CORRECTION OF THE MUSCLE DISEASE IN GSDIII</b>	67.
I. <u>INTRODUCTION</u>	68.

II. <u>MATERIAL AND METHODS</u>	70.
III. <u>RESULTS</u>	75.
1. GENERATION OF A SINGLE VECTOR APPROACH FOR GSDIII BASED ON A FUNCTIONAL, TRUNCATED GDE	75.
2. rAAV ENCODING ΔNTER2-GDE RESCUES THE CARDIAC AND MUSCLE PHENOTYPE OF <i>AGL</i> <sup>-/-</sup> MICE	80.
3. rAAV-ΔNTER2-GDE CORRECTS GLYCOGEN ACCUMULATION IN AN <i>AGL</i> <sup>-/-</sup> RAT MODEL OF GSDIII	82.
4. ΔNTER2-GDE CLEARS GLYCOGEN IN A HUMAN MUSCLE CELL MODEL OF GSDIII WITHOUT ANY SPECIFIC TOXICITY	84.
IV. <u>DISCUSSION</u>	86.
V. <u>SUPPLEMENTARY MATERIALS</u>	89.
<b>PART II: CORRECTION OF THE LIVER DISEASE IN GSDIII</b>	110.
I. <u>INTRODUCTION</u>	111.
II. <u>MATERIAL AND METHODS</u>	111.
III. <u>RESULTS</u>	114.
1. CHARACTERIZATION OF THE LIVER INVOLVEMENT IN <i>AGL</i> <sup>-/-</sup> MICE	114.
2. TRANSIENT CORRECTION OF THE LIVER INVOLVEMENT IN 4-5-MONTH-OLD <i>AGL</i> <sup>-/-</sup> MICE	116.
3. LONG-TERM CORRECTION OF THE LIVER INVOLVEMENT IN 6-WEEK-OLD <i>AGL</i> <sup>-/-</sup> MICE	119.
IV. <u>DISCUSSION</u>	122.
V. <u>SUPPLEMENTARY MATERIALS</u>	127.
<b>CONCLUSION</b>	133.
<b>REFERENCES</b>	135.
<b>SCIENTIFIC COMMUNICATIONS</b>	148.
<b>RÉSUMÉ EN FRANÇAIS</b>	150.

## **ABBREVIATIONS**

List of abbreviations present at least twice in the manuscript:

AAV: adeno-associated virus

ALT: alanine aminotransferase

ApoE-hAAT: human apolipoprotein E enhancer with human  $\alpha$ 1-antitrypsin promoter

ASO: antisense oligonucleotide

AST: aspartate aminotransferase

CMV: cytomegalovirus

CpG: cytosine phospho-guanine

CPK: creatine phosphokinase

CRISPR: clustered regularly interspaced short palindromic repeats

cDNA: complementary deoxyribonucleic acid

DMD: Duchenne muscular dystrophy

DNA: deoxyribonucleic acid

G1P: glucose-1-phosphate

G6P: glucose-6-phosphate

G6Pase: glucose-6-phosphatase

GAA: acid  $\alpha$ -glucosidase

GC:  $\alpha$ 1,6-glucosidase activity

GDE: glycogen debranching enzyme

GSD: glycogen storage disease

GSDIII: glycogen storage disease type III

GT: glycosyltransferase activity

Gyg: glycogenin

Gys: glycogen synthase

HCC: hepatocellular carcinoma

HES: hematoxylin-eosin stain

hiPSC: human induced pluripotent stem cells

HPS: hematoxylin-phloxine saffran

ITR: inverted terminal repeats

LGMD: limb-girdle muscular dystrophy

LNP: lipid nanoparticles

MMA: methylmalonic acidemia

mRNA: messenger ribonucleic acid

MTM: myotubular myopathy

NAbs: neutralizing antibodies  
PAS: periodic acid Schiff  
PBS: phosphate buffer saline  
PCR: polymerase chain reaction  
PFIC: progressive familial intrahepatic cholestasis  
PH1: primary hyperoxaluria type I  
Poly-A: poly-adenylation  
rAAV : recombinant adeno-associated virus  
RNA: ribonucleic acid  
siRNA: small interfering ribonucleic acid  
SEM: standard error of the mean  
SMA: spinal muscular atrophy  
tRNA: transfer ribonucleic acid  
TIR: terminal inverted repeats  
TSS: transcription starting site  
TTR: transthyretin  
vg: viral genome  
VGCN: vector genome copy number

## **LIST OF FIGURES**

**Figure 1:** Overview of the different general modalities of treatment for a disease

**Figure 2:** Three main gene therapy strategies

**Figure 3:** Different types of vectors for gene therapy

**Figure 4:** Schematic representation of the structure and genome of the AAV

**Figure 5:** Entry of the AAV into the cell

**Figure 6:** Schematic representation of the rAAV genome

**Figure 7:** Methods for capsid discovery and engineering

**Figure 8:** Different dual vectors approaches allowing gene transfer of cDNA above the encapsidation limit of rAAVs

**Figure 9:** Schematic representation of immune system activation following rAAV injection

**Figure 10:** Schematic representation of the hepatic sinusoid

**Figure 11:** Different rAAV-based *in vivo* gene therapy strategies targeting the liver

**Figure 12:** Mechanism of action of the *TTR* gene-targeting CRISPR-Cas9 strategy used in the clinical trial for TTR-mediated amyloidosis

**Figure 13:** Mechanism of action of the nuclease-free integrative strategy used for MMA

**Figure 14:** Gene therapy strategy using the *Piggybac* transposase system

**Figure 15:** Structure of glycosidic bonds and overall structure of glycogen

**Figure 16:** Structure of a glycogen particle

**Figure 17:** Glycogen synthesis

**Figure 18:** Glycogen breakdown

**Figure 19:** Simplified overview of glycogen metabolism

**Figure 20:** Overview of glycogen storage diseases (GSD)

**Figure 21:** Locus of the human *AGL* gene and schematic representation of the different mRNA isoforms

**Figure 22:** Representation of the 3D crystal structure of *Candida glabrata* GDE and of the prediction of the human GDE 3D structure using the Alpha-Fold2 software

**Figure 23:** Generation of truncated GDE polypeptides in the C domain

**Figure 24:** Generation of truncated GDE polypeptides in the N-terminal domain

**Figure 25:** Rescue of muscle and heart impairment in the *Agf*<sup>-/-</sup> mouse model with rAAV-ΔNter2-GDE vector

**Figure 26:** Correction of muscle and heart impairment in the *Agf*<sup>-/-</sup> rat model with rAAV-ΔNter2-GDE vector

**Figure 27:** Clearance of glycogen in human induced pluripotent stem cells-derived muscle cells treated with rAAV-ΔNter2-GDE.

**Figure 28:** Characterization of the liver biochemical phenotype of *Agf*<sup>-/-</sup> mice

**Figure 29:** Characterization of liver histology in *Agf<sup>-/-</sup>* mice

**Figure 30:** Transient correction of the liver phenotype of *Agf<sup>-/-</sup>* mice injected at 5-month of age with a rAAV containing an hepatic promoter

**Figure 31:** Long-term correction of the hepatic phenotype in some *Agf<sup>-/-</sup>* mice injected at 6 weeks of age with an rAAV vector optimized for liver targeting

**Figure 32:** Expression of GDE in the liver of treated mice and normalization of glycogen content in some *Agf<sup>-/-</sup>* mice injected at 6 weeks of age with an rAAV optimized for liver targeting

**Figure 33:** Correction of liver histology in some young *Agf<sup>-/-</sup>* mice injected at 6 weeks of age with an rAAV optimized for liver targeting

### **LIST OF TABLES**

**Table I:** Tropism of natural and selected chimeric AAV capsids

**Table II:** List of gene therapy clinical trials targeting liver

**Table III:** Comparison of liver growth and embryonic development between human and C57BL/6 mice

**Table IV:** List of gene therapy clinical trials targeting inherited muscle and heart diseases

### **LIST OF SUPPLEMENTARY MATERIALS**

**Figure S1:** Selection of a short promoter able to efficiently express a reporter gene in muscle

**Figure S2:** Evaluation of the requirement for an intron in rAAV vectors expressing GDE

**Figure S3:** Expression of the full-length GDE in heart and muscle of *Agf<sup>-/-</sup>* mice using an optimized rAAV vector

**Figure S4:** Partial correction of muscle glycogen accumulation in *Agf<sup>-/-</sup>* mice treated with a rAAV vector encoding full-length GDE

**Figure S5:** Structural organization of human GDE protein

**Figure S6:** Generation of functional truncated GDE polypeptides in the C domain

**Figure S7:** Generation of functional truncated GDE polypeptides in the N-terminal domain

**Figure S8:** Molecular dynamics of human GDE and  $\Delta$ Nter2-GDE showing conformational changes occurring along the simulation

**Figure S9:** Improved productivity and quality of vectors bearing truncated  $\Delta$ Nter2-GDE

**Figure S10:** Rescue of the muscle and heart phenotype of the *Agf<sup>-/-</sup>* mouse model using rAAV vector encoding for  $\Delta$ Nter2-GDE

**Figure S11:** Correction of the muscle and heart phenotype of the *Agf<sup>-/-</sup>* rat model using rAAV vector encoding for  $\Delta$ Nter2-GDE

**Figure S12:** Elevation of Myom3 level in the plasma of 6-week-old *Agf<sup>-/-</sup>* rats

**Figure S13:** Absence of specific toxicity in human induced pluripotent stem cells (hiPSC)-derived muscle cells treated with rAAV- $\Delta$ Nter2-GDE.

**Figure S14:** PAS staining in the liver of *Agf*<sup>+/+</sup> and *Agf*<sup>-/-</sup> mice

**Figure S15:** Absence of correction of the liver phenotype in *Agf*<sup>-/-</sup> injected with rAAV containing the miniCMV promoter

**Figure S16:** Absence of GDE expression in the liver of *Agf*<sup>-/-</sup> mice injected with rAAV containing the mLSP promoter.

**Figure S17:** Absence of correction of the muscle phenotype in *Agf*<sup>-/-</sup> mice with rAAV containing the mLSP promoter

**Figure S18:** Selection of a liver promoter using the mSEAP reporter gene

**Figure S19:** Correction of the liver phenotype according to the correction of the glycemia.

**Table S1:** Glycogen levels expressed as mmol/g proteins from Figure S4

**Table S2:** Glycogen levels expressed as mmol/g proteins from Figure 23

**Table S3:** Glycogen levels expressed as mmol/g proteins from Figure 25

**Table S4:** Glycogen levels expressed as mmol/g proteins from Figure 26

## **STRUCTURE OF THE MANUSCRIPT**

This manuscript describes our approach aiming to **optimize a gene therapy approach using recombinant adeno-associated viral vectors (rAAV) for glycogen storage disease type III (GSDIII)**, which is an inherited metabolic disease affecting muscles, heart and liver.

The manuscript begins with an **introduction** which gives an overview of gene therapy, **focusing on the use of rAAV**. We briefly describe the structure of the wild type AAV and the use of rAAV as potent gene therapy viral vectors. Then, we describe the **specificities of rAAV use for the treatment of diseases affecting either liver or muscles**, highlighting both the advantages and limitations of rAAV when targeting these two very different organs. Finally, the **last part of the introduction presents GSDIII** as well as the metabolism of glycogen in both liver and muscles. We also highlight the successes and limitations of prior gene therapy approaches designed to correct GSDIII and **why we think further optimizations is required** before such treatment could be applied to GSDIII patients.

The manuscript is then structured into **two main, largely independent, parts**: the first one focuses on the **treatment of the muscular impairment**, which represents the major disease burden in adult patients affected by GSDIII. We describe the optimizations that led to the generation of a potent, muscle-specific, rAAV vector able to correct the muscle and heart phenotype while de-targeting the liver. This part has been submitted for publication. The second part describes our **attempts to target the liver** in order to correct the metabolic and liver disease. In this part, we discuss the difficulties that we have encountered and the potential solutions to overcome them.

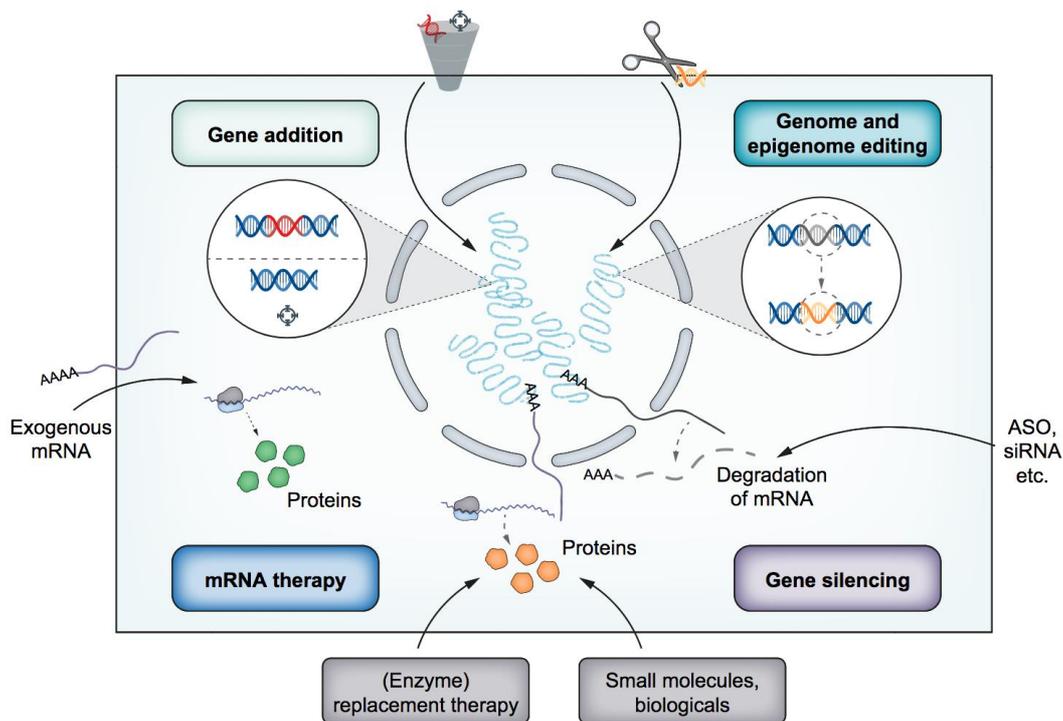
Finally, the manuscript ends with a **short conclusion** in which we describe the perspectives for potential applications of such gene therapy approaches to GSDIII patients. At the end, a list of the scientific communications performed during the 3 years of PhD is provided, as well as a short summary in French, as requested by the doctoral school.

# INTRODUCTION

## I. GENE THERAPY: GENERAL CONCEPTS

Treatment of a disease, especially of genetic origin, can rely on one of the following modalities, shown in **Figure 1** :

- **Intervention on DNA** with the use of exogenous DNA (gene addition) or the modification of the patient's genome or of its expression (gene silencing, gene edition).
- **Intervention on RNA** with the use of messenger RNAs (mRNA) to express a protein, the use of small interfering RNAs (siRNA) or antisense oligonucleotides (ASO) to suppress the expression of a protein, or the use of RNA editors or small molecules able to modulate splicing in order to modify the RNA stability and/or the produced protein.
- **Intervention on proteins** with the replacement of a missing protein (enzyme replacement therapy) or modulation of cellular proteins targeting a dysregulated pathway (small molecules).



**Figure 1. Overview of the different general modalities of treatment for a disease, from Gardin et al<sup>1</sup>.**

Gene therapy is defined by the American Society of Gene and Cell Therapy (ASGCT) as “the introduction, removal, or change in the content of a person’s genetic code with the goal of treating or curing a disease”<sup>2</sup>. Therefore, “gene therapy” consists of the first level and include gene addition, gene silencing and gene editing strategies<sup>2</sup>.

## 1. EX VIVO VS. IN VIVO GENE THERAPY

There are two main types of gene therapy: *ex vivo* gene therapy and *in vivo* gene therapy<sup>2</sup>.

**Ex vivo gene therapy** requires sampling of the patient's cells, their modification outside of the body and their reinjection to the patient<sup>2</sup>. Modification of cells can be achieved using either non-viral vectors, such as electroporation of ribonucleoprotein complexes, or viral vectors, such as lentiviruses<sup>3,4</sup>. This approach quasi-exclusively targets **hematopoietic stem cells** for now, usually for inborn diseases of the bone marrow, and requires an autologous hematopoietic stem cell transplantation<sup>3,4</sup>. More than 60 years of clinical use of hematopoietic stem cell transplantation have allowed the development of efficient hematopoietic stem cell collection and reinjection procedures. Since the early 2000s, several clinical trials have been performed using *ex vivo* gene therapy approaches for **primary immune deficiencies, inborn errors of metabolism and hemoglobinopathies**<sup>3-5</sup>. Some gene therapy products now have market approval, such as for metachromatic leukodystrophy (Libmeldy®),  $\beta$ -thalassemia major (Zynteglo®) or adenosine deaminase deficiency (Strimvelis®).

**In vivo** gene therapy relies on the injection of a vector, either viral or non-viral, able to carry the genetic material or the genome editing machinery to the right cells in the right tissue. This vector can be administered either **intravenously** or **locally**, for example intrathecally, intracerebrally or intraocularly<sup>2</sup>. Once in the target cell, usually a terminally-differentiated cell or its precursor, the cargo can perform the targeted modification (gene expression, silencing or editing). Both non-viral and viral vectors can be used, the most frequently one being the recombinant **adeno-associated viral** vector<sup>2,6</sup>. These approaches are evaluated in diseases involving the liver, the muscles, the central nervous system, the eye and the heart, among others<sup>6</sup>. Most of the diseases that could be treated by gene therapy are of genetic origin, but some can be acquired diseases such as neurodegenerative diseases (Alzheimer's disease, NCT05040217) or heart failure (NCT04179643) for instance. Several gene therapy products now have market approval, for example for **hemophilia A** (Roctavian®), **hemophilia B** (Hemgenix®), **spinal muscular atrophy** (Zolgensma®) or RPE65-mutated **retinitis pigmentosa** (Luxturna®). Many more clinical trials are currently ongoing<sup>1,2</sup>.

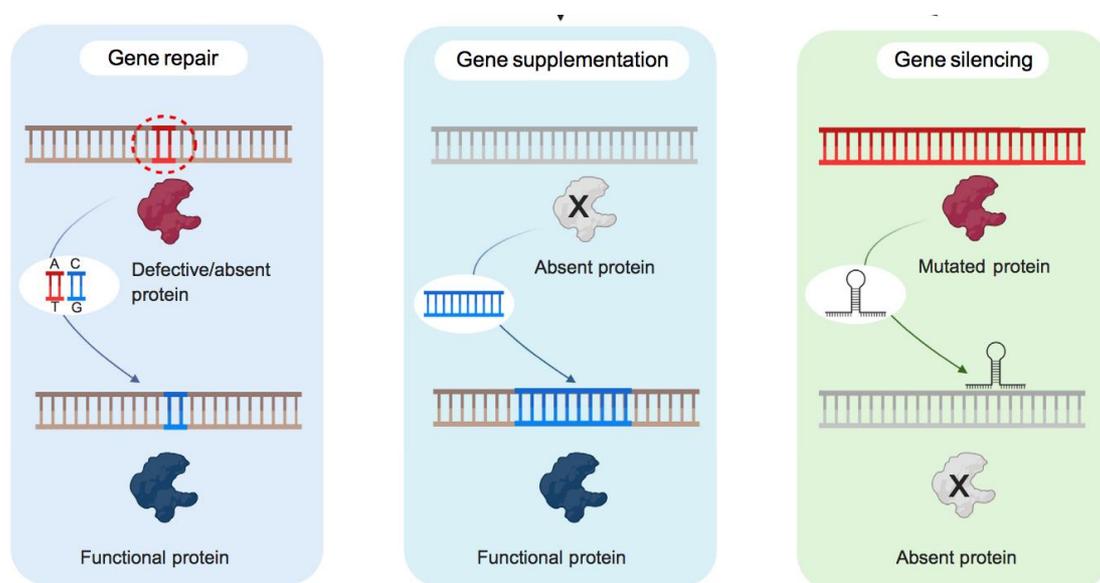
## 2. THREE MAIN STRATEGIES OF IN VIVO GENE THERAPY

The choice of the gene therapy strategy depends on the disease which is targeted. Mendelian diseases can be broadly classified into diseases with either **loss-of-function** or **gain-of-function** variants. For the first ones, the complete absence (autosomal or X-linked recessive inheritance) or an insufficient amount (haploinsufficiency, autosomal dominant inheritance) of a functional protein leads

to the disease, and the gene therapy approach mainly aims at **restoring the level of this protein**. For the diseases with gain of function variants (autosomal or X-linked dominant inheritance), the production of a toxic protein leads to the disease, either through abnormal protein aggregation or dysregulation of cellular signaling. In this case, the goal of the gene therapy is to **reduce the level of the toxic protein**, and if possible, to restore the expression of the normal protein. Indeed, the complete suppression of the toxic protein would also prevent the protein to play its physiological role and therefore could lead to a specific disease (equivalent of a loss of function variant). Such diseases, as well as diseases due to haploinsufficiency, have a small therapeutic index, with the risk of toxicities if the expression level of the protein is not maintained within its physiological (narrow) range.

Three different strategies of *in vivo* gene therapy can be used and are depicted in **Figure 2** : **gene addition, gene silencing** and **gene repair**<sup>2</sup>.

**Gene addition** aims to restore the expression of a missing protein, usually by the delivery of an exogenous copy of the complementary DNA (cDNA) of this protein<sup>2</sup>. This strategy is mostly intended for mendelian diseases with loss of function variants. But it could also be used in order to treat non-genetic diseases, by boosting the expression of a specific protein in order to improve or compensate the phenotype (a growth factor for instance). Potential toxicities could be observed if supraphysiologic levels of the transgene are achieved.



**Figure 2. Three main gene therapy strategies, from Maestro et al.<sup>2</sup>.**

**Gene silencing** aims to reduce the expression of a **protein**, either by delivering an inhibitor of this protein or by blocking its expression, using siRNA or CRISPR-Cas9 for instance<sup>1,2</sup>. This strategy is

mostly intended for mendelian diseases with gain of function variants in which a mutation in a specific gene leads to the **accumulation of a toxic protein**. It can also be used in some metabolic recessive diseases with accumulation of a toxic metabolite due to an **enzymatic blockade**, in order to silence the enzyme responsible for the production of this metabolite (primary hyperoxaluria for instance)<sup>2</sup>. Potential toxicities could be observed in case of complete knock-down of a vital protein.

**Gene repair** aims to modify directly the patient's genome in order to correct the DNA mutation and to **restore the expression** of a missing protein for diseases with loss of function variants or to **suppress the expression of a toxic protein** for diseases with gain of function variants<sup>1,2</sup>. One of the advantages of this strategy is the restoration of the physiological regulation of the protein, as well as to allow the expression of the normal protein in diseases with gain-of-function variants. The approach relies mostly on the **CRISPR-Cas9 technology** (and its by-products such as base editing and prime editing), allowing either gene repair using **homologous recombination** in presence of a DNA template, or gene silencing using indels generation through **non-homologous end-joining**<sup>1,2</sup>.

### 3. TYPES OF VECTORS

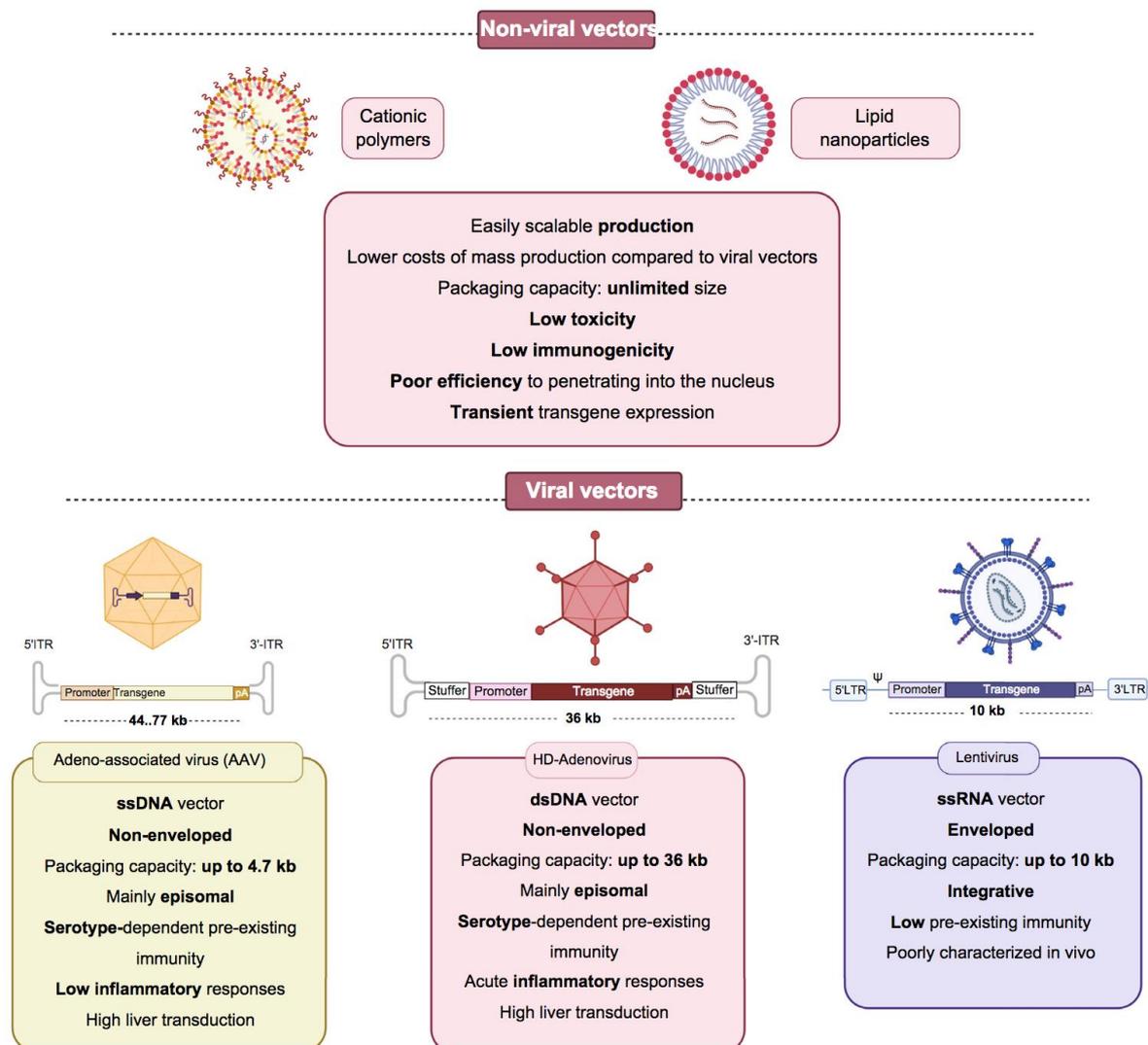
Many types of viral and non-viral vectors can be used depending on the selected disease, the route of administration and the targeted organ<sup>2</sup>. A brief description is provided in [Figure 3](#)<sup>2</sup>.

**Non-viral vectors** consist mostly in **lipid nanoparticles** (LNP) and **cationic polymers**<sup>2</sup>. These lipophilic molecules are well tolerated, have low immunogenicity and have been optimized for strong **liver targeting** through binding of N-acetylgalactosamine residues located on the vector to asialoglycoprotein receptor (ASGPR) in hepatocytes<sup>2</sup>. In contrary, most of the other organs are yet poorly targeted by this approach. Another important limitation is the **weak entry into the nucleus** of cargos delivered through this therapeutic modality, with often transient expression of the gene of interest<sup>2</sup>. Such vectors are therefore not used for gene addition strategy, but rather to deliver the gene edition machinery to the liver, ensuring a transient expression of the Cas9 enzyme in order to reduce potential toxicities<sup>1,7</sup>. It can also be used for **messenger RNA (mRNA) or siRNA delivery**, and several RNA-based liver-targeting products now have market approval with impressive results (Patisiran®, Givosiran®, Lumasiran®)<sup>1,2</sup>.

**Viral vectors** include recombinant **adeno-associated viral vectors** (rAAV), **lentiviruses**, Helper-dependent **adenoviruses** and others **less studied viral vectors** (Herpes simplex, vaccinia).

**Recombinant AAV vectors** (rAAVs) are the most frequently used viral vectors for *in vivo* gene therapy<sup>2,6</sup>. rAAV are small, single stranded non-enveloped DNA virus, in which the viral genes have

been replaced by an expression cassette<sup>8</sup>. These vectors have **low immunogenicity**, are well tolerated after intravenous injection and have an overall good safety profile<sup>1,2</sup>. rAAVs remain mostly in the nucleus as **episome**, i.e. in a circularized extra-chromosomal form, and exhibit only a low frequency of integration in the host genome<sup>2,8</sup>. As such, cell proliferation leads to **progressive loss of the episomes** (“dilution” of the vector) and rAAVs are mostly used to target tissues with **terminally-differentiated quiescent cells** (central nervous system, heart) or **slowly regenerating tissues** under normal conditions (muscles, liver in adults)<sup>2,6</sup>. These vectors are detailed further in the next section (II.).



**Figure 3. Different types of vectors for gene therapy, from Maestro et al.<sup>2</sup>.**

**Lentiviruses** are the most frequently used viral vectors for **ex vivo gene therapy**<sup>2,4</sup>. They consist in large single stranded enveloped RNA virus, able to integrate in the host genome<sup>8</sup>, therefore allowing persistent expression in proliferating tissues such as the bone marrow. A risk of **insertional**

**mutagenesis** was initially present with the earliest  $\gamma$ -retroviruses and seems to be now very low with the newest “self-inactivating” lentiviruses, able to integrate only one time in the genome (and not to perform serial cycles of integration as  $\gamma$ -retroviruses)<sup>4,8</sup>. One limitation of this kind of vector is their **poor tropism for organs such as the liver**, limiting their use for *in vivo* gene therapy<sup>2</sup>. Nonetheless, several recent studies have shown the ability of an optimized lentivirus to efficiently target the liver and to express therapeutic levels of factor IX in both mice and non-human primates<sup>9,10</sup> or to correct a mouse model of glycogen storage disease type Ia (GSD Ia)<sup>11</sup>.

**Helper-dependent adenoviruses** are mostly used for **oncolytic gene therapy** approaches for cancers. They consist in large double stranded defective DNA viruses with a large packaging ability<sup>2,8</sup>. They remain as episomes in the nucleus. The two major limitations are their **strong immunogenicity** and the presence of a high **frequency of immunization** against adenoviruses during childhood<sup>2,8</sup>. The use of adenovirus in early 2000s in gene therapy trials for the ornithine transcarbamylase (OTC) urea cycle defect resulted in a poor clinical tolerance of this vector with fever, inflammation and death of one of the participant<sup>12</sup>.

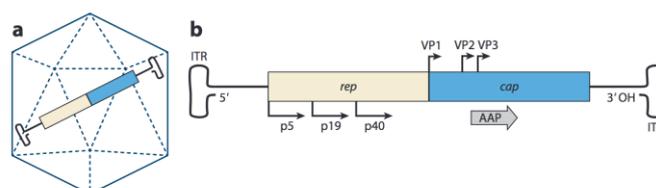
**Other vectors**, derived from Herpes simplex or the vaccinia virus, have also been used in oncolytic gene therapy<sup>8</sup>, and are beyond the scope of this *Introduction*.

## II. GENE THERAPY AND rAAV VECTORS

rAAVs are the most commonly used vectors for *in vivo* gene therapy targeting liver and muscle and are the one chosen in our study to develop a gene therapy approach for glycogen storage disease type III.

### 1. WILD TYPE AAV

Wild type AAVs are **natural viruses** identified in 1965 as cell culture contaminants in adenoviral preparations and have been considered for long as non-pathogenic<sup>13</sup>. AAVs are single stranded non-enveloped DNA viruses with a genome of 4.7 kb flanked by two *inverted terminal repeats* (ITR) sequences<sup>8</sup>. ITRs are mandatory elements for **viral genome packaging** into the capsid and play a role in genome replication.



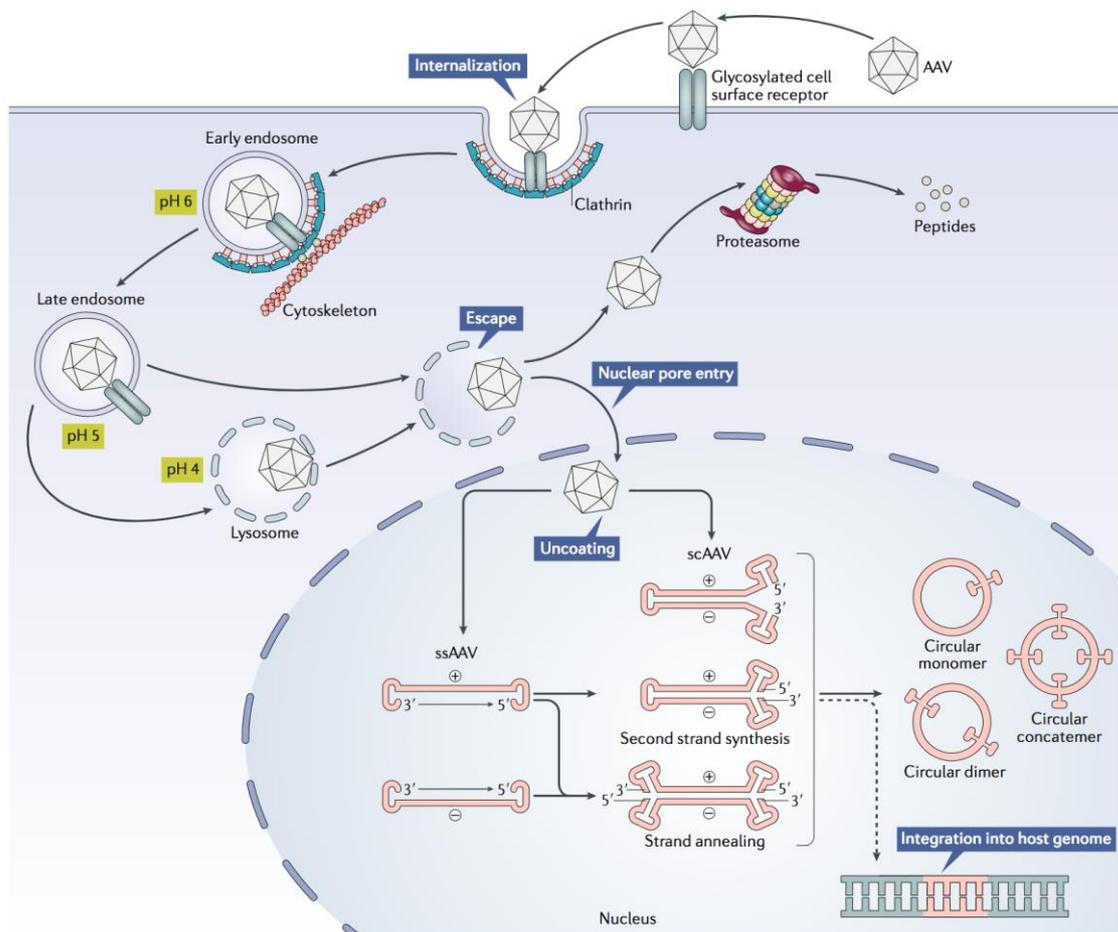
**Figure 4. Schematic representation of the structure (a) and genome (b) of the AAV.** Figure issued from Kotterman *et al.*<sup>8</sup>.

Two main **open reading frames** are present in the genome: *Rep* and *Cap* (Figure 4). *Rep* encodes for 3 proteins (p5, p19 and p40) responsible for **virus replication** and integration into the host genome, whereas *Cap* encodes for the **3 capsid proteins** (VP1, VP2, VP3) as a result of alternative splicing of a single mRNA<sup>8,14</sup>. Several alternative open reading frames exist in the *Cap* gene, allowing the production of accessory proteins of unclear function, such as AAP<sup>8</sup>, AAVX<sup>15</sup>, MAAP<sup>16,17</sup>. The AAV **viral capsid** contains 60 subunits with a VP1:VP2:VP3 stoichiometry of 5:5:50<sup>14</sup>. More than one hundred natural serotypes have been identified, not only from human or non-human primate tissues but also from distant species (goats, sea lions, bats, snakes...), and at least as many chimeric serotypes have been generated<sup>8,18</sup>. Most of the differences between serotypes are located within the **VP3 subunit** in multiple hypervariable regions<sup>19</sup>. The VP3 subunit is exposed at the surface of the AAV capsid and plays a role in the **tropism of the serotype** (Table I)<sup>20,21</sup>.

Serotype	Origin of isolation	Receptors & coreceptors	Tissue tropism (in monkey/human)
<b>Natural AAV capsids</b>			
AAV1	Monkey	Sialic acid, AAVR	Muscle, CNS, Airway, Liver
AAV2	Human	Heparin, AAVR, Integrin, FGFR, HGFR, LamR	Liver, CNS, Retina, Muscle, Airway
AAV3	Human	Heparin, AAVR, HGFR, LamR	Liver, CNS
AAV4	Monkey	Sialic acid	CNS, Eye, Airway
AAV5	Human	Sialic acid, AAVR, PDGFR	Liver, CNS
AAV6	Human	Heparin, Sialic acid, AAVR, EGFR	CNS, Muscle, Heart, Airway
AAV7	Monkey	Unknown	CNS, Muscle
AAV8	Monkey	AAVR, LamR	Liver, Heart, Muscle, CNS, Eye
AAV9	Human	Galactose, AAVR, LamR	Liver, Heart, Muscle, CNS, Eye
AAVrh10	Monkey	LamR	Muscle, CNS
AAV11	Monkey	Unknown	Unknown
AAV12	Human	Unknown	Unknown
AAVrh74	Monkey	Unknown	Muscle
<b>Selected AAV chimeric capsids</b>			
AAV-LK03	Rational shuffling (AAV3B/4)	Not tested	Liver <sup>22,23</sup>
AAV-DJ	Rational shuffling (AAV2/8/9)	Not tested	Liver, CNS <sup>24,25</sup>
AAV-MYO	Peptide insertion	Integrin	Muscle (cells) <sup>26,27</sup>

**Table I. Tropism of natural and selected chimeric AAV capsids**, based on Pupo et al.<sup>18</sup>, Lisowski et al.<sup>28</sup>, Li et al.<sup>29</sup>. AAVR = AAV receptor; CNS = central nervous system; EGFR = epidermal growth factor receptor; FGFR= fibroblast growth factor receptor; HGFR = hepatocyte growth factor receptor; LamR = laminin receptor 1; PDGFR = platelet-derived growth factor receptor.

The entry of AAVs into the cell is complex and does not only rely on a sole interaction between single ligand and receptor<sup>20</sup>. The AAV initially binds to cell surface glycans such as heparan sulfate, sialic acid or galactose according to the serotype, then most serotypes rely on the **AAV receptor** (AAVR) for endocytosis, cellular sorting and trafficking<sup>20,30</sup>. Involved mechanisms are summarized in **Figure 5**. Once internalized, the AAV leaves the endosomal compartment (*endosomal escape*) and reaches the nuclear pore complex through a poorly understood mechanism<sup>31</sup>. The AAV enters the nucleus using the cellular nuclear import machinery, the capsid opens (*uncoating*) and the **genome is released into the nucleus**<sup>32</sup>. Some steps of this pathway require the correct function of the intracellular trafficking machinery as well as autophagy. Although not studied for wild-type AAV cell transduction, the **induction of autophagy** has been shown to promote higher *in vivo* transduction by rAAV vectors<sup>33</sup>.



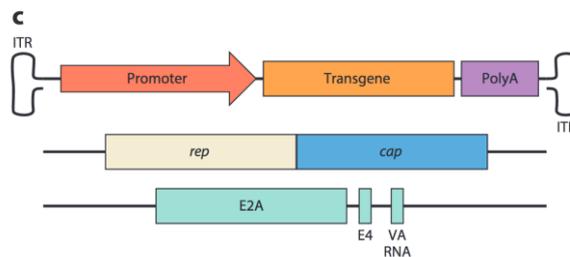
**Figure 5.** Entry of the AAV into the cell. Figure from Wang et al<sup>31</sup>.

Once in the nucleus, the genome is single stranded and the synthesis of the second strand is required, generating a double-stranded DNA genome<sup>34</sup>. In wild-type AAVs, the presence of the proteins encoded by the *Rep* gene allows the **integration of the genome** into the host DNA and long term persistence<sup>8,31</sup>. A co-infection by viruses such as adenoviruses or human herpes virus 6 (HHV6)<sup>35</sup>

and production of *early genes* (E1, E2A, E4) will induce expression of the AAV genome and allow AAV replication<sup>31</sup>.

## 2. RECOMBINANT AAV VECTORS (rAAVs)

In a rAAV, the viral *Rep* and *Cap* genes have been replaced by an expression cassette containing the gene of interest (transgene) under the control of regulatory sequences (promoter ± enhancers) and followed by a poly-adenylation signal, as shown in **Figure 6**<sup>2,8</sup>. The presence of AAV2-ITRs flanking the expression cassette allows its encapsidation<sup>2</sup>. One method of production of rAAVs is the **triple transfection** of competent cells (such as HEK293T) with a plasmid containing the expression cassette, a plasmid with *Rep* and *Cap* genes (possibly different from AAV2) and a “helper” plasmid with adenovirus *early genes*<sup>2,31</sup> (**Figure 6**). The *Rep* and *Cap* proteins will allow genome replication and viral capsid assembly, and then encapsidation of the expression cassette into the nascent virions<sup>2,8</sup>. The *Cap* gene determines the **serotype** of the rAAV and therefore the properties of the rAAV in term of tropism (**Table I**) and/or immune escape. The size of the expression cassette will determine the rate of genome encapsidation (full particles), with efficient encapsidation for expression cassette up to 5 kb<sup>31</sup>.



**Figure 6. Schematic representation of the rAAV genome.** Accessory plasmids required for rAAVs production are shown, one encoding for viral *Rep* and *Cap* genes and the other (helper plasmid) for the adenovirus genes E2A and E4. HEK293T cells constitutively express the required E1 antigen<sup>18</sup>. Figure issued from *Kotterman et al.*<sup>8</sup>.

Because of the absence of the *Rep* gene, rAAVs present only **low frequency random integrations** in the host genome, close to 0.1% of cells but difficult to quantify precisely<sup>8,36</sup>. rAAVs persist in the nucleus as **episomes** and is used to target tissues with no or low renewal potential, since the episome will be diluted after sequential mitosis<sup>2,8</sup>. One of the limiting steps of rAAVs genome expression is the second strand synthesis in the nucleus. This step can be bypassed by using a **self-complementary** vector in which one half of the expression cassette is complementary of the second half (**Figure 5**)<sup>37</sup>. The limitation of this approach is the short maximum size of the expression cassette<sup>37</sup>, which is reduced by a factor 2.

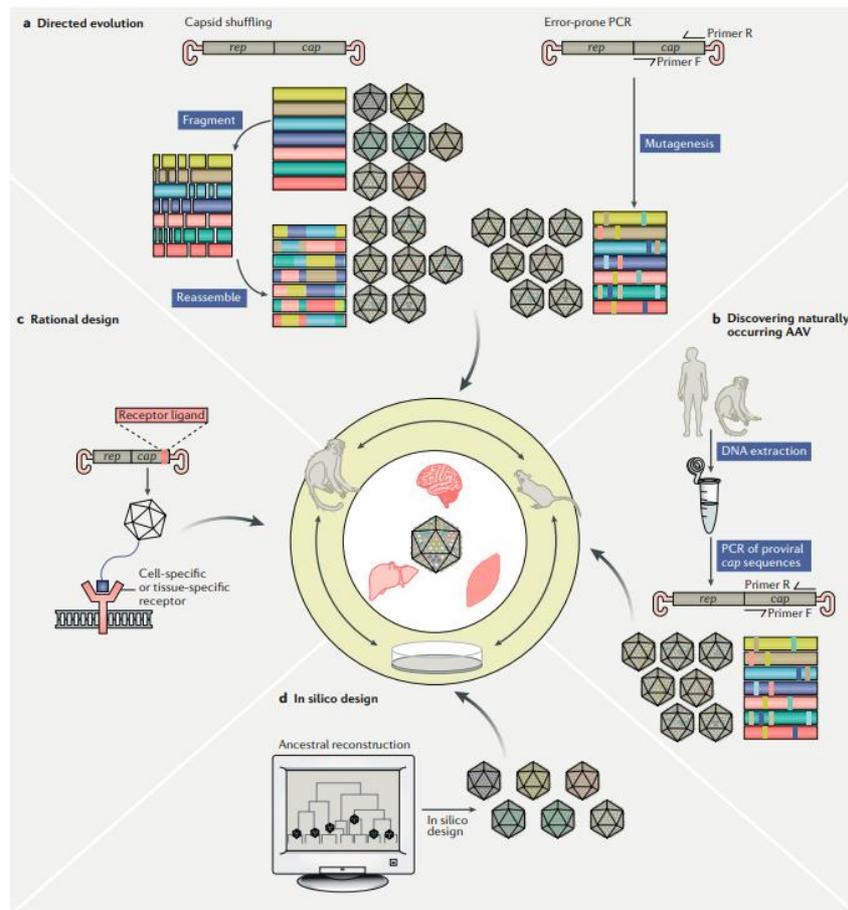
Once in the nucleus as a double stranded DNA, the transgene can be expressed depending on the **regulatory sequences** placed upstream (**Figure 6**). These sequences consist of a **promoter** and sometimes **enhancer** sequences<sup>38</sup>. The promoter usually exhibits the **transcription initiation site** (TSS or “+1”) and binding sites for the polymerase and general transcription factors, the most well-known being the *TATA box* and *Initiator* sequences<sup>38,39</sup>. Enhancer sequences increase the transcription through different transcription factor binding sites, without having the other attributes of the promoter (TSS, TATA box, Initiator)<sup>38</sup>. The transcription factor binding sites present in both types of sequences are responsible for the tissue selectivity of the promoter<sup>38</sup>. Some are specific of a given tissue for instance the apolipoprotein E enhancer and human  $\alpha$ 1-antitrypsine promoter (ApoE-hAAT) for the liver<sup>40</sup> or desmin promoter for the muscles<sup>38</sup>. Other are ubiquitous, such as the cytomegalovirus promoter (CMV) or CMV enhancer and chicken beta actin promoter (CAG). To date, it is very difficult to predict the strength and the tissue selectivity of the promoter/enhancer based only on its sequence and an evaluation *in vitro* or preferably *in vivo* is required. As recently described, potent **dual-specificity tandem promoters** can be obtained by combining different regulatory sequences for liver and muscle for instance<sup>41,42</sup>.

Finally, another method to increase transgene expression is **codon optimization**<sup>43</sup>. The genetic code is redundant, with between 1 and 6 synonymous codons for each amino acid, and all transfer RNA (tRNA) responsible for codon decoding are not expressed in the same amount in the cell<sup>43</sup>. Therefore, selecting the synonymous codons associated with the higher frequency of tRNA can optimize the **level and speed of transgene mRNA translation**, without altering the protein sequence of the transgene<sup>43</sup>. Interestingly, the most expressed genes in the cell also have a bias toward these types of codons<sup>43</sup>. It is currently not possible to predict accurately the effect of such modifications, which have to be evaluated *in vitro* or preferably *in vivo*. An open question is whether the same codon optimization will be of similar efficacy in different species, since it is most often selected in mice to be used in humans. Nonetheless, most clinical trials use this approach for the design of the expression cassette<sup>1</sup>.

### 3. CAPSID DISCOVERY AND ENGINEERING

Multiple novel chimeric capsids have been engineered during the past 10 years, which largely broaden the AAV capsid repertoire available for gene therapy applications<sup>18,31</sup>. These capsids have new interesting properties such as a **better targeting** of a specific organ, such as the liver<sup>23</sup>, the central nervous system<sup>44</sup> or the muscles<sup>26</sup>, could **reduce unwanted targeting** and therefore improve safety<sup>26</sup>, and could be **less impacted by neutralizing antibodies** (NAbs) derived from natural AAV infections<sup>31</sup>. Several methods have been developed for the generation of novel interesting capsid and include

natural discovery and/or directed evolution and/or rational design (reviewed in Wang et al.<sup>31</sup>, Figure 7) with potential combinations of these different techniques.



**Figure 7. Methods for capsid discovery and engineering**, issued from Wang et al.<sup>31</sup>.

The first historical method was the **natural discovery** of AAV capsids in human or non-human primate tissues, such as liver. Among the most potent and widely used rAAVs have been obtained this way, such as AAV8 and AAV9 for instance<sup>45,46</sup>. Since the **liver is a frequent site of natural infection** in both humans and non-human primates, many AAVs have been retrieved from liver samples<sup>31</sup>. This method requires the screening of tissue samples for complete proviral DNA sequences integrated in the host genome. Although this approach usually leads to the retrieval of efficient human-adapted AAV capsids when using human samples, one major drawback is the fact that these capsids are often **endemic in humans** with frequent pre-existing immunity, which limits their use in patients<sup>47,48</sup>. To circumvent this issue, several studies have focused on large animal models such as pigs in order to obtain efficient capsids with potential lower NAb seroprevalence<sup>49</sup>. This strategy can also lead to the discovery of AAV capsids with new properties, such as the recently reported **AAVHSC**, isolated from human CD34<sup>+</sup>-hematopoietic stem cells, which display nuclease-free gene editing abilities<sup>50</sup>. These

serotypes, belonging to clade F, have higher site-specific homology-dependant integration abilities compared to conventional serotypes in human cells and *in vivo* in mice<sup>50</sup>.

The second strategy is the **rational design** (Figure 7), which aims at introducing specific modifications in the VP3 protein in order to alter the tropism or the immunogenicity of the capsid<sup>31</sup>. This requires the **elucidation of structure-function relationships**, of the 3D-structure of the AAV capsid and/or of the mechanisms of cellular entry of the AAVs (reviewed in Meyer *et al.*<sup>20</sup>). In the VP3 protein, most of the variations between serotypes are clustered in specific **hypervariable regions**<sup>19</sup>. Many rationally-designed AAV capsids harbor modifications in one of these hypervariable regions, for instance the insertion of a specific peptide improving the transduction of a specific cell type (reviewed in Büning *et al.*<sup>51</sup>). One example is the use of an **RGD-containing peptide** which allows the improved targeting of muscles through the recognition of specific integrins expressed by muscle cells<sup>26</sup> (further discussed below). Another example is the introduction of point mutations to **remove the exposed tyrosine residues** within the VP3 protein of AAV2, aiming at reducing the ubiquitination and subsequent degradation of virions once inside the cell<sup>52</sup>. These modifications were shown to enhance transgene expression *in vivo*<sup>52</sup>. Finally, a last example is the use of *in silico* analyses to generate potential novel AAV capsids<sup>31,53</sup>. This approach was used to reconstruct the phylogenetic tree of current “natural” AAVs serotypes, in order to predict the sequence of a **potential AAV ancestor**, named Anc80<sup>53</sup>. This capsid showed interesting liver tropism in non-human primates with low cross-reactivity with natural serotypes<sup>53</sup>.

Another method to generate new AAV capsid variants is **directed evolution** (Figure 7). This approach simulates natural evolution by applying a selective pressure on AAV capsids in order to select capsid variants with interesting properties such as specific tissue tropism, immune escape or transgene expression<sup>31</sup>. Interestingly, this strategy **does not require knowledge on the 3D structure of AAV capsids** or on the mechanisms of cellular entry. It requires the generation of a large library of AAV capsids followed by a step of selection. A high diversity of AAV capsids can be generated either using **error-prone PCR** or **capsid shuffling**<sup>31</sup>. With the first strategy, amplification of the *Cap* gene using an error-prone polymerase will lead to the generation of a library of AAV capsid variants with random mutations. In the second strategy, the *Cap* genes of different serotypes are fragmented and randomly reassembled in order to generate a pool of new *Cap* genes. Then, a selection step is performed, for instance by injecting these AAV variants in mice or primates, and the most interesting capsids are retrieved from the organ of interest. One limitation of this approach is that, since it is not targeted, most of the capsids obtained this way are not functional and the final yield is low unless a **high-throughput screening** can be performed<sup>54</sup>. The development of next-generation sequencing along with the “barcode” technology allows the screening of several thousands of AAV variants at the same time, with iterative steps of generation and selection<sup>26,31</sup>. Directed evolution has led to the generation

of interesting AAV capsid variants with **high liver targeting** such as LK03<sup>23</sup> or AAV-DJ<sup>25</sup>. A potential limitation is the selection of a specie-specific AAV variant, especially if the selection is performed in mice. For instance, the AAV capsid PHP.B has been selected for its high central nervous system tropism in mouse, but failed to outperform AAV9 in non-human primates<sup>44</sup>. To circumvent this issue, the use of humanized mouse models for the selection step, as reported for the identification of LK03<sup>23</sup>, is an interesting strategy.

#### 4. MAIN TECHNICAL LIMITATIONS

rAAVs are a very potent platform for gene transfer in tissues such as the liver, muscles or central nervous system, but face two major technical limitations: **neutralizing antibodies**, either preexisting or produced after rAAV injection, and a **low packaging capacity**<sup>55</sup>.

##### a. NEUTRALIZING ANTIBODIES

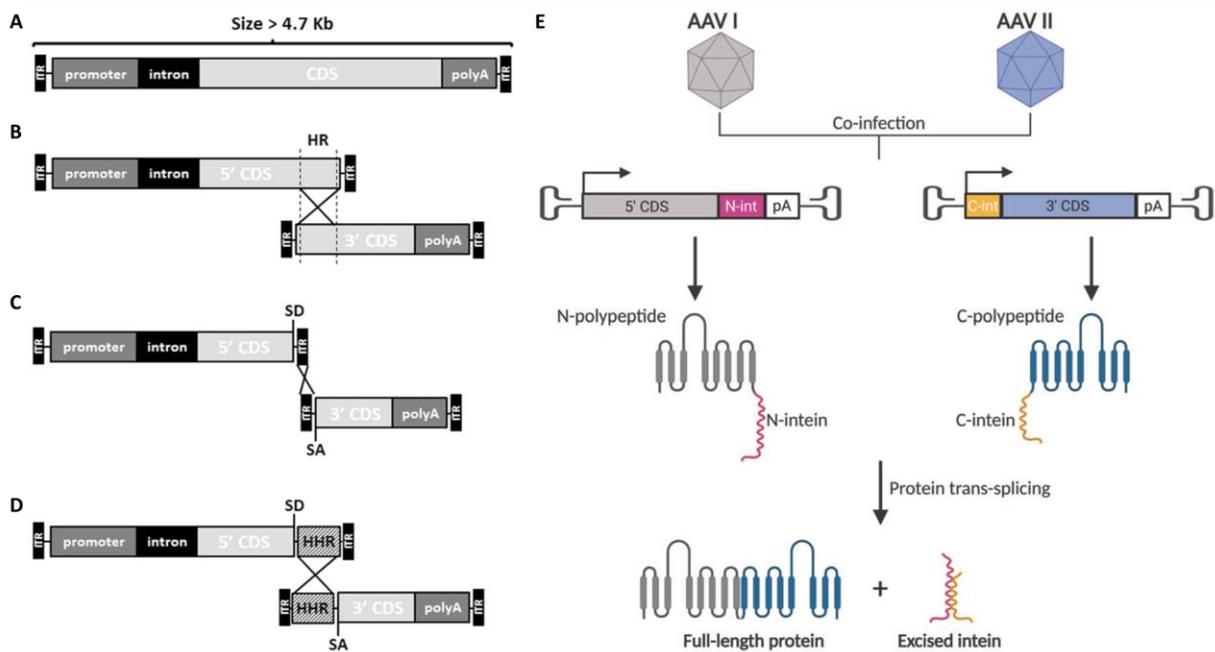
The **seroprevalence of preexisting neutralizing antibodies** (NAbs) depends on the serotype and can reach 50 - 60% for AAV1 and AAV2, 20 - 30% for AAV8 and AAV9 and < 10% for AAV5 in adults<sup>48</sup>. NAbs seroprevalence varies with age, with **lower rates in young children** in line with the progressive anti-AAV natural immunization during childhood<sup>47,56</sup> and maybe a potential cross-reactivity with other common parvoviruses. **Elevated anti-AAV NAbs titers prevent the use** of the particular serotype<sup>57,58</sup>. This affects up to 50% of patients for some serotypes<sup>48</sup> and is a major limitation of the use of the rAAV technology. Furthermore, very high anti-rAAV NAbs titers persist after rAAV injection for as long as 15 years after initial gene transfer as shown in the earliest hemophilia B gene therapy clinical trials<sup>59</sup>, **precluding any possible reinjection**. Cross-reactivity of these Nabs occurring after rAAV injection is frequent, although at lower titers, and can be problematic for potential readministration using a different serotype at least in humans<sup>59</sup>.

Therefore, several approaches have been designed to escape the preexisting NAbs, to reduce their titers or to prevent their development after rAAV injection and include, among others (reviewed in *Gross et al.*<sup>60</sup>) :

- The **generation of chimeric AAV capsids** with lower seroprevalence in the population<sup>31</sup>
- The **escape from preexisting NAbs** using exosome-enveloped rAAVs<sup>61</sup> or empty capsids used as decoy<sup>62</sup>
- The **depletion of preexisting NAbs** using plasmapheresis<sup>63</sup> or the immunoglobulin G-cleaving endopeptidase IdeS<sup>64</sup>
- The **prevention of NAbs generation after rAAV injection** using co-administration of rapamycin-loaded nanoparticles<sup>65</sup> or B-cell depleting agents such as rituximab<sup>66</sup>

b. LOW PACKAGING CAPACITY

The **size of the rAAV genome** that can be packaged is another important limitation of the use of rAAVs in gene therapy<sup>55</sup>. Although the size of the wild type AAV genome is 4.7 kb<sup>8</sup>, studies of rAAV production have shown that an expression cassette with a size up to 5 to 5.2 kb could be efficiently encapsidated<sup>67</sup>. Above 5.2 kb, multiple truncated genomes are encapsidated<sup>67</sup>. Importantly, many cDNA of clinical interest have a size close to or above 5 kb<sup>55,68</sup>, mostly structural protein or membrane transporters involved in liver diseases (familial cholestasis, Wilson disease) or ophthalmic diseases (retinitis pigmentosa). One extreme case is the *dystrophin* gene, involved in Duchenne Muscular Dystrophy (DMD) and whose cDNA is close to 14 kb. In contrast, most enzymes usually have a small cDNA size, except the glycogen debranching enzyme which is the focus of our work.



**Figure 8. Different dual vectors approaches allowing gene transfer of cDNA above the encapsidation limit of rAAVs**, issued from *Colella et al.*<sup>55</sup> for panel A-D and from *Trapani et al.*<sup>68</sup> for panel E. HR = homology region; HHR = highly recombining region; SD = splicing donor site ; SA = splicing acceptor site ; CDS = coding sequence.

Two main strategies were developed to use rAAV-based gene therapy when the expression cassette including the transgene and mandatory regulatory sequences is above 4.7 to 5 kb (**Figure 8A**): the use of a **dual vector strategy** or the use of **truncated version of the transgene** which could be encapsidated in a single rAAV vector.

The **dual vector strategy** consists in splitting the expression cassette in two, with a first vector containing the promoter and the 5' part of the cDNA and the second vector containing the 3' part of

the cDNA and the poly-A signal<sup>55</sup>. In order to allow the reconstitution of the full-length transgene *in vivo*, several strategies can be used:

- The use of a **common region from the cDNA** (in 3' of the 1<sup>st</sup> vector and 5' of the 2<sup>nd</sup> vector) which will allow homologous recombination in the cell and the reconstitution of a full length expression cassette (**Figure 8B**)<sup>55</sup>
- The use of an **exogenous recombining region (HHR)** in both vectors (in 3' of the 1<sup>st</sup> vector and 5' of the 2<sup>nd</sup> vector), which will allow the homologous recombination in the cell. In order to restore the reading frame, donor and acceptor splicing sites must be added in order to remove this HHR from the final matured mRNA (**Figure 8D**)<sup>55</sup>.
- To rely on rAAV concatemerization (binding through ITRs) and trans-ITR transcription to generate a mRNA containing the full-length transgene. In order to restore the reading frame, donor and acceptor splicing sites must be added in order to remove the 3' ITR from the 1<sup>st</sup> vector and 5' ITR from the 2<sup>nd</sup> vector from the final matured mRNA (**Figure 8C**)<sup>55</sup>

A more recently described dual vector strategy does not rely on homologous recombination but requires the use of small peptides named *inteins* allowing **protein trans-splicing**<sup>68</sup>. Inteins are auto-catalytic microbial proteins able to ligate two proteins while excising themselves, as shown on the **Figure 8E**<sup>68</sup>. By placing the coding sequence of the N-terminal intein in 3' of the 1<sup>st</sup> vector and the coding sequence of the C-terminal intein in 5' of the 2<sup>nd</sup> vector, one can obtain the generation of a full-length protein after removal of the two inteins<sup>68</sup>.

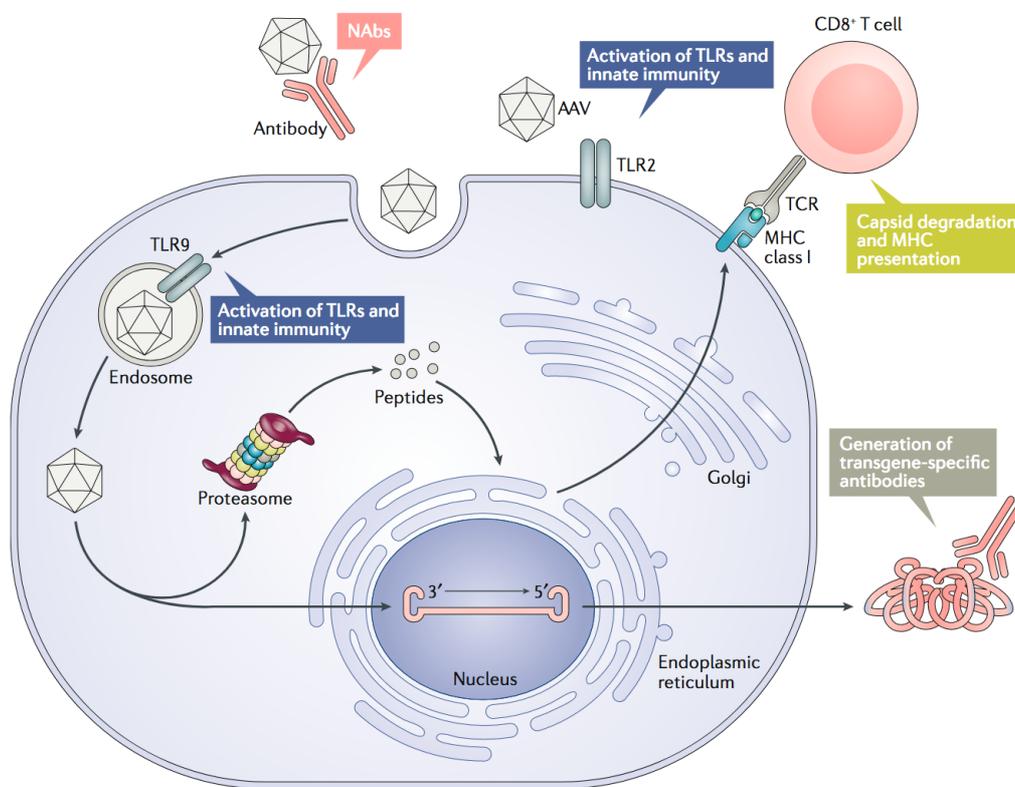
This technology of dual vector has been used **successfully in preclinical models** of glycogen storage disease type III (GSDIII)<sup>69</sup>, DMD<sup>70</sup>, dysferlinopathy<sup>71</sup>, Wilson disease<sup>72</sup> and of retinitis pigmentosa<sup>73</sup>. Specific limitations include the **requirement of two vectors** which doubles the vector dose received by the patient and therefore potentially the toxicity, the **suboptimal efficiency of recombination**, which can reduce the efficacy and result in the accumulation of truncated N-terminal proteins, and finally the **use of microbial peptides** which could potentially induce an immune response after accumulation of excised inteins<sup>55</sup>. Such an approach could nonetheless be interesting for the eye, because of a **locally-injected vector** which could favor the co-infection of the same cell by the two vectors and therefore an increased efficacy, and the **relative immune privilege** of the eye<sup>74</sup> that may result in a lower immune-mediated toxicity.

The second strategy that can be used when the expression cassette is larger than the packaging capacity of rAAVs is the **truncation of the transgene**. To generate a truncated version of a protein can be challenging, as the truncation can affect the function, the stability and/or the regulation of the protein. Nonetheless, several successful examples have been published, such as the B-domain deletion

mutant of the factor VIII (FVIII-SQ) used in the clinic for hemophilia A<sup>75</sup> and which has now market approval (Roctavian®). Other examples include the deletion mutants for the dystrophin and for the liver protein ATP7B, used in clinical trials for DMD<sup>76</sup> and for Wilson disease<sup>77</sup>, respectively.

## 5. IMMUNE RESPONSE-ASSOCIATED TOXICITIES

Although rAAV vectors have overall a good safety profile, clinical trials have revealed several unexpected complications, poorly predicted by preclinical studies in large animal models: **cellular immune response** to the capsid antigens, **acute toxicities** at high vector dose (liver injury, thrombotic microangiopathy) and possible **immune response to CpG** present in the expression cassette.



**Figure 9. Schematic representation of immune system activation following rAAV injection.** From Wang *et al*<sup>31</sup>.

### a. CELLULAR IMMUNE RESPONSE

The **development of an immune response** is an important complication of the use of a viral vector<sup>78</sup>. After entry of the vector into the cell, capsid proteins are degraded through the proteasome pathway and presented to CD4<sup>+</sup> and CD8<sup>+</sup>-T cells via the major histocompatibility complex<sup>79</sup> as shown in **Figure 9**. These peptides induce a **cellular immune response** toward the capsid and lysis of the transduced cells. This response is absent of most of the animal models. Early clinical trials for hemophilia have revealed the presence of a **dose-dependent T-cell response** to the capsid antigens

occurring between **4 and 8 weeks** post-injection that can result in the loss of the transduced cells and a loss of correction<sup>80,81</sup>. This immune response can be controlled using **corticosteroids**<sup>81,82</sup>. Peptide presentation to CD4<sup>+</sup>-T cells also leads to B-cell priming and **NAbs generation**, preventing any further rAAV administration (see above). In contrast, no significant immune responses toward the transgene have been demonstrated in liver-targeted gene therapy<sup>81</sup>, whereas some have been reported in clinical trials targeting muscles especially for DMD<sup>83,84</sup>. Species-specific humoral response to the transgene is present in preclinical models, likely due to the expression of human transgenes in rodents<sup>85</sup>. Such an immune response has not been documented in clinical trials targeting the liver (especially hemophilia), even in patients with complete absence of the protein, possibly in line with liver-induced tolerance (see below)<sup>81</sup>.

#### b. ACUTE TOXICITY

Signs of **acute liver toxicity** have been observed after high-dose rAAVs injection in clinical trials for spinal muscular atrophy (SMA) and myotubular myopathy (MTM)<sup>86,87</sup>. Concerning SMA, almost 90% of patients displayed elevation of liver enzymes and few had liver failure within a week of the injection<sup>86</sup>. Concerning MTM, four patients died from fulminant liver failure within a month after rAAV infusion<sup>87</sup>. In these trials, **very high-doses of rAAVs** were used (1 to 3 x 10<sup>14</sup> viral genome (vg)/kg)<sup>86,87</sup> and MTM patients presented with a **poorly characterized preexisting liver disease** known as peliosis hepatis<sup>87,88</sup>. Such a toxicity seems too early for being due to a T-cell mediated response (at least in the SMA trial). Furthermore, the absence of T-cell infiltrates in liver biopsies and the absence of response to classical immunosuppressive drugs in the 4 MTM patients who died led to suspect the **direct toxic role of high-dose rAAVs** and/or an **activation of innate immune response** ( $\pm$  complement)<sup>89</sup>. One recent study performed in non-human primates reports the occurrence of dose-dependent liver injuries within a week from the injection of an rAAV vector close to the one used in the SMA trial, without immune infiltrates, thus suggesting a **direct hepatic toxicity of the vector**<sup>90</sup>. In any cases, data are scarce and a close monitoring of both liver damage and potential immune response is mandatory. It is interesting to note that the presence of **preexisting liver damage** seems to increase the risk of severe liver decompensation following the injection of high-dose rAAVs, as this could also be relevant for GSDIII.

Another unexpected complication is the **occurrence of thrombotic microangiopathy** in patients involved in the clinical trials for DMD<sup>83</sup> and for methylmalonic acidemia (MMA), as well as in patients treated for SMA with Zolgensma<sup>89</sup>. Most of the patients responded well to symptomatic treatment, but one death was recently reported in a patient with SMA<sup>91</sup>. The exact mechanism of this complication is unknown, but **complement activation** is known to play a central role in “conventional”, gene therapy-unrelated, thrombotic microangiopathy<sup>92</sup>.

### C. CPG-INDUCED IMMUNE RESPONSE

A role for the **CpG content of the expression cassette** in immune response to the rAAV and long-term persistence in the liver has recently been suggested in two clinical trials for hemophilia B<sup>93,94</sup>. CpG from the expression cassette can be methylated during the synthesis of the plasmid used for rAAVs production. These **methylated CpG** are recognized by the innate immune system through the **TLR9 protein**, and could trigger both inflammation and loss of corrected cells<sup>95</sup>. In both aforementioned clinical trials, the rAAV vector had a CpG-rich expression cassette and almost all patients displayed a **sudden loss of expression** of the factor IX between 6 and 12 weeks post-injection, despite an immunosuppression by corticosteroids<sup>93,94</sup>. Several patients also displayed T-cell response toward the capsid antigens in ELISPOT<sup>93</sup>. Partial data in mice suggest that CpG-rich expression cassette have **increased immunogenicity**<sup>93,95</sup>. In contrast, this sudden loss of expression was not observed in other clinical trials in which the expression cassette did contain only a small number of CpG (reviewed in *Nathwani et al.*<sup>81</sup>) and it is possible that the activation of the innate immune system may have resulted in a subsequent adaptive immune response and loss of the transduced cells.

## 6. TUMORIGENESIS

The **potential risk of tumor development** after the use of viral vectors has been a concern since the occurrence of leukemia in the first clinical trials with  $\gamma$ -retroviruses in the 2000s<sup>4</sup>. Evidences gathered so far suggest a **good safety profile for rAAVs** as for tumor development<sup>1</sup>. Nonetheless, controversies exist about potential rAAV-induced tumors.

The wild type AAV has the ability to **integrate in the host genome** through the Rep proteins<sup>8</sup>. Although considered for long as non-pathogenic, several studies reported the occurrence of wild type AAV clonal integration in **hepatocellular carcinoma** (HCC) in patients, often young and without cirrhosis, and with **insertional mutagenesis** and oncogene expression dysregulation<sup>35,96</sup>. This suggest a potential direct oncogenic risk in humans<sup>35,96</sup>. The **3'-ITR is almost constantly observed** within these integrations and is able to activate the expression of nearby oncogenes through a **liver-specific enhancer sequence** located within the 3'-ITR of AAV2<sup>97</sup>.

In contrast, the **risk of HCC seems low** with rAAVs mainly due to the replacement of the *Rep* gene by the expression cassette. Therefore rAAVs display only very low integration abilities<sup>2,8</sup>. Of note, the 3'-ITR liver-specific enhancer sequence has also been removed from the newest rAAV vectors. Nonetheless, **random genomic insertions** were observed with a frequency ranging from 0.1 to 1% of hepatocytes in the liver in preclinical models<sup>98,99</sup> and were lower than 0.5% in biopsies from treated patients in one clinical trial<sup>36</sup>. Studies in **adult mice** did not find an increased risk of HCC or of gene dysregulation after rAAVs injection<sup>100,101</sup>. In opposite, the incidence of HCC is very important

(sometimes above 50%) in mice treated from the **neonatal period**, because of insertion of the vector in the **Rian locus** within the *Mir341* gene<sup>102–104</sup>. This locus is highly expressed during embryonic development and in neonates<sup>105</sup>, which could explain the high risk of HCC restricted to this period of life. *Rian* locus orthologs exist in most species, but without the **rodent-specific *Mir341* gene**, and this region is not an integration hotspot in adult non-human primates and in patients<sup>36</sup>. No HCC were identified in cats more than 8 years post-injection<sup>106</sup> as well as in dogs<sup>99</sup>. Nonetheless, one recent study in dogs found clonal insertions of the vector in the liver of two dogs, close to known oncogenes, with **clonal expansion** of the cells but without HCC 6 to 10 years after rAAV injection<sup>99</sup>.

In non-human primates as well as in patients, **no insertional mutagenesis** has been described, with integration distributed randomly throughout the genome<sup>36</sup>. Data from clinical trials also show the **safety of rAAV vectors**, with no rAAV-induced HCC with follow-up duration up to 12 to 15 years for the earliest gene therapy trials for hemophilia B<sup>59</sup>. Of note, one patient included in the hemophilia A trial from Uniqure (AMT-061) developed **HCC one year after rAAV injection**<sup>107</sup>. This patient displayed further risk factors (viral hepatitis B and C, fatty liver disease) and **did not have any rAAV-related clonal insertion**, thus suggesting an unrelated HCC<sup>107</sup>. To conclude, although preclinical studies showed the occurrence of random integrations and low-frequency clonal expansions, clinical data suggest a good safety profile with prolonged follow-up, up to 15 years.

### **III. LIVER-TARGETED GENE THERAPY WITH A FOCUS ON RAAV**

The liver is the **most frequently targeted organ** in rAAV-based gene therapy, accounting for a third of all clinical trials on the period 2015-2019<sup>6</sup>. Most trials target **hemophilia A and B** or **inborn errors of metabolism**.

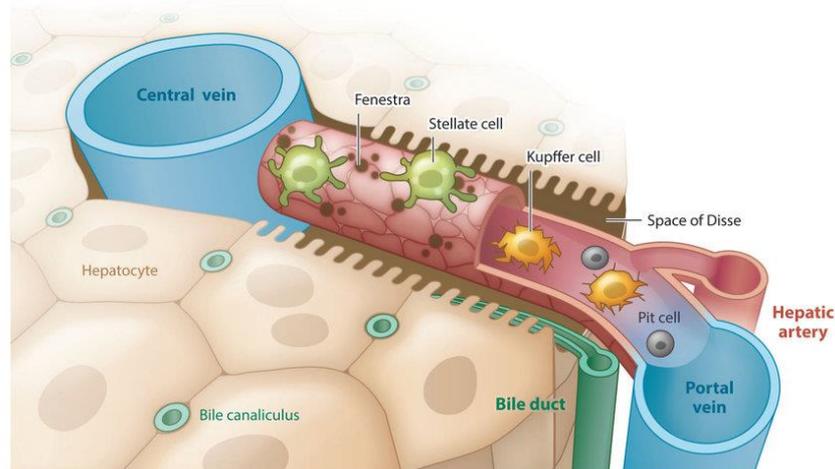
#### **1. ADVANTAGES OF LIVER TARGETING**

Several features make the liver an ideal target for rAAV-based gene therapy approaches: its **important blood flow**, the presence of **endothelial fenestrations**, its role in the **biosynthesis of plasma proteins**, its role in **tolerance** and importantly the **high tropism** of AAV for the liver.

##### a. ANATOMICAL ASPECTS

The liver receives both **arterial and venous blood** from the hepatic artery and the portal vein respectively and an **important blood flow** estimated between 20 to 25% of cardiac output<sup>108</sup>. Blood flow travels through the hepatic lobule inside sinusoids to reach the central veins and thus the systemic venous circulation (**Figure 10**). In the sinusoids, the **endothelium is discontinued** because of the

presence of fenestrations between endothelial cells, allowing the flow of small molecules<sup>109</sup>. These fenestrations have a diameter from 50 to 100 nm<sup>109</sup>, much larger than an AAV particle (around 25 nm)<sup>31</sup>. Therefore, rAAVs can easily cross the vascular barrier and access to the hepatocytes, favoring liver transduction.



**Figure 10. Schematic representation of the hepatic sinusoid, from Ehrlich et al.<sup>110</sup>.**

#### b. PLASMA PROTEINS BIOSYNTHESIS: THE “BIOFACTORY”

The liver is responsible for the synthesis of most of the plasma proteins, such as albumin,  $\alpha$ -globulins,  $\beta$ -globulins and blood clotting factors, and hepatocytes have a **highly developed secretory pathway**<sup>111</sup>. It is possible to use this particular feature in order to produce high amount of a protein in the plasma by targeting the liver: it is the **concept of “biofactory”**<sup>112</sup>. Many preclinical studies have shown the feasibility to produce proteins **normally synthesized by the hepatocyte** (factor IX) **or liver endothelial cells** (factor VIII), but also proteins **not normally secreted** by the liver but able to act afar in order to correct other organs<sup>81,112,113</sup>. Such proteins include lysosomal enzymes, for instance acid  $\alpha$ -glucosidase involved in Pompe disease<sup>112</sup>, or Fgf23 inhibitors involved in X-linked hypophosphatemia (XLH)<sup>113</sup>.

#### c. IMMUNE TOLERANCE

The liver has peculiar **immune tolerance abilities**, allowing the tolerance of food antigens and commensal microorganisms<sup>114</sup>. These features are known for long and explain the occurrence of **chronic viral hepatitis**<sup>115</sup> and the possibility to perform **liver transplantation** in absence of human leukocyte antigen (HLA) matching and with a lower immunosuppression compared to heart and kidneys transplantations<sup>116,117</sup>. This **tolerance goes beyond the liver** and explains the better outcome of combined liver and kidney transplantation compared to isolated kidney transplantation<sup>116</sup>.

Similar findings have been demonstrated for rAAV gene therapy. Indeed, liver transduction is associated with partial tolerance and long-term transgene expression, as demonstrated in clinical trials for hemophilia<sup>81</sup>. Furthermore, preclinical studies have also suggested that targeting of the liver allows a **tolerance toward a transgene expressed in the periphery**<sup>118,119</sup>, similar to the combined liver-kidney transplantation: in these studies, the intramuscular injection of an rAAV induces a potent immune response to the transgene and a loss of corrected cells. In contrast, intravenous injection of the rAAV, in combination to the intramuscular injection, allows a tolerance and persistence of muscular (and liver) transduced cells<sup>118</sup>. Precise mechanisms are not yet fully understood and might involve the generation of regulatory T-cells (Treg) and/or the deletion or exhaustion of effector T-cells<sup>114,115,118</sup>.

#### d. NATURAL TROPISM OF AAV

Above all, a major advantage of the rAAV for liver targeting is their **potent tropism for this organ**. Almost all natural and chimeric capsids have strong liver transduction, even at **relatively “low” doses**<sup>31</sup>. Although it is a strong advantage for liver targeting, on the other side this is a major limitation for targeting other organs such as muscles and the central nervous system. These tissues being difficult to target, **high vector doses are required** (usually above  $5 \times 10^{13}$  vg/kg in human) with the risk of liver overloading. A recent study reporting the autopsy of two patients deceased from SMA despite having received the gene therapy product Zolgensma® showed a transduction at least 100 times more important of the liver compared to the central nervous system or the muscles<sup>120</sup>. This could lead to an rAAV vector overload and/or **unwanted transgene expression** in the liver, with potential toxicities (as discussed above).

## 2. LIVER GENE THERAPY STRATEGIES

Many liver-targeting gene therapy clinical trials are currently ongoing and can be grouped into 3 different situations, depicted in the **Figure 11** and **Table II**. Almost all of them use rAAVs.

#### a. USING THE LIVER AS A “BIOFACTORY”

The first strategy is the **use of the liver as a biofactory** in order to produce large amount of the desired protein in the bloodstream, as described above. These clinical trials are intended to treat either a disease in which a **protein in the plasma is lacking** (hemophilia A and B, C1-inhibitor deficiency) or a **lysosomal storage disease without neurological involvement** (type VI mucopolysaccharidosis, late-onset Pompe disease, Fabry disease, Gaucher disease). In the latter diseases, the missing lysosomal enzyme will be produced in the bloodstream in high amount and reuptaked by peripheral tissues via a mechanism known as cross-correction<sup>112</sup>. Indeed, lysosomal enzymes harbor a mannose-6-phosphate group which allow their trafficking to the lysosome, including from the plasma membrane<sup>121</sup>. Then, the

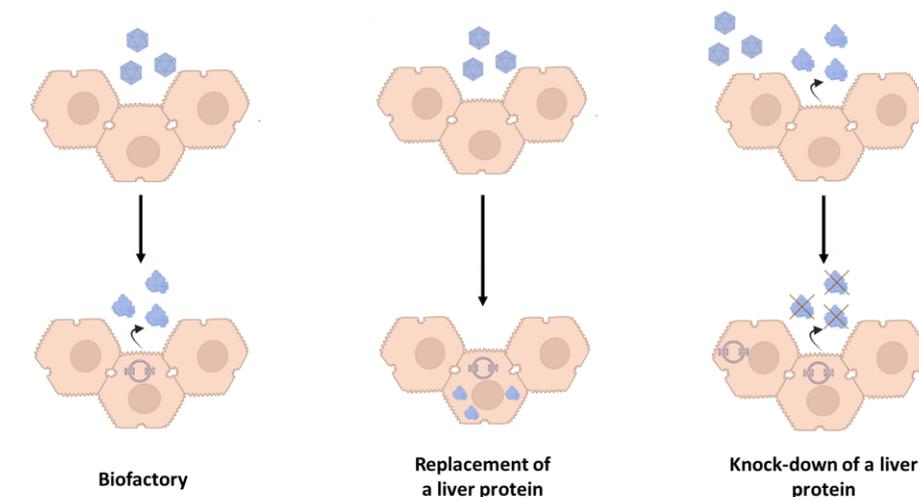
enzyme from the bloodstream can correct peripheral tissues, if present in sufficient amount. Since the crossing of the blood-brain barrier is not efficient enough, this strategy cannot be efficiently used for lysosomal storage diseases with neurological involvement<sup>121</sup>.

Strategy	Disease	Sponsor	Serotype ( <i>Transgene</i> )	Clinical trials
« Biofactory »	Hemophilia B	Sparks Therapeutics	<b>Spark100</b> ( <i>F9</i> -Padua)	Phase III (NCT03861273)
		Uniqure	<b>AAV5</b> ( <i>F9</i> -Padua)	Market approval
	Hemophilia A	BioMarin	<b>AAV5</b> ( <i>F8</i> -SQ)	Market approval
		Pfizer	<b>AAV6</b> ( <i>F8</i> -SQ)	Phase III (NCT04370054)
	Mucopolysaccharidosis type VI	Fondazione Telethon	<b>AAV8</b> ( <i>ARSB</i> )	Phase I/II (NCT03173521)
	Late onset Pompe disease (LOPD)	AskBio	<b>AAV8</b> ( <i>GAA</i> )	Phase I/II (NCT03533673)
		Sparks Therapeutics	? ( <i>GAA</i> )	Phase I/II (NCT04093349)
	Fabry disease	Sangamo Therapeutics	<b>AAV6</b> ( <i>GLA</i> )	Phase I/II (NCT04046224)
		Freeline Therapeutics	<b>AAV53</b> ( <i>GLA</i> )	Phase I/II (NCT04040049)
	Gaucher type I disease	Freeline Therapeutics	<b>AAV53</b> ( <i>GBA</i> )	Phase I/II (NCT05324943)
C1-inhibitor deficiency	BioMarin	<b>AAV5</b> ( <i>SERPING1</i> )	Phase I/II (NCT05121376)	
Replacement of a liver protein	Crigler-Najjar disease	Généthon	<b>AAV8</b> ( <i>UGT1A1</i> )	Phase I/II (NCT03466463)
	Urea cycle defects	Ultragenyx	<b>AAV8</b> ( <i>OTC</i> )	Phase III (NCT05345171)
		UCL	<b>LK03</b> ( <i>OTC</i> )	Phase I/II (NCT05092685)
	Glycogen storage disease Ia	Ultragenyx	<b>AAV8</b> ( <i>G6PC1</i> )	Phase III (NCT05139316)
	Phenylketonuria	BioMarin	<b>AAV5</b> ( <i>PAH</i> )	Phase I/II (NCT04480567)
		Homology Medicine	<b>AAVHSC15</b> ( <i>PAH</i> )	Phase I/II (NCT03952156)
	Porphyria	Uniqure	<b>AAV5</b> ( <i>OBGD</i> )	Phase I (NCT02082860)
	Methylmalonic acidemia	LogicBio	<b>LK03</b> ( <i>MMUT</i> integrative)	Phase I/II (NCT04581785)
	Wilson disease	Vivet Therapeutics	? ( <i>mini-ATP7B</i> )	Phase I/II (NCT04537377)
Ultragenyx		<b>AAV9</b> ( <i>ATP7B</i> )	Phase I/II (NCT04884815)	
Knock-down of a protein	TTR-amyloidosis	Intellia Therapeutics	<b>LNP</b> expressing CRISPR & <i>TTR</i> -targeting sgRNA	Phase I/II (NCT04601051)

**Table II. List of gene therapy clinical trials targeting liver**, available on clinicaltrials.gov on 1<sup>st</sup> June 2023. *F8-SQ* = *B domain-deleted factor VIII*.

Using this strategy, reaching only a **small number of hepatocytes** can be enough to obtain a therapeutic level of the desired protein, especially since even a small increase of plasma factor VIII or IX or of lysosomal enzyme amount is sufficient to **drastically improve the phenotype**<sup>81</sup>. The synthesis capacities of hepatocytes are very important<sup>111</sup>, therefore targeting too many hepatocytes can also lead to toxicities as recently demonstrated in a clinical trial for hemophilia B with one patient requiring anticoagulant therapy for factor IX levels above 200% of normal<sup>122</sup>. A relatively “low” dose (< 5 x 10<sup>12</sup>

vg/kg) can be sufficient to obtain a therapeutic effect and toxicity is likely to be low<sup>81</sup>. The liver is **usually not damaged** in these diseases with no proliferation or fibrosis. Importantly, transgene expression in the hepatocyte leads to **immune tolerance** toward the secreted protein, which is very important considering the high rate of antibodies developed against clotting factor concentrate in hemophilia patients (25%) and against the enzyme replacement therapy in Pompe disease (60%)<sup>123,124</sup>. When considering all these elements, rAAV appears as the vector of choice, using a capsid with liver tropism (AAV8, AAV5, AAV3-derived capsid such as AAVS3 or LK03).



**Figure 11. Different rAAV-based *in vivo* gene therapy strategies targeting the liver. Made with BioRender.**

Clinical trials for hemophilia A and B have shown the possibility to reach a clotting factor level above 20% of normal and a **full correction of the bleeding phenotype** in most of the patients<sup>75,125</sup>. Two gene therapy products **recently received market approval** for hemophilia B (etranacogene dezaparvovec, Hemgenix<sup>®</sup>) and hemophilia A (Valoctocogene Roxaparvovec, Roctavian<sup>®</sup>), based on published data showing a significant reduction of annualized bleeding rates and clotting factor concentrates use<sup>126,127</sup>. Of note, a high frequency of elevation of liver enzymes was reported (up to 85% of patients in the clinical trial using Roctavian), which was managed using transient immunosuppression<sup>127</sup>. In lysosomal storage diseases, clinical trials are still ongoing except for **type VI mucopolysaccharidosis**. In this trial, the highest dose resulted in an enzymatic activity in the plasma of 30% of control value and a modest clinical benefit<sup>128</sup>. For Pompe disease, preclinical studies evaluating the targeting of the liver in a mouse model using an rAAV demonstrated complete correction of the muscle phenotype, above the one obtained using the enzyme replacement therapy<sup>112,129</sup>. One of the advantage of targeting the liver compared to the enzyme replacement therapy is the continuous production of the enzyme by the liver, avoiding fluctuations in the plasma enzymatic activity level<sup>112,129</sup>.

#### b. REPLACEMENT OF A LIVER PROTEIN

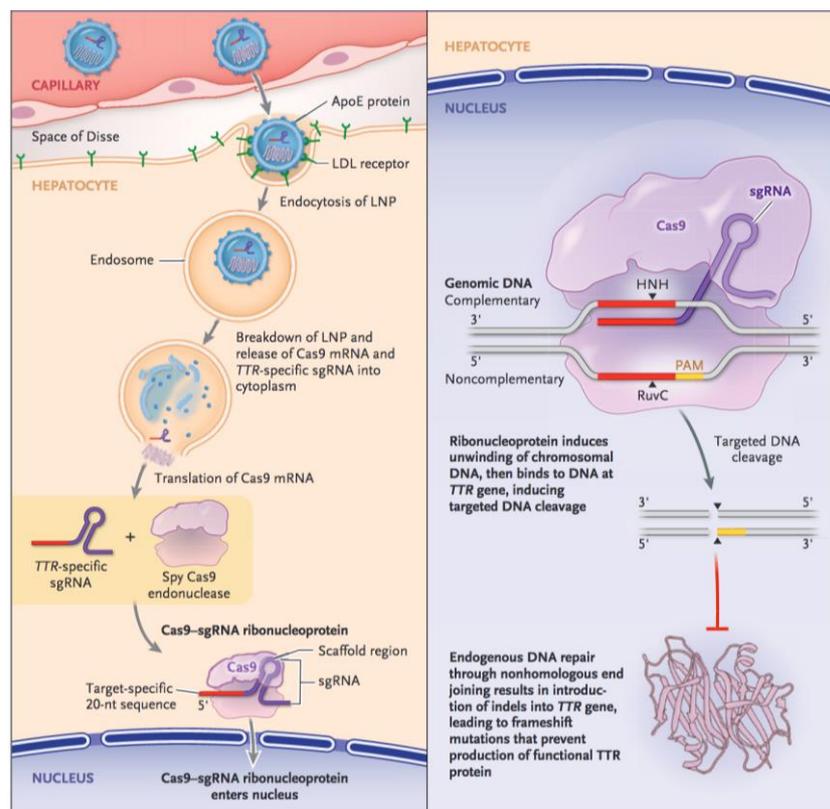
Another strategy is the **replacement of a missing protein**, usually an enzyme, in the liver. Most of the diseases in this group are either inborn errors of metabolism, impacting a specific metabolic pathway in the liver, or chronic fibrotic disorders. In inborn errors of metabolism, the liver is usually not proliferative and without significant fibrosis, and the restoration of a modest enzymatic activity (5 to 10% of normal) is usually sufficient to improve the phenotype of patients<sup>130,131</sup>. Accumulated metabolites, such as unconjugated bilirubin (Crigler-Najjar disease), ammonia (urea cycle defect) or phenylalanine (phenylketonuria) can be **transferred between hepatocytes**. Thus, it is not required to correct all hepatocytes to achieve a significant detoxification of the metabolites. In these diseases, the rAAV is therefore a vector of choice, at least in adult patients. Preclinical studies have shown the possibility to fully correct the phenotype of Crigler-Najjar disease and GSD Ia using rAAV expressing the missing enzyme in adult mice and rats<sup>132</sup> or in dogs<sup>133</sup>. Several clinical trials are currently ongoing in these diseases ([Table II](#)), with promising results<sup>1</sup>.

In contrast, in chronic fibrotic disorders such as progressive familial intrahepatic cholestasis (PFIC), Wilson disease or GSDIII, the liver exhibit chronic damages with **hepatocyte loss**, compensatory **proliferation** and progression toward **fibrotic liver disease**<sup>1</sup>. In these diseases, the accumulation of limit dextrin (GSDIII), of copper (Wilson disease) or of bile salts (PFIC3) is a **cell autonomous defect** and is restricted to the diseased cell. Therefore, it is **necessary to correct as many cells as possible**, since the correction of a cell will not correct neighboring cells. In this context, the use of rAAV gene therapy is more challenging and is discussed in further details below. To date, clinical trials have begun only for **Wilson disease** ([Table II](#))<sup>1</sup>. These trials are intended for patients with stable liver disease under copper chelation, in order to reduce as much as possible hepatocyte loss and proliferation, and thus to achieve a long-term correction with rAAVs<sup>1</sup>.

#### c. KNOCK-DOWN OF A LIVER PROTEIN

In some diseases, an improvement of the phenotype can be obtained by **knocking-down the expression of a particular protein** in hepatocytes. The use of **siRNA enveloped in lipid nanoparticles** has been evaluated in humans in 4 inherited diseases, such as primary hyperoxaluria type I (PH1), Transthyretin (TTR)-mediated amyloidosis, acute intermittent porphyria and ZZ-associated  $\alpha$ 1-antitrypsin deficiency, and currently have market approval in Europe for the first three<sup>1,134</sup>. The goal is either to reduce the **accumulation of a toxic mutated protein** (TTR,  $\alpha$ 1-antitrypsin ZZ) or to reduce the accumulation of the toxic metabolite by **suppressing the enzyme responsible of its synthesis** (PH1, acute intermittent porphyria). This siRNA-based approach is only transient and requires repeated administration of the treatment<sup>1</sup>.

An alternative gene therapy strategy is currently evaluated in patients with amyloidosis due to mutations in the *TTR* gene. In order to permanently reduce the production of the mutant protein, patients have received a lipid nanoparticle containing the mRNA of the Cas9 protein and a guide RNA directed against the endogenous *TTR* gene (depicted in [Figure 12](#))<sup>7</sup>. It allows a transient expression of the editing machinery in hepatocytes, but a permanent knock-down of the *TTR* gene. Preliminary data have shown a significant reduction of the plasma TTR concentration in patients, without toxicity after a very short follow-up<sup>7</sup>. A similar approach has also been used in mice with PH1<sup>135</sup>, type I tyrosinemia<sup>136</sup> or type I glutaric aciduria<sup>137</sup>, with promising results.



**Figure 12. Mechanism of action of the *TTR* gene-targeting CRISPR-Cas9 strategy used in the clinical trial for *TTR*-mediated amyloidosis, issued from *Gillmore et al.*<sup>7</sup>.**

### 3. SPECIFIC ISSUES ASSOCIATED WITH LIVER TARGETING

The main limitation of rAAVs for liver targeting relates to the **proliferation of hepatocytes** due to liver growth or chronic liver disease with subsequent vector dilution<sup>1</sup>. The generation of high NABs titers following the initial rAAV treatment prevents rAAV readministration and therefore the use of rAAV for the **treatment of children and of patients with chronic liver damage**<sup>1</sup>, two patient populations with high unmet need.

a. LIVER GROWTH

The **notion of vector dilution** with liver growth has been demonstrated in rodents, with the inevitable loss of rAAV vector 2 to 3 weeks after neonatal rAAV injection<sup>138</sup>. Nonetheless, human liver growth is not equivalent to the one of mice, as summarized in [Table III](#)<sup>138,139</sup>. In humans, most of the liver growth has occurred by 6 years of age, whereas liver growth in mice is both more intense and logically faster. Furthermore, these data only take into account the weight of the liver, and cannot distinguish **hepatocyte proliferation from hepatocyte hypertrophy**. Moreover, another difference is the timing of **liver embryogenesis** between the two species ([Table III](#)), since part of liver development occurs in the post-natal period in mice<sup>140,141</sup>. Thus, it is challenging to extrapolate to humans the results of an rAAV injection performed in mice at birth (P0), especially since the latter cannot be realistically performed in humans considering the time required to obtain a diagnosis and to prepare the rAAV injection.

	Human	Mouse (C57BL/6)
<b>Liver weight growth between the neonatal period and adulthood</b>		
Total expansion	X 16 <u>including</u> : X 4 before 2 years of age X 4 between 2 and 18 years	X 32 <u>including</u> : X 12 before 4 weeks of age X 2 between 4 and 6 weeks
Age at liver doubling	4-month-old 16-month-old 6-years-old During puberty	During the 1 <sup>st</sup> week of life Day 7 Day 15 Day 30 6 weeks of age
Age at “adult” liver weight	At adulthood	At 6 weeks
<b>Embryonic development</b>		
Liver bud	4 – 8 weeks	9.5 – 14.5 days (E9.5-E14.5)
Hepatocyte and cholangiocyte differentiation	8 – 30 weeks	13 .5 – 18.5 days (E13.5-E18.5)
Ductal plate formation	12 weeks	15.5 days (E15.5)
End of the biliary differentiation	30 weeks	Neonatal period
Liver vascular development	Prenatal	Postnatal

**Table III. Comparison of liver growth and embryonic development between human and C57BL/6 mice.** From *Cunningham et al.*<sup>138</sup> and *Coppoletta et al.*<sup>139</sup> for liver growth and from *Raynaud et al.*<sup>140</sup> and *Gordillo et al.*<sup>141</sup> for embryonic development.

## b. CHRONIC LIVER DAMAGE

As mentioned earlier, the replacement of a missing protein in a **background of chronic liver damage**, with **hepatocyte proliferation** and **progressive fibrosis**, is challenging<sup>1,142</sup>. Many of these pathologies, such as PFICs, will lead to liver transplantation early on during childhood, whereas for others there is an available treatment or regimen able to slow down or stop the progression of the disease (Wilson disease, type I tyrosinemia)<sup>1,142</sup>. Considering gene therapy trials, specific issues arise such as the **necessity of an early treatment** (before death or liver transplantation), the management of **hepatocyte proliferation** (with possible rAAV dilution) and the **high inherent risk of HCC** present in many of these diseases<sup>1,142,143</sup>. Glycogen storage disease type III (GSDIII) also belongs to this group of diseases, although the degree of fibrosis observed both in patients and in animal models is much less pronounced in GSDIII compared other fibrotic diseases. Thus, it might be possible to use rAAVs in this setting to treat GSDIII.

Several preclinical studies have demonstrated the ability of an rAAV vector to correct the phenotype of murine models of Wilson disease<sup>77,144</sup> or PFIC type 3<sup>145,146</sup>, although some of these models are less severe than human patients. Studies in animal models have suggested that the correction of the phenotype requires **high doses of rAAV vectors** (usually  $> 10^{13}$  vg/kg)<sup>77,145,146</sup> and that the **fibrosis stage reduces liver transduction by rAAVs**<sup>147</sup>. Therefore, rAAVs treatment should be given early in life, before significant fibrosis developed, but this will require to treat a growing liver. Furthermore, the correction of an insufficient number of cells could lead to a persisting low grade inflammation and hepatocyte proliferation, with a risk of subsequent **loss of correction by vector dilution** and recurrence of the disease<sup>146,148</sup>.

Importantly, many of these diseases (type I tyrosinemia, PFIC2, PFIC3) display a **very high risk of liver tumors** with two specific issues, one about the safety of the vector, the second about the ability of the vector to prevent the occurrence of the tumors<sup>142</sup>. First, although rAAVs display a good safety profile, few data are available in a background of chronic liver damage, at increased risk of HCC because of the underlying disease itself<sup>149</sup>. In case of HCC, how to distinguish between a tumor induced by rAAV and a tumor due to the cirrhotic background? Integration analyses within the tumors could help to distinguish between these two hypotheses. Second, in order to reduce as much as possible the occurrence of tumors, the broadest transduction of hepatocytes must be achieved, since HCC could arise from focal areas of untransduced hepatocytes.

## c. INTEGRATIVE STRATEGY AND GENE EDITION

Only one clinical trial has evaluated the injection of an liver targeting-rAAV in children, for patients with type VI mucopolysaccharidosis due to arylsulfatase B deficiency<sup>128</sup>. Several patients received the rAAV between the age of 5 to 10 years of age, and displayed a **stable arylsulfatase B**

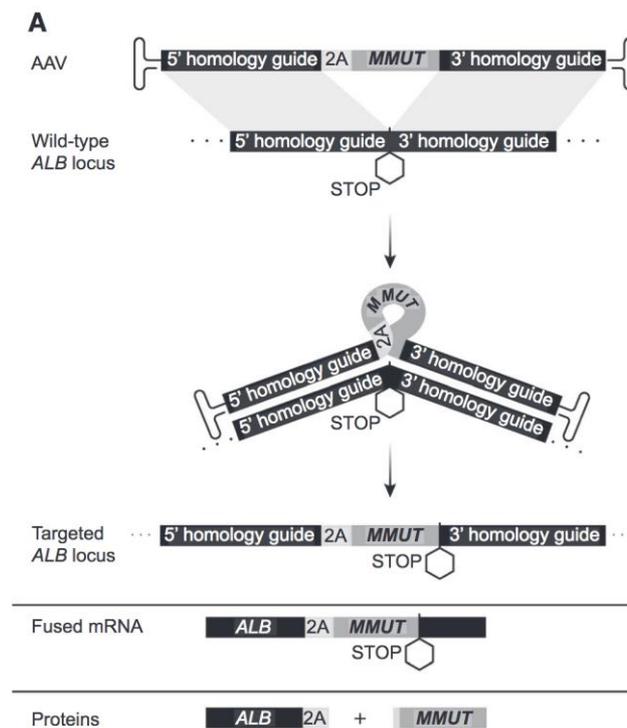
**enzymatic activity** (25-30% of normal) for a follow-up of at least 2 years<sup>128</sup>. It seems therefore reasonable to inject patients older than 6-10 years of age, especially since most of the liver growth has already occurred by this age ([Table III](#)). Nonetheless, several diseases would require an earlier treatment in order to prevent **early liver transplantation or death** (PFIC), to prevent **potential sequelae** (urea cycle defect, methylmalonic acidemia) and/or because of **poor quality of life** (Crigler-Najjar syndrome).

In these patients, two strategies could be considered: the **readministration of rAAVs** or the use of an **integrative or gene editing strategy**<sup>1</sup>. The first one has already been discussed above, and consists in preventing NAb production or removing NAb to allow readministration of the rAAV. Sequential rAAV injections could allow the treatment of children from an early age and in case of vector dilution due to high hepatocyte proliferation in patients with chronic liver disease. Several strategies have been evaluated, only in preclinical models so far, and might allow a second injection of a therapeutic vector. One recent publication using co-injection of rapamycin-loaded nanoparticle along with rAAVs reports promising results in a mouse model of PFIC3, with successful readministration of rAAVs encoding the missing transporter, long-term correction of the cholestatic phenotype and prevention of liver tumors<sup>148</sup>.

The second strategy relies on the **permanent modification of the genome of hepatocytes**, using an integrative transgene, an integrative vector and/or a CRISPR-Cas9 gene editing strategy<sup>1</sup>. Few interesting examples are discussed below.

An elegant strategy has been developed for **methylmalonic acidemia** (MMA), a severe inborn error of metabolism of protein intoxication with manifestations in the first few months of life. The missing gene (*MMUT*) has been **integrated in the albumin locus**, which is a locus highly expressed in hepatocyte<sup>150</sup>. *Mmut*<sup>-/-</sup> mice received in the neonatal period an rAAV8 encoding for the *MMUT* gene flanked by two regions of homology allowing its integration in the 3' part of the albumin locus by homologous recombination ([Figure 13](#)). Once integrated, the *MMUT* gene is located upstream of the stop codon of the albumin gene and allows the production of a **fusion mRNA** containing both the *Albumin* and the *MMUT* coding sequences separated by a **P2A sequence**. Upon mRNA translation, the ribosome will stop at the P2A sequence and re-initiate the translation downstream of the P2A sequence (*ribosome skipping*) in the *MMUT* coding sequence. This allows the production of both the albumin (tagged with a P2A sequence) and the missing MMUT protein. The level of P2A-labeled albumin in the plasma can be used to indirectly assess the level of integration. The integration is not efficient, with less than 1% of integration 1 month after treatment but 10 to 20% of integration 1 year after the treatment in mice, resulting in a full correction of the metabolic phenotype<sup>150</sup>. One hypothesis is the **selective advantage of corrected cells**, combined with the **high proliferation rate** associated to the neonatal injection (as well as to the underlying disease)<sup>150</sup>. Indeed, homologous recombination

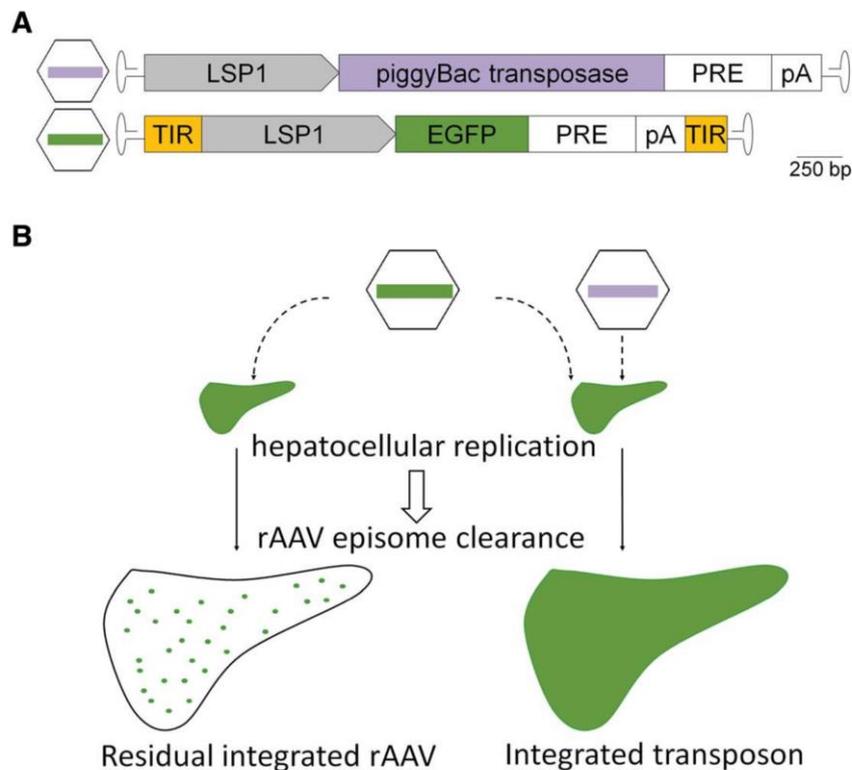
allowing the integration of the transgene is more efficient in a proliferative liver, therefore the timing of injection in the neonatal period will favor the integration of the *MMUT* gene. Because the cells are permanently corrected, their number will progressively increase over time because of persisting proliferation, and corrected cells are positively selected unlike uncorrected cells. This strategy is currently evaluated in a clinical trial led by LogicBio intended for young MMA patients (NCT04581785).



**Figure 13. Mechanism of action of the nuclease-free integrative transgene used for MMA, from Chandler et al.<sup>150</sup>.**

Another interesting strategy is the **successful use of a transposase** in two different mouse models of **urea cycle disorders**<sup>151</sup> and in a mouse model of **PFIC3**<sup>147</sup>. This enzyme, originating from a transposon, is able to perform the excision of DNA sequence flanked by terminal inverted repeats (TIR) and allows the transposon to “move” in the genome through cycles of excision-insertion. The transposase can be encapsidated in a rAAV and **does not induce mutagenesis at the time of excision** (known as “seamless excision”)<sup>147</sup>. Its application in gene therapy is depicted in the **Figure 14**. Two rAAV vectors are required, one encoding for the transposase under the control of a liver promoter and the second one containing the gene of interest within a classical expression cassette (with a promoter and poly-A signal) **flanked by TIR sequences**. Upon the synthesis of the transposase, the expression cassette is integrated within the host genome. In both aforementioned studies, the injection was performed in neonatal mice, at a time of maximal liver growth. Corrected cells will divide and spread

in the liver, especially if a selective advantage is present, whereas both rAAV vectors will be progressively lost because of vector dilution<sup>147,151</sup>. The loss of the transposase-containing vector ensures that the transposase is only transiently expressed. One alternative would be the use of **lipid nanoparticles** to deliver both constructs, avoiding the potential long-term persistence of the transposase. This strategy allowed the restoration of an enzymatic activity around 30% of normal in both urea cycle disorder mouse models, with prevention of premature death<sup>151</sup>. Similarly, restoration of biliary secretion and prevention of liver disease and occurrence of tumors was demonstrated in a mouse model of PFIC3<sup>147</sup>. A genotoxicity study performed in the first study showed integration **mostly within genes** and close to the transcription starting site in 15% of cases<sup>151</sup>. The study in PFIC3 did only evaluate the occurrence of tumors, reduced from 100% of untreated-mice to 7% in treated mice<sup>147</sup>.



**Figure 14. Gene therapy strategy using the *Piggybac* transposase system**, from Cunningham *et al.*<sup>151</sup>. *LSP1*= liver specific promoter 1; *TIR*= terminal inverted repeats.

Another recently published approach relies on the **optimization of a lentivirus** able to target efficiently the liver<sup>9,10</sup>. The integration of the expression cassette in hepatocytes allows a persistent expression of the transgene, even in case of cell proliferation<sup>8,9</sup>. Other advantages of this strategy are the **low seroprevalence** of NAbS against lentiviruses (i.e. against human immunodeficiency virus (HIV) or the VSV-G glycoprotein used for pseudo-serotyping) in the general population<sup>9,10</sup>. Moreover, this vector can deliver **large expression cassette** up to 10 kb<sup>2,8</sup>. Earliest generation of optimized lentiviruses

could correct the phenotype of mouse models of hemophilia, although 3 to 5 times more vector was present in **Kupffer cells** and dendritic cells than in hepatocytes, in line with a fast clearance of the vector<sup>9,10</sup>. Beside the reduced efficacy, the uptake of the vector by immune cells induces cytokine secretion<sup>9</sup>. Further optimizations, by increasing the expression of the **phagocytosis-inhibitor CD47** at the surface of the vector, improved the efficacy of hepatocytes transduction while reducing the transduction and activation of immune cells<sup>9,10</sup>. The safety profile of these vectors seems good in both mice and non-human primates, without insertional mutagenesis<sup>9,10</sup>.

Finally, several proof-of-concepts have been published in preclinical models using the **CRISPR-Cas9** technology. **Knock-down of a liver-expressed protein**, such as *PCSK9* or *AGPTL3* for primary hypercholesterolemia<sup>152–154</sup> or *HAO1* for PH1<sup>135</sup> relies on indels generation by non-homologous end joining or base editing. Fewer studies have shown the possibility to **restore the expression of a missing protein**, such as in type I tyrosinemia through homologous recombination<sup>155</sup> or base editing<sup>156</sup> or in phenylketonuria using a strategy of *prime editing*<sup>157</sup>. The main limitation is the **low efficiency of homologous recombination** and therefore the low number of corrected cells. Nonetheless, its use in a disease combining a **proliferative liver** (*a fortiori* in children) which increases the rate of homologous recombination, a potential **selective advantage of corrected cells** and the possibility to improve the phenotype with **only a fraction of the normal enzymatic activity** could confer a significant benefit. Type I tyrosinemia is a good example of such a disease. A recent study showed the possibility to correct the phenotype in rabbit with type I tyrosinemia after neonatal injection of an rAAV expressing the gene editing machinery<sup>155</sup>. Interestingly, only 3 to 6% of corrected cells were present in the liver at last evaluation, but were sufficient to fully correct the phenotype<sup>155</sup>.

#### **IV. MUSCLE-TARGETED GENE THERAPY BASED ON RAAV**

Muscles are frequently targeted in gene therapy trials based on rAAV, although less than the liver, the retina and the central nervous system in term of number of trials<sup>6</sup>. Muscle is the **largest organ within the body** and includes **striated skeletal muscles, heart, diaphragm** as well as **smooth muscles**. This organ is difficult to target, mostly because of its extension without any possibility of local injection and its peculiar cellular structure with very long multinucleated muscle cells.

##### **1. UNDERLYING MUSCLE DISEASES**

Muscle-targeted gene therapy trials include two main groups of diseases, each having specificities: **muscular dystrophies** and **metabolic and congenital myopathies**. Ongoing clinical trials are shown in [Table IV](#).

Group	Disease	Sponsor	Serotype ( <i>Transgene</i> )	Clinical trial
Muscular dystrophy	Duchenne muscular dystrophy (DMD)	Pfizer	<b>AAV9</b> ( $\mu$ Dys)	Phase III (NCT05689164)
		Sarepta	<b>AAVrh74</b> ( $\mu$ Dys)	Phase III (NCT05881408)
		Solid Biosciences	<b>AAV9</b> ( $\mu$ Dys)	Phase I/II (NCT03368742)
		Généthon	<b>AAV8</b> ( $\mu$ Dys)	Phase I/II (EudraCT 2020-002093-27)
		Regenxbio	<b>AAV8</b> ( $\mu$ Dys)	Phase I/II (NCT05693142)
	DMD with exon 2 duplication	Audentes	<b>AAV9</b> (exon skipping)	Phase I/II (NCT04240314)
	LGMD 2I/R9	AskBio	<b>AAV9</b> ( <i>FKRP</i> )	Phase I/II (NCT05230459)
	LGMD 2B/R2	Généthon	<b>AAV9</b> ( <i>FKRP</i> )	Phase I/II (NCT05224505)
	LGMD 2E/R4	Sarepta	<b>AAVrh74</b> ( <i>DYSF</i> )	Phase I/II (NCT05906251)
LGMD 2E/R4	Sarepta	<b>AAVrh74</b> ( <i>SGCB</i> )	Phase I/II (NCT03652259)	
Congenital myopathy	Myotubular myopathy	Astellas	<b>AAV8</b> ( <i>MTM1</i> )	Phase I/II (NCT03199469)
Metabolic myopathy & cardiomyopathy	Danon disease	Rocket Pharmaceuticals	<b>AAV9</b> ( <i>LAMP2B</i> )	Phase I/II (NCT03882437)
	Infantile-onset Pompe disease	Chinese PLA General Hospital	<b>AAV9</b> ( <i>GAA</i> )	Phase I/II (NCT05567627)
	Late-onset Pompe disease	Astellas	<b>AAV8</b> ( <i>GAA</i> )	Phase I/II (NCT04174105)
Isolated cardiomyopathy	Friedreich ataxia	Lexeo Therapeutics	<b>AAVrh10</b> ( <i>FXN</i> )	Phase I/II (NCT05445323)
	Fabry with isolated heart involvement	4D Molecular Therapeutics	? ( <i>GLA</i> )	Phase I/II (NCT05629559)
	Hypertrophic cardiomyopathy	Tenaya Therapeutics	<b>AAV9</b> ( <i>MYBPC3</i> )	Phase I/II (NCT05836259)
	Arythmogenic cardiomyopathy	Rocket Pharmaceuticals	<b>AAVrh74</b> ( <i>PKP2A</i> )	Phase I/II (NCT05885412)

**Table IV. List of gene therapy clinical trials targeting inherited muscle and heart diseases**, available on clinicaltrials.gov on 1<sup>st</sup> June 2023. Early trials using intramuscular injections were excluded. *DMD*= *Duchenne muscular dystrophy*; *LGMD* = *Limb-girdle muscular dystrophy*;  $\mu$ Dys = *microdystrophin*.

#### a. MUSCULAR DYSTROPHIES

This group of diseases include **DMD** and the late onset **Limb-girdle muscular dystrophies** (LGMD). The molecular anomaly usually involves a structural protein, such as proteins of the dystrophin-associated protein complex, linking the cytoskeleton to the extracellular space through numerous membrane-anchored proteins. These diseases display a dystrophic phenotype, with **destruction of muscle fibers**, intense **inflammation** and **necrosis-regeneration** phenomenon<sup>158</sup>. It

leads to muscle tissue degeneration with progressive **fibrosis**, myocytes loss and **replacement by adipose tissue**<sup>158</sup>. In line with this, it is usually difficult to obtain a reversion of the muscle disease once the degeneration of muscle fibers is already advanced and early treatment is therefore required<sup>159</sup>. Moreover, the high inflammatory environment observed in these dystrophies could prime the immune system and worsen potential immune responses directed toward the rAAV or the transgene, as described in mouse models of sarcoglycanopathies<sup>119</sup>.

Five clinical trials using rAAV expressing a truncated version of the dystrophin transgene are currently ongoing for DMD (reviewed in *Elangkovan et al.*<sup>160</sup>), one trial is ongoing using an exon-skipping strategy and one trial using the CRISPR-Cas9 technology<sup>161</sup>, following promising preclinical results in mice<sup>162</sup>, rats<sup>163</sup> and dogs<sup>164</sup> models of DMD. A high dose of rAAV is required (between  $5 \times 10^{13}$  and  $3 \times 10^{14}$  vg/kg), as well as a capsid with muscle tropism (AAV9, AAV8, AAVrh74) and a muscle promoter (MHCK7, CK8, SpC5.12)<sup>160</sup>. Beyond DMD, many preclinical studies have been published in various mouse models of LGMD (FKRP, CAPN3, LAMA2, BVES) with positive results<sup>159,165–167</sup> and several clinical trials are currently starting for patients with LGMD ([Table IV](#)).

#### b. CONGENITAL OR METABOLIC MYOPATHIES

This group of diseases includes **glycogen storage diseases with muscle involvement** (GSDIII, Pompe disease, Danon disease) and **congenital myopathies**, such as myotubular myopathy (MTM). The molecular anomaly usually impacts a gene required to provide energy to the muscle fiber, such as an enzyme of glycogen metabolism, or a protein regulating muscle contraction, such as the phosphatase MTM1. There is a **blockade of the muscle contraction** and the occurrence of a myopathy that is present since birth (MTM1) or progressively following glycogen accumulation (GSDs). But **no or few signs of muscular dystrophies are present**, such as necrosis-regeneration, inflammation or fibrous degeneration of the muscle<sup>168,169</sup>. Then, in contrast to muscular dystrophies, it should be possible to get **some degree of reversal of the muscle impairment**, even in a relative advanced stage. Preclinical studies in animal models of Pompe disease<sup>170</sup>, GSDIII<sup>69</sup> or MTM<sup>171</sup> have shown the possibility to restore muscle strength in animals already significantly impaired. Interim results for the clinical trial involving patients with MTM (NCT03199469) also suggest so, with patients able to acquire walking and to be weaned off mechanical ventilation<sup>172</sup>. As described for the hepatic inborn errors of metabolism, the restoration of even a small fraction of the normal enzymatic activity can have beneficial effects in metabolic myopathies, as shown for GSDIII<sup>69</sup>. In the GSDIII mouse model, a restoration of 10% of the normal GDE expression in heart and 2% in quadriceps was sufficient to achieve a major reduction of glycogen > 80%<sup>69</sup>. Some clinical trials are ongoing in Pompe disease and Danon disease, without published results to date ([Table IV](#)), as well as for metabolic cardiomyopathy in Fabry disease and Friedreich ataxia.

## 2. ISSUES ASSOCIATED WITH MUSCLE TARGETING USING RAAV

Gene therapy using rAAVs faces two challenges to target skeletal muscles: the **poor muscle targeting** achieved with most rAAV serotypes, requiring in the use of **high doses of vector**, and the **resulting toxicities** due to liver, but also heart, overloading.

### a. POOR MUSCLE TARGETING WITH CURRENT RAAV

Natural rAAVs have a **poor tropism for skeletal muscles** compared to the liver and spleen for instance, including serotypes commonly referred to as “muscle targeting” (AAV9, AAVrh74). Indeed, a recent study reporting the autopsies of two children with SMA treated with Zolgensma® at a high dose ( $1.1 \times 10^{14}$  vg/kg) and who died despite the treatment confirmed the low muscle targeting ability of the AAV9 capsid<sup>120</sup>. Vector genome copy numbers (VGCN) were around **1 vector genome per diploid genome (vg/dg)** for skeletal muscles and slightly higher for diaphragm and heart (2-3 vg/dg) while the **liver exhibited more than 100 vg/dg**<sup>120</sup>.

To reduce the vector dose used, **design of new rAAV serotypes** having better muscle targeting abilities is an interesting solution. Several methods have been described, such as natural discovery, rational design, random selection and/or a combination of several of them<sup>31</sup>, as discussed above. Two recent studies report the identification of chimeric capsids close to AAV9, harboring a **peptide insertion in the VP3 protein** (AAVMYO and MyoAAV) allowing enhanced muscle targeting, while reducing liver targeting<sup>26,27</sup>. A similar capsid was generated in our team, by inserting a RGD-containing peptide (allowing integrin binding) in the VP3 protein of an AAV9-AAVrh74 chimeric capsid<sup>173</sup>, allowing strong muscle targeting and liver detargeting. This capsid has been used in the work presented in this manuscript and named AAV-MT.

Nonetheless, the **translation of these vectors**, usually optimized in mice, to other species such as humans is still an open question. Their evaluation in non-human primates could give some information about their potential usefulness for further translation in human, as performed for the aforementioned study reporting MyoAAV<sup>27</sup>. Indeed, **discrepancies** in term of central nervous system transduction have been reported between evaluations in mice and non-human primates for some chimeric rAAV capsids such as PHP.B<sup>44</sup>.

### b. LIVER AND HEART TOXICITIES AFTER HIGH DOSE OF RAAV

Most of the severe toxicities of rAAV gene therapy were reported in **clinical trials intended for neuromuscular diseases**<sup>1</sup>, likely because of the **high doses of rAAV** required. The hepatic toxicity has already been discussed above, seems correlated to the dose received by the liver and could result from a possible direct toxicity of rAAVs. One other mechanism of toxicity could be the **inappropriate expression of the transgene** in the liver, even if a “specific” promoter is used. Indeed, considering the

high level of liver transduction at these doses of vector, promoter leakage is possible and could induce toxicities. This has been demonstrated in a mouse model of **Friedreich ataxia**, in which the missing gene (Frataxin/*FXN*) was restored in the heart using rAAV vectors with an ubiquitous promoter<sup>174</sup>. At high dose, unwanted Frataxin overexpression in the liver resulted in a dose-dependent toxicity due to its mitochondrial action<sup>174</sup>. It recalls the importance to **control the expression dosing** of the transgene (especially if the therapeutic index is small) not only in the targeted organ but also in off-target organs, such as the liver. Considering this, the generation of a chimeric capsid with both enhanced muscle targeting and strong liver detargeting (as discussed above) could be a valuable tool to reduce the potential liver toxicity by reducing the total vector dose received by the liver<sup>26</sup>. This could also be interesting for **muscle diseases with some degree of preexisting liver disease**, such as MTM and GSDIII, as there is a significant risk of severe liver toxicities associated with liver overload at high doses<sup>87</sup>.

Another issue is the **potential cardiotoxicity** following rAAV injection, especially in patients with preexisting heart disease. The **heart is often better transduced** than peripheral muscles<sup>26,120</sup>. A high dose of rAAV used to target both muscles and heart can result in the **overload of the heart** with potential toxicity and/or immune responses<sup>83,161</sup>. First, severe cardiotoxicity occurred in clinical trials for DMD, with several cases of myocarditis including two deaths reported in multiple clinical trials<sup>83,161</sup>. Although the mechanism is not well understood, data gathered so far suggest the occurrence of innate and/or adaptive immune response to the vector. Beyond the potential immune response, overexpression of a transgene in the heart could also induce some toxicity. As in the liver, the overexpression of the Frataxin protein in the heart of a mouse model of Friedreich ataxia led to **severe cardiotoxicity** due to the mitochondrial action of the protein<sup>175</sup>. This mechanism could be important in GSDIII, considering the central role of glycogen metabolism in heart and the potential toxicities in case of sudden and/or profound depletion of glycogen in this organ.

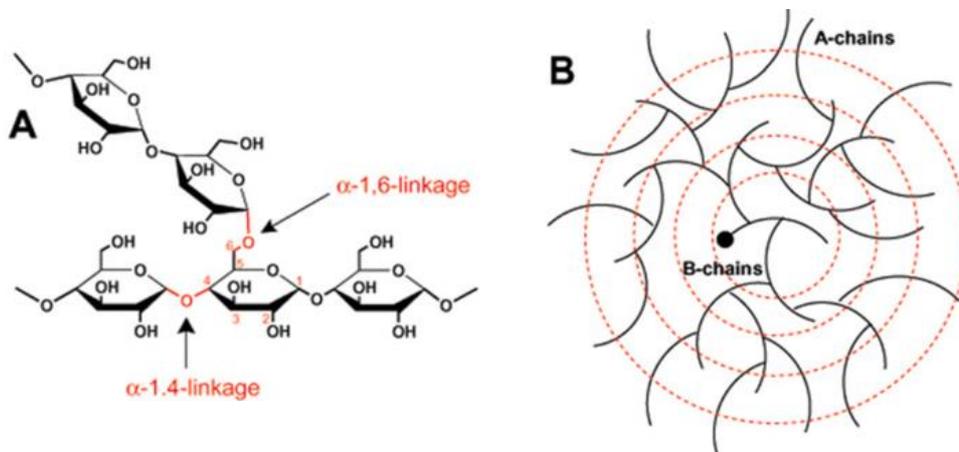
## **V. GLYCOGEN METABOLISM AND GLYCOGEN STORAGE DISEASE TYPE III**

Glycogen storage disease type III (GSDIII) is a rare inborn error of glycogen metabolism of autosomal recessive inheritance due to **pathogenic variants in the *AGL* gene** encoding the **glycogen debranching enzyme** (GDE)<sup>176,177</sup>. It results in accumulation of an abnormal glycogen known as **limit dextrin** in the **liver**, the **skeletal muscles** and the **heart** of patients<sup>178</sup>.

### **1. GLYCOGEN METABOLISM**

Glycogen is a **branch polymer of glucose** allowing the osmotically neutral storage of glucose in anticipation of fasting<sup>179</sup>. All species from bacteria to humans, except plants, have glycogen

metabolism which is therefore extremely conserved throughout evolution<sup>179</sup>. Glycogen plays a major role in the physiology of liver, of heart and skeletal muscles, and also of the brain<sup>180</sup>.

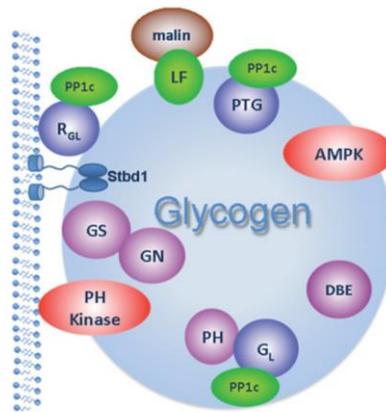


**Figure 15. Structure of glycosidic bonds (A) and overall structure of glycogen (B).** The black dot represents glycogenin. From Roach *et al.*<sup>179</sup>.

#### a. GLYCOGEN STRUCTURE

In glycogen, glucose molecules are linked through **linear bonds** («  $\alpha 1,4$  » bonds) and **branching points** («  $\alpha 1,6$  » bonds), resulting in a structure resembling the one of a tree (Figure 15)<sup>179</sup>. These branching points are required for **proper solubilization** of glycogen: in absence of the branching enzyme (glycogen storage disease type IV), the glycogen is only linear (“amylopectin-like”), insoluble and strongly cytotoxic<sup>178,179</sup>. The branching pattern of the glycogen is not exactly the same between all the glycogen molecules : therefore, **glycogen does not have a unique structure**, but rather is a mixture of glucose polymers with variable number of residues, of branches and *in fine* of molecular weight<sup>179</sup>. In the current accepted model, glycogen contains inner chains with one or two branching points (B-chains) and outer chains without branching points (A-chains) with a mean length of 13 residues per chain (Figure 15)<sup>179</sup>. In this model, the maximal number of residues would be 55,000 corresponding to a total molecular weight of  $10^7$  Da (10 MDa), equivalent to the one of a ribosome<sup>179</sup>. The glycogen molecule is centered by a protein known as **glycogenin** (Gyg), required to initiate the synthesis of glycogen (see below)<sup>181</sup>.

Importantly, glycogen is not isolated in the cytosol, but permanently associated with multiple proteins involved in its metabolism (Figure 16), such as **enzymes involved in the synthesis** (glycogen synthase) or **breakdown** (GDE, glycogen phosphorylase) of glycogen and **regulatory proteins** (phosphorylase kinase, phosphatase PP1c, AMPK, Laforin)<sup>179</sup>. The protein **STBD1** allows its anchorage to the membranes and plays a role in the glycopagy pathway (see below)<sup>179</sup>.



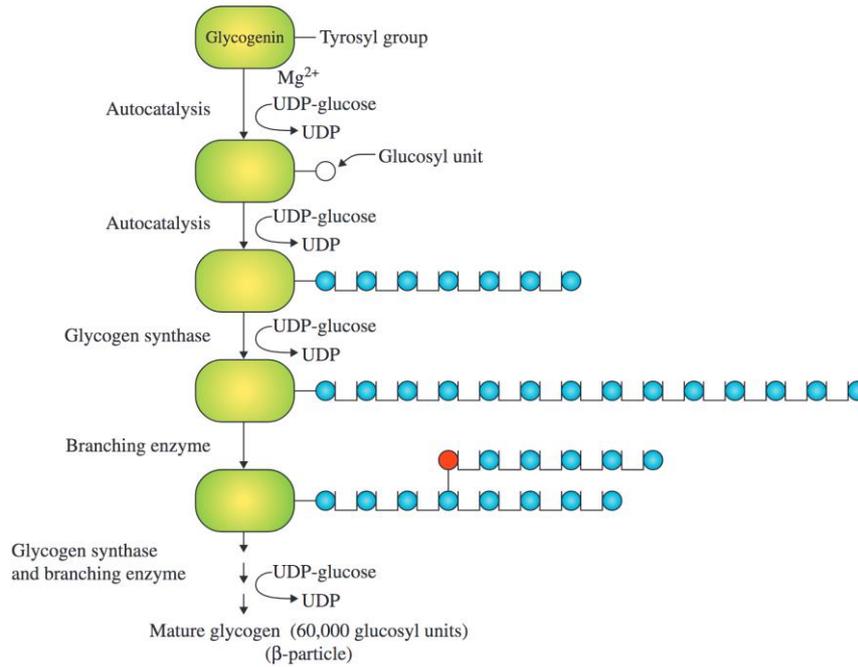
**Figure 16. Structure of a glycogen particle** with associated proteins involved in its metabolism (in purple), in its regulation (kinases in red; phosphatase in green) or allowing targeting to the glycogen (in blue). From *Roach et al.*<sup>179</sup>.

#### b. GLYCOGEN SYNTHESIS

The synthesis of glycogen from glucose requires the use of an activated form of glucose, known as **UDP-glucose**<sup>180</sup>. The glucose from the bloodstream is picked up by the cell through specific glucose transporters and phosphorylated in **glucose-6-phosphate** (G6P) by the first enzyme of the glycolysis pathway (hexokinase ou glucokinase). Following its isomerization in **glucose-1-phosphate** (G1P), it can be combined to uridine triphosphate by the UGP enzyme (UDP-glucose pyrophosphorylase) generating **UDP-glucose**, a compound reactive enough to be incorporate in the synthesis of glycogen<sup>180</sup>.

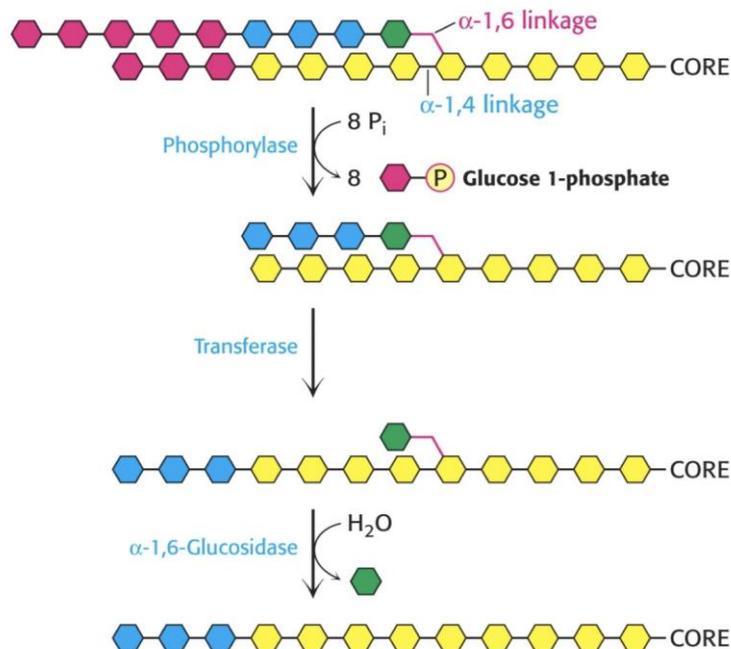
The enzyme enabling glycogen synthesis of linear bonds is **glycogen synthase** (Gys), but Gys acts only on preexisting linear chains and cannot begin the synthesis *de novo*<sup>180</sup>. The initiation of the synthesis of glycogen relies on the enzyme **glycogenin** (Gyg)<sup>178,180</sup>. Gyg is able to glycosylate itself on a conserved tyrosine residue and then to elongate the nascent chain up to 10 glucose residues<sup>180,181</sup>. Two isoforms of Gyg exist: Gyg1 in skeletal muscles and heart and Gyg2 in the liver<sup>178</sup>. Once the nascent chain is long enough, Gys takes over and continues the synthesis of linear chains (**Figure 17**)<sup>180</sup>. Two isoforms of Gys exist, as for Gyg: Gys1 in skeletal muscles, heart and brain and Gys2 in the liver<sup>178</sup>.

Then, a **branching enzyme** (GBE1) creates a **branching point** by displacing a series of 7 glucose residues from the nascent chain on position 6 of a glucose residue (**Figure 15**) located elsewhere in the molecule, generating an  $\alpha$ 1,6 bond (**Figure 17**)<sup>180</sup>. This branching point must be located at least 4 residues away from a previous one. Unlike the debranching reaction, the branching reaction only requires one single enzymatic activity (glucosyltransferase)<sup>180</sup>. The branching enzyme first transfers 7 glucose residues from the glycogen to a conserved residue of its catalytic site, move to the desired location and catalyzes the inverse reaction, thus creating the branching point<sup>180</sup>.



**Figure 17. Glycogen synthesis**, from *Bhagavan et al.*<sup>181</sup>.

The regulation of glycogen synthesis relies mostly on the **glycogen synthase kinase 3 $\beta$**  (GSK3 $\beta$ ). Under fed conditions, insulin inactivates GSK3 $\beta$ , enabling the dephosphorylation of Gys and therefore the synthesis of glycogen<sup>180</sup>. A second layer of regulation exists: the abundance of G6P (and therefore G1P) allosterically activates Gys and results in glycogen synthesis in fed conditions<sup>180</sup>.



**Figure 18. Glycogen breakdown**, from *Biochemistry, Stryer et al.*, 6<sup>th</sup> edition<sup>180</sup>.

### c. GLYCOGEN BREAKDOWN

The breakdown of glycogen requires **two distinct enzymes**: the **phosphorylase** for  $\alpha$ 1,4 linear bonds and the **glycogen debranching enzyme** (GDE) for  $\alpha$ 1,6 branched bonds (**Figure 18**)<sup>180</sup>.

The **phosphorylase** uses inorganic phosphate (Pi) to cleave  $\alpha$ 1,4 bonds and does not require any source of energy. It generates G1P molecules, which can fuel the glycolysis and produce energy after isomerization in G6P<sup>180</sup>. Different isoforms of phosphorylase are present in **muscles** (PygM), **liver** (PygL) and **brain** (PygB)<sup>178</sup>. Because these three enzymes have different roles in physiology, they are regulated in different ways (see below).

Once the phosphorylase encounters a branching point, it is unable to continue the breakdown of glycogen and leaves a **short branch of 4 residues** (**Figure 18**)<sup>180</sup>. The GDE enzyme is required to remove these branching points. GDE has **two distinct enzymatic activities**: a glucosyltransferase activity (GT), close to the one of GBE1, and an  $\alpha$ 1,6-glucosidase activity (GC)<sup>180</sup>. The **GT activity** results in the transfer of the 3  $\alpha$ 1,4-linked residues (in blue in **Figure 18**) on the non-reducing end of another branch of the glycogen molecule. Then, the **GC activity** removes the last glucose residue (in green in **Figure 18**) and therefore the  $\alpha$ 1,6 bond, generating a molecule of glucose. The phosphorylase can then resume the degradation of the linear chain<sup>180</sup>. A detailed structure of GDE is provided in a next section.

The **regulation of glycogen breakdown** is also both hormonal and allosteric<sup>180</sup> and is different between muscles and liver. Since the **muscle phosphorylase** aims at degrading glycogen in order to generate energy to sustain muscle contraction during exercise, it is mostly activated allosterically by **increase of adenosine monophosphate**, which reflects an energy deficit<sup>180</sup>. Increase of adenosine triphosphate and G6P at rest rather blocks glycogen breakdown and favors glycogen synthesis<sup>180</sup>. In contrast, since the major role of glycogen in the liver is to produce glucose in order to maintain the glycemia, the liver phosphorylase is activated in **absence of glucose** rather than by the energy deficit<sup>180</sup>. These differences are also highlighted by the different **hormonal regulation**: whereas the muscle phosphorylase is mostly activated by epinephrine released during exercise, the liver phosphorylase is mostly activated by glucagon released during fasting<sup>180</sup>. The signaling pathways of both hormones converge on the activation of a common kinase (although with tissue-specific subunits), the **phosphorylase kinase**, which activates the phosphorylase and therefore the breakdown of glycogen<sup>180</sup>.

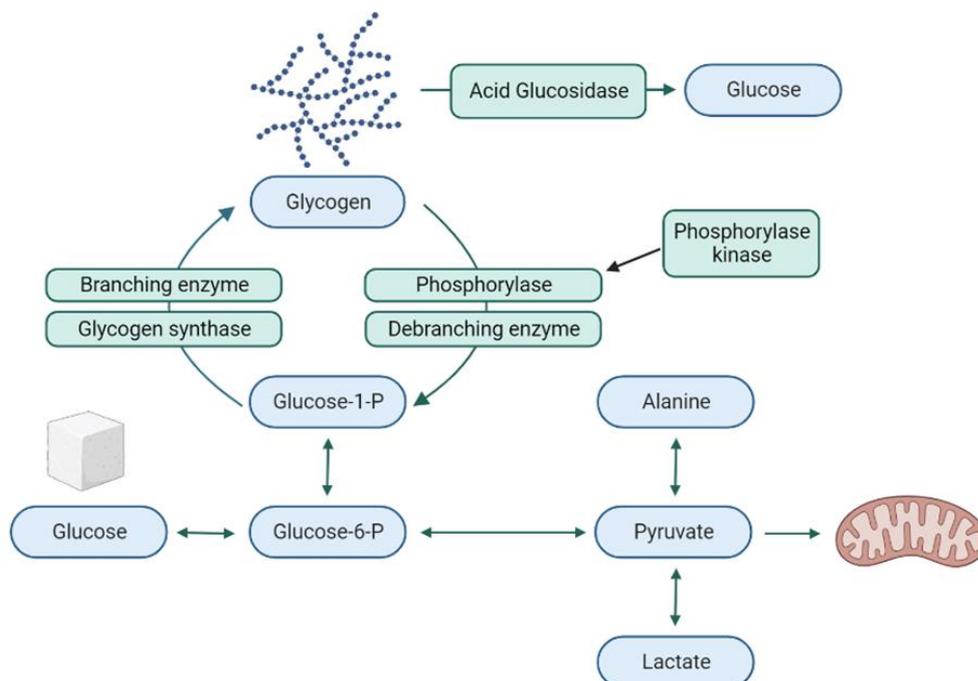
A less understood pathway of glycogen degradation involves the lysosome and the lysosomal enzyme **acid  $\alpha$ -glucosidase** (GAA)<sup>179</sup>. The exact mechanism and physiologic functions of this pathway are still unclear, but its **relevance in humans is important**, as demonstrated by the extremely severe muscle phenotype of patients with deficiency of the GAA enzyme (Pompe disease)<sup>179</sup>. It seems that glycogen is able to reach the lysosome through a specific form of autophagy known as « **glycophagy** »,

which requires both the STBD1 and GABARAPL1 proteins in order to anchor glycogen to autophagosome and to recruit the autophagic machinery respectively<sup>179</sup>.

#### d. GLYCOGEN IN PHYSIOLOGY

A simplified overview of glycogen metabolism is depicted in **Figure 19**. Glycogen is mostly present in **liver** (10% of liver weight), in **muscles** (2% of weight) and **brain**, with a different function according to the organ<sup>179,180</sup>.

In the **liver**, glycogen enables a continuous production of glucose to **maintain the glycemia**<sup>180</sup>. In fed conditions, high insulin levels result in the uptake and storage of glucose as glycogen, by the activation of Gys. In fasting, high glucagon levels result in the breakdown of glycogen and the increase of G6P. The latter can be converted in glucose by the enzyme **glucose-6-phosphatase (G6Pase)**<sup>180</sup>. If fasting continues and glycogen stores are exhausted, the liver uses the **gluconeogenesis pathway** to produce glucose from **lactate** and **alanine** (released by muscles) and from **pyruvate** obtained by the oxidation of fatty acids (released by white adipose tissues)<sup>180</sup>. G6Pase is only expressed in the liver, kidney and intestine, the only organs capable to export glucose into the bloodstream<sup>180,182</sup>.



**Figure 19. Simplified overview of glycogen metabolism.** Made with Biorender.

In the **muscle**, glycogen is used as an energy source to **sustain fast muscle contraction** during exercise<sup>180</sup>. After a meal, the elevated level of insulin results in the entry of glucose in myocytes and its storage as glycogen<sup>180</sup>. During exercise, the **energy deficit** (elevation of adenosine monophosphate) and eventually the **rise of plasma epinephrine** lead to the breakdown of glycogen and increase G1P

and G6P<sup>180</sup>. Since muscle do not express G6Pase, no glucose production is possible, and G6P is therefore used within the glycolytic pathway to produce energy and lactate. **Lactate is sent back to the liver** to produce glucose through gluconeogenesis (known as the Cori cycle)<sup>180</sup>. Once the glycogen stores are exhausted (endurance exercise), muscles will mostly use fatty acid oxidation as an energy source<sup>180</sup>. The **heart** deserves specific considerations: this muscle is constantly active and relies mostly on fatty acid oxidation for adenosine triphosphate generation<sup>183</sup>. But under sudden increase of energy demand, such as following a rise of plasma epinephrine to increase heart rate, the heart cannot increase the rate of fatty acid oxidation, already elevated, and will induce glycogen breakdown<sup>183</sup>.

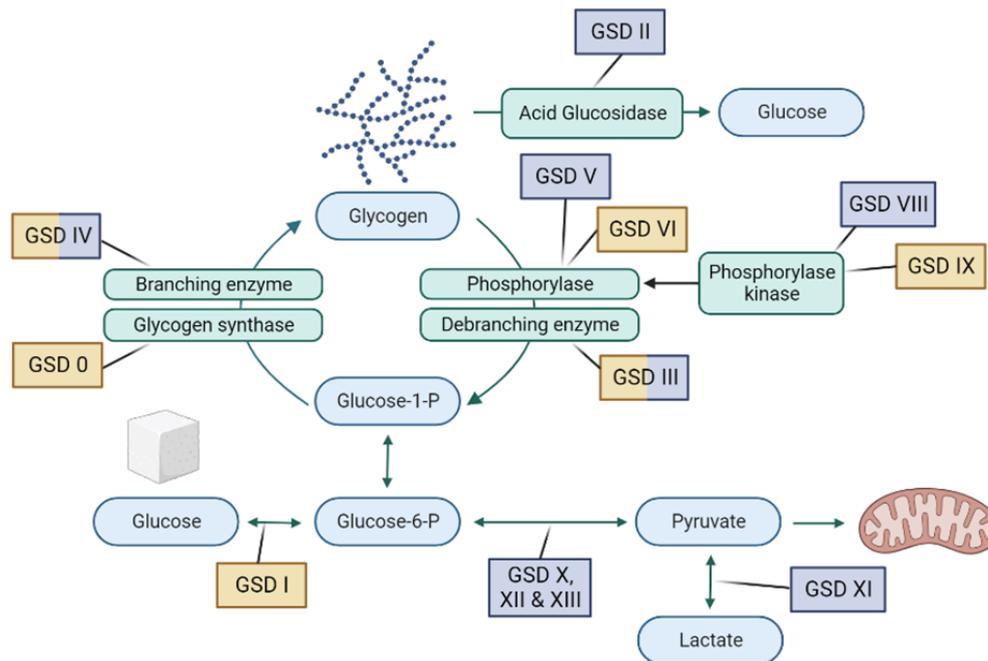
Finally, the specific role of glycogen in the **central nervous system** is still unclear. Amounts of glycogen are too little to play a role as an energy source and glucose is anyway constantly provided to the brain by the liver through the blood<sup>184</sup>. Glycogen is mostly present in **astrocytes** and one recent study reported that brain glycogen is made up to 25% of **glucosamine** instead of glucose, unlike in muscles or liver<sup>185</sup>. Glycogen from astrocytes supports the neurons, for instance by enabling the synthesis of the **neurotransmitter glutamate** and the proper **glycosylation of proteins** (through glucosamine release)<sup>184,185</sup>. Interestingly, glycogen can also be produced in neurons in stress conditions and activates **autophagy**<sup>184,186</sup>. An excessive production of neuronal glycogen, as observed in Lafora disease, leads to the loss of neurons<sup>184,186</sup>, emphasizing that its metabolism requires fine regulation.

#### e. GLYCOGEN STORAGE DISEASES

The **absence of an enzyme of glycogen metabolism** leads to a glycogen storage disease (GSDs) that can affect the liver, the muscles, the heart and/or the brain<sup>178</sup>. The most frequent GSDs are depicted in the **Figure 20**.

Clinical manifestations of GSDs are secondary either to the **absence of glycogen use** and/or to direct toxicity related to the **accumulation of normal or abnormal glycogen**<sup>178</sup>. In **liver GSDs** (GSD0, GSDI, GSDIII, GSDVI, GSDIX), the absence of glycogen use leads to fasting **hypoglycemia** which are particularly severe in GSDI because of the combined blockade of glycogen breakdown and gluconeogenesis<sup>178</sup>. In GSD I, further metabolic complications secondary to the accumulation of G6P, such as elevation of lactate, triglycerides and uric acid, occur<sup>187</sup>. In **muscle GSDs** (GSDIII, GSDV, GSDVIII, GSDX à XIII), the **energy deficit during exercise** manifests as muscle pain, muscle weakness or muscle lysis (rhabdomyolysis) leading to exercise arrest<sup>178</sup>. In some diseases, a **direct toxic effect of glycogen accumulation** is present such as in debranching- (GSDIII) or branching enzyme (GSDIV) deficiencies<sup>177,178</sup>. In both these diseases, the accumulation of **abnormal glycogen**, highly branched (limit dextrin) or poorly branched (amylopectin-like) leads to cell death resulting in **chronic liver disease** (cirrhosis) and **permanent (cardio)myopathy**<sup>178</sup>. In contrast, diseases with accumulation of glycogen of overall normal structure (GSDI, GSDV, GSDVI, GSDIX) usually do not present with such severe complications<sup>178</sup>.

Finally, Pompe disease (GSDII) behaves as a lysosomal storage disease, with lysosomal dysfunction and autophagy impairment, and is definitely aside compared to other GSDs<sup>188</sup>.



**Figure 20. Overview of glycogen storage diseases (GSD).** Each GSD is depicted in a box (**yellow** for GSDs affecting the liver; **blue** for GSDs affecting the muscles/heart) close to their enzymatic defect. GSD V and VI affect the muscle and hepatic isoforms of the phosphorylase respectively, and GSD VIII and IX affect the muscle and hepatic isoforms of the phosphorylase kinase respectively. GSD X, XII and XIII affect the glycolysis pathway at different levels. *Made with Biorender.*

## 2. CLINICAL ASPECTS OF GLYCOGEN STORAGE DISEASE TYPE III

GSDIII manifests early in life with initial symptoms related to **liver and metabolic** impairments, then a **muscle and heart involvement** appears during late childhood or early adulthood<sup>176,177</sup>. This rare disease has an incidence estimated to be around 1/100 000 births<sup>177</sup>.

### a. CLINICAL MANIFESTATIONS

The earliest manifestations of GSDIII are noticed around 1 year of age, with **hypoglycemia after short fasting, hepatomegaly** and **growth retardation**<sup>177</sup>. Elevation of plasma aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels are present since childhood<sup>189</sup>. The improvement of the AST and ALT levels with age<sup>189</sup>, along with the spontaneous increase of the time to hypoglycemia during fasting with age has led to the assumption that the liver and metabolic diseases improve with age in GSDIII. But two recent studies have suggested that **it is not the case**. The first study analyzed liver biopsies of 13 patients with GSDIII<sup>189</sup> and has shown the presence of glycogen

accumulation in hepatocytes, portal and lobular inflammation (in 9 out of 13) and more importantly the **presence of significant fibrosis** in all patients, including portal fibrosis (n = 2), septal fibrosis (n = 9) and cirrhosis (n = 2)<sup>189</sup>. The fibrosis is present **since 9 months of age** and is more severe in adult patients, suggesting progression of the disease with age<sup>189</sup>. The second study reports the outcome of patients at adulthood and revealed that adult patients displayed hypoglycemia (40%), hepatomegaly (50%) and **cirrhosis** (40%)<sup>190</sup>. Liver transplantation have been reported, as well as the occurrence of liver tumors (adenoma, HCC)<sup>177,190</sup>. To conclude, although the metabolic phenotype improves with age, the liver disease is not improving in GSDIII patients, but rather evolves toward **severe liver fibrosis** with cirrhosis and a risk of liver tumors.

A **muscle impairment** appears in late childhood or in teenagers in 85% of cases (GSDIIIa)<sup>176,177</sup>. Initially, the myopathy manifests as muscular pain during exercise and can progress toward loss of ambulation in one third of adult patients<sup>190</sup>. An elevation of the plasma creatine phosphokinase (CPK) level is usually absent in young children and apparent at a median age of 10 years<sup>177</sup>. Around 60% of patients will develop a **cardiac impairment**, initially only present in ultrasonography (left ventricle hypertrophy) and around 15% of patients have overt symptoms at adulthood (dyspnea)<sup>177,190</sup>. Heart transplantation has been reported in few cases, as are cases of sudden deaths<sup>190</sup>. In contrast, 15% of patients will **not present any muscle or heart phenotype** (GSDIIIb). These patients have specific *AGL* mutations (see below)<sup>176,191</sup>.

Others rarely reported features includes bone fractures, type 2 diabetes and polycystic ovary syndrome<sup>177</sup>. No overt neurological involvement is present, although one study reported the presence of mild impairments in social cognition and executive functions in some patients<sup>192</sup>. Finally, almost all patients reach adulthood when treated, but with a significant motor handicap<sup>177</sup>.

#### b. DIAGNOSIS AND THERAPEUTIC APPROACHES

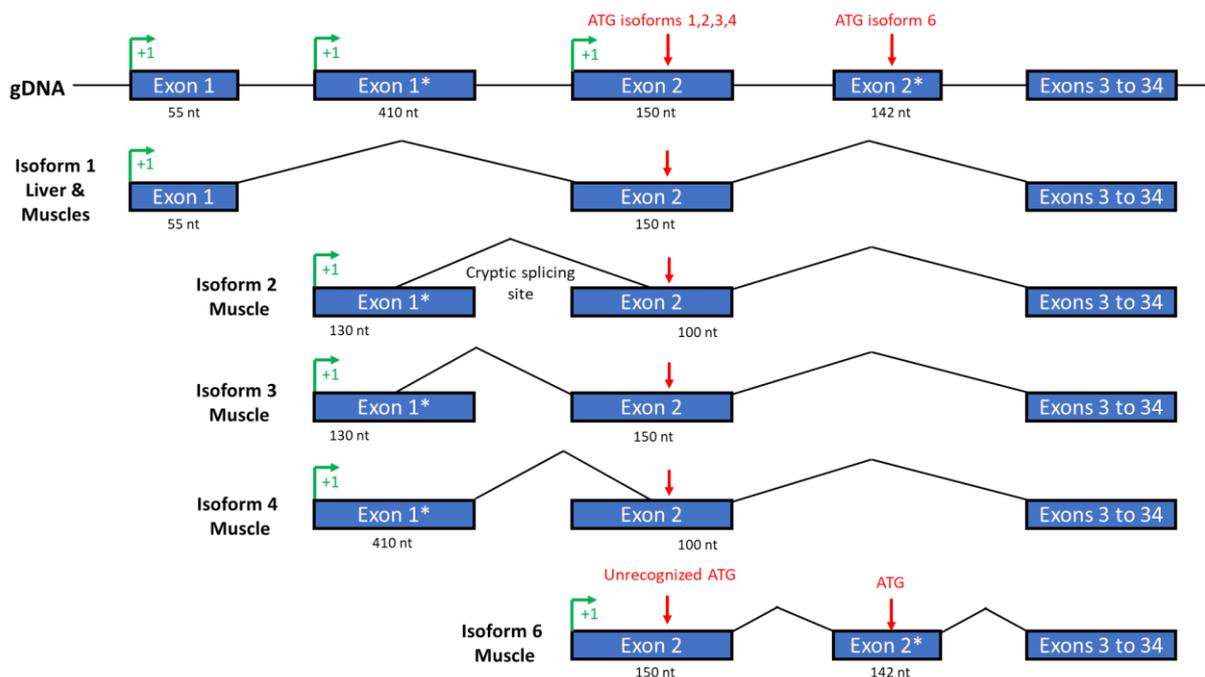
In case of clinical suspicion, the diagnosis can be made by the measurement of the **enzymatic activity of GDE** in leukocytes or rarely on muscle biopsies, or with the **molecular analysis** of the *AGL* gene<sup>193</sup>. The treatment is only supportive, since no curative treatment is currently available for patients with GSDIII, and relies mostly on the **prevention of the hypoglycemia**<sup>193</sup>. Children have frequent meals, at specific hours including during the night to avoid prolonged fasting, and can require the use of a enteral feeding tube (during the night)<sup>193</sup>. The diet is meant to reduce the use of simple sugars, which are going to increase the synthesis of glycogen and worsen disease manifestations, especially in heart, and the patients usually require the addition of complex sugars such as **uncooked cornstarch** to control the hypoglycemia. The diet can be **enriched in proteins** in order to improve the production of glucose by the liver from amino acids (gluconeogenesis)<sup>193</sup>, and **ketones bodies** have been tested in patients with significant cardiac involvement with promising results<sup>194</sup>. These guidelines allow a **good control**

of hypoglycemia and can improve in part the cardiac phenotype, but have **few or no effect on the progression of the muscle disease**, which remains a significant burden at adulthood<sup>177,190,193</sup>.

### 3. PHYSIOPATHOLOGY OF GSDIII

#### a. AGL GENE LOCUS

The **AGL gene** is located on chromosome 1 (1p21.2 locus) and contains **34 coding exons** (exon 2, exon 2\* et exons 3 à 34) and two non-coding exons (exon 1 et exon 1\*)(**Figure 21**). The naming of the exons evolved with time and is based on the exons of the main isoform (isoform 1). Indeed, **5 different mRNA isoforms** can be produced (mRNA isoforms 1, 2, 3, 4 et 6), because of the presence of 3 different transcription starting sites (exon 1, exon 1\*, exon 2). The first four mRNA isoforms encode the **same 1532-amino acid long protein** (protein isoform 1), with a translation beginning in exon 2. This is the main isoform, present in **both liver, muscles and heart**. In contrast, the minor mRNA isoform 6 include exon 2, exon 2\* and exons 3 to 34 and lead to the production of a **shorter 1516-amino acid long protein** (protein isoform 6). The ATG start codon in exon 2 is not recognized and translation begin in exon 2\*, thus generating a protein with a shorter N-terminal sequence. This isoform is only expressed **in muscles and heart**.

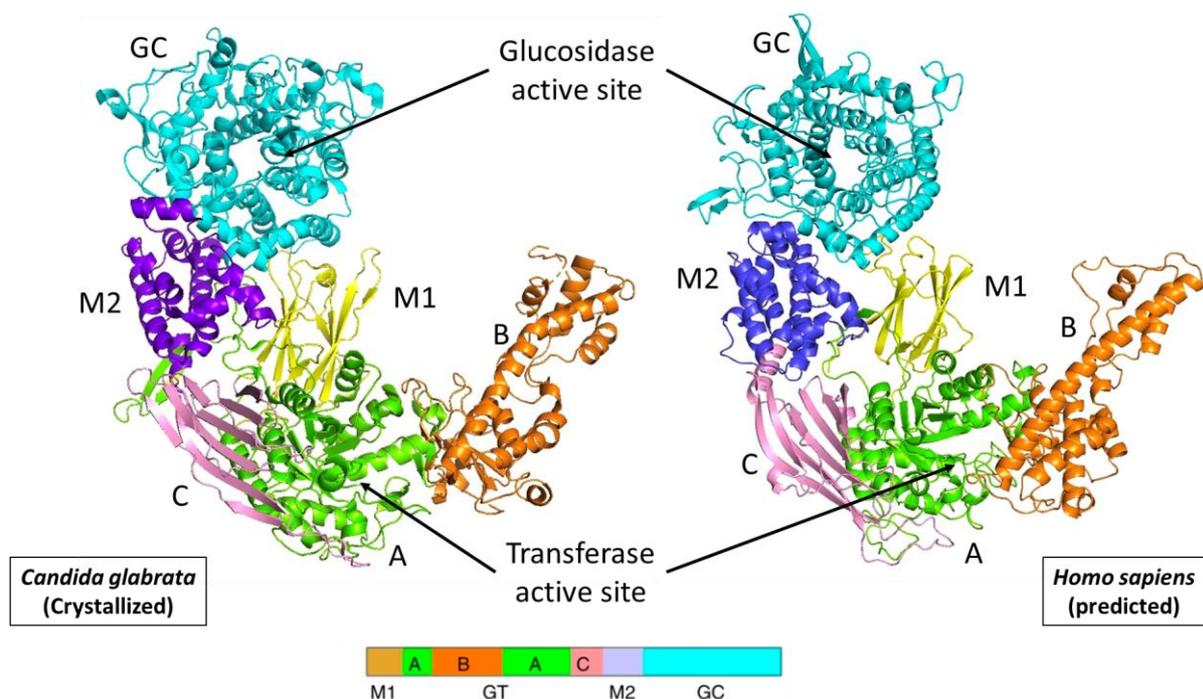


**Figure 21. Locus of the human AGL gene and schematic representation of the different mRNA isoforms.** The four first mRNA isoforms encode protein isoform 1 (translation in exon 2) and have a variable first non-coding exon (complete or partial exon 1 ou exon 1\*). The mRNA isoform 6 includes exon 2 (here non-coding) and the specific exon 2\* in which translation of a shorter protein isoform (isoform 6) begins. *Data extracted from NCBI.*

The latter isoform explains the **occurrence of a GSDIIb phenotype**, with only an hepatic and metabolic disease without any muscular or cardiac phenotype (15% of patients). These patients have **one of two recurrent mutations** present in exon 2 (exon 3 in older publications): p.Gln6\* et c.18\_19delGA<sup>177,191</sup>. Both truncated mutations affect only the ubiquitous protein isoform 1, leading to a **complete absence of GDE in the liver**. On the contrary, since the translation of the protein isoform 6, expressed only in muscle, begins in exon 2\* downstream of these two mutations, complete and normal **GDE is then expressed in muscles and heart**, which explains the GSDIIb phenotype.

#### b. GDE ENZYME STRUCTURE

The structure of GDE is **largely conserved throughout evolution**, with four main domains: a N-terminal domain, a first catalytic domain (GT activity), a central domain and a 2<sup>nd</sup> catalytic domain (GC activity)<sup>195</sup>. This organization is present in almost all eukaryotes, but prokaryotes and archaea have two distinct enzymes for each activity: an enzyme close to the branching enzyme for the GT activity and an enzyme with only the GC activity (known as pullulanase)<sup>195</sup>.



**Figure 22. Representation of the 3D crystal structure of *Candida glabrata* GDE (left) and of the prediction of the human GDE 3D structure using the Alpha-Fold2 software (right).** Structures retrieved from the Protein Data Bank/PDB (N°5D06) and the Alpha-Fold Protein Structure Database (<https://alphafold.ebi.ac.uk>, N°AF-G1UI17-F1) for the *Candida glabrata* and human GDE respectively. Limits of domains based on the publication by *Zhai et al.*<sup>196</sup>.

There is no crystallographic structure available for the human GDE protein. The only studies which have analyzed the structure of an eukaryotic GDE focused on the one of the fungus *Candida glabrata*<sup>196,197</sup>. The latter displays 38% of identity to human GDE, which is important considering the evolutive distance between humans and *Candida glabrata*<sup>196</sup>. The authors reported a 3D structure with 4 main domains, as predicted by the phylogenetic analysis (**Figure 22**): a **N-terminal domain named M1** (in yellow), a first **GT catalytic domain** with two sub-domains A (in green) and B (in orange), a **central domain** with two sub-domains C (in pink) and M2 (in purple), and a C-terminal **GC catalytic domain** (in cyan)<sup>196</sup>. By mutagenesis of key amino acid residues, the authors demonstrated the presence of the **GT and GC enzymatic activities in the A-B and GC domains** respectively, as well as the role of the B domain in glycogen binding<sup>196</sup>. By analyzing the positions of 20 missense mutations described in GSDIII patients, they observed that the orthologs of these amino acids were almost **all located within catalytic sites or glycogen binding domains**<sup>196</sup>.

During the summer 2021, the predicted structures of all human proteins using the Alpha-Fold 2 software (DeepMind) were released in the Alpha-Fold Protein Database (<https://alphafold.ebi.ac.uk>), including the one of the human GDE. Its structure is displayed along with the one of *Candida glabrata* in **Figure 22**. Although the sequence identity between of the two proteins is only of 38%, the structure is almost completely conserved with the **four domains arranged in an identical way**. Few functional studies have been performed on human GDE, but at least one study confirmed the presence of the GT catalytic site in N-terminal and GC catalytic domain in C-terminal<sup>198</sup>. This suggests that the overall structure and the structure-function relationships are largely conserved between species.

### C. AVAILABLE ANIMAL MODELS

**Two mouse models**<sup>199,200</sup> and a **dog model**<sup>201</sup> of GSDIII have been reported until now. Our team has also recently generated a **GSDIII rat model** (*manuscript in preparation*). All models recapitulate most of the features of the human disease.

The **first mouse model** has been generated by constitutive deletion of exons 5 to 10 of the *Agf* gene<sup>69,199</sup>. This is the mouse model used in this thesis. *Agf*<sup>-/-</sup> mice have a complete GDE deficiency and exhibit **progressive mortality** after repeated fasting<sup>199</sup>. Liver disease is present with hepatomegaly, glycogen accumulation, AST and ALT elevation and **liver fibrosis**, especially after 18 weeks of age. Liver tumors are present in 14-month-old animals (unpublished observations). Fasting hypoglycemia are present, and fasting worsens the biochemical parameters (plasma ALT, AST and CPK levels). In the **muscles and heart**, significant glycogen accumulation is present, with **motor strength impairment** (in grip-strength test). CPK are not elevated except in fasting conditions.

The **second reported mouse model** harbors a constitutive deletion of the exons 32 to 34 of the *Agf* gene in a C57BL/6 background<sup>200</sup>. Disease features are close to the former mouse model, with

**progressive lethality** (50% at 9 months of age), **glycogen accumulation** in liver, heart and muscles, the occurrence of **hypoglycemia** in fed and long fasting (16h) and **elevation of plasma AST, ALT and CPK levels**. Of note, there is **no liver fibrosis** (even in old animals) or tumors. In line with glycogen accumulation in muscles, mice display **impaired muscle strength** in the *treadmill* test (endurance) and in the *rotarod* test (coordination) starting from 3 months of life. The muscle phenotype worsens at 9 months of age, emphasizing the **progressive nature** of the disease. Histology reveals the presence of glycogen with alteration of the architecture of the muscle fiber in electronic microscopy. No overt neurological phenotype is observed except the accumulation of PAS-positive inclusions in the cerebellum<sup>200</sup>.

Finally, a **spontaneously occurring dog model** of GSDIII has been reported, with a frame-shift mutation in the *AGL* gene leading to the **loss of the last 126 amino acids**<sup>201</sup>. These dogs display a liver and muscle phenotype and the heart phenotype has not been analyzed in the initial publication. In the **liver**, an early accumulation of glycogen is present with hepatomegaly and elevation of AST and ALT from 2 to 4 months, increasing with time. **Glycemia is normal** in fed conditions (without evaluation at fasting). A **very severe liver fibrosis** is present with cirrhosis and nodules at 16 months of age, this model being much more severe than its murine counterparts<sup>201</sup>. An increase of glycogen is also present in muscles from 4 months of age, but the increase of CPK in plasma **occurs late in life** (after 1 year), as observed in human<sup>177,201</sup>. Histology reveals that accumulated glycogen leads to **distortion of the myofibrils**.

#### d. PATHOPHYSIOLOGIC MECHANISMS

Animal models, as well as a few studies performed in patients, give some insights about the progression of the disease and the mechanisms involved<sup>169,199-202</sup>.

In the **liver**, the accumulation of glycogen is an early feature of the disease and is associated with **hepatocytes death** and elevation of liver enzymes<sup>199-201</sup>, **compensatory cell proliferation**<sup>202</sup> and progressive activation of myofibroblasts, **fibrosis deposition** and cirrhosis with nodules<sup>199,201,202</sup>, in a pattern common to many chronic liver diseases<sup>203</sup>. Nonetheless, the direct mechanistic link between the accumulation of limit dextrin and hepatocyte death is still unclear. This mechanism is different of the one leading to the metabolic phenotype (hypoglycemia, low lactate, increase of triglycerides). Indeed, a recent study evaluating a **substrate reduction therapy** using a glycogen synthase inhibition by siRNA demonstrated a complete normalization of the liver phenotype (AST and ALT levels, hepatocyte proliferation, liver fibrosis, occurrence of nodules) with the normalization of the glycogen in the liver, but without any effect on hypoglycemia<sup>202</sup>. Therefore, this phenotype is really secondary to a **toxic effect of limit dextrin accumulation** rather than by a consequence of the inability of the liver to use glycogen.

In the **muscle**, studies in animal models<sup>200,201</sup> and in patients<sup>169</sup> suggest an early accumulation of glycogen but a free interval before elevation of CPK levels and onset of a clinically meaningful motor impairment. Accumulation of glycogen results in the distortion of myofibrils, impeding muscle contraction, **without signs of muscular dystrophies**, such as necrosis-regeneration and myofibrillar degeneration<sup>169</sup>. Myocyte loss is likely a **late feature of the disease**, as evidenced by the late increase of CPK in plasma of patients and dogs<sup>177,201</sup>. Architectural distortion is probably not the sole pathogenic mechanism, as histologic studies in patients report the presence of “**vacuolar myopathy**” with many **glycogen-rich autophagosomes**<sup>169</sup>. This feature suggests an impairment of autophagy, and such histologic pattern has already been described in other models with autophagy impairment such as the Danon disease rat model<sup>204</sup> or the *Vps15*<sup>-/-</sup> mouse model<sup>205</sup>. Therefore, it is possible that a reduced lysosomal clearance of the glycogen could be present, in addition to a direct toxic effect of glycogen accumulation. No studies have evaluated the potential impairment of autophagy in the liver. The use of **mTOR inhibitors**, known to induce autophagy, in the GSDIII dog model, led to a rather modest improvement of the liver and muscle phenotypes<sup>206</sup>. Since mTOR inhibitors can also block the synthesis of glycogen through the inhibition of glycogen synthase, it is difficult to decipher the precise role of autophagy inhibition.

#### 4. GENE THERAPY APPROACHES IN ANIMAL MODELS OF GSDIII

Several innovative gene therapy approaches have been evaluated in the GSDIII mouse models. The main issues are the **large size of the GDE cDNA** (4.6 kb), limiting its packaging in rAAV vectors, and the **requirement to treat several organs** (liver, heart, muscles)<sup>143</sup>. Three main strategies have been published: the **overexpression of GAA**, a **dual rAAV vector** and the **use of a bacterial GDE ortholog**.

##### a. OVEREXPRESSION OF THE GAA ENZYME

This approach relied on an rAAV vector initially developed for Pompe disease, expressing the **GAA enzyme under the control of a liver promoter**<sup>112,129</sup>. The rationale was to enhance the lysosomal breakdown of glycogen to compensate for the deficient cytosolic breakdown due to the absence of GDE. After injection of this vector in *Agf*<sup>-/-</sup> mice, a **supraphysiologic dose-dependent increase of GAA** activity was observed in liver and muscles, suggesting correct targeting of GAA to lysosomes of these organs<sup>69</sup>. Interestingly, **glycogen clearance was observed in the liver**, but not in heart and muscles<sup>69</sup>, suggesting that the glycogen did not reach the lysosome in heart and muscles. Similar results were observed in GSDIII primary muscle cells treated with the recombinant GAA (Myozyme®) replacement therapy, with low if any glycogen correction<sup>207</sup>. These data are in line with a possible blockade of autophagy in muscles (and heart) of both patients and GSDIII mouse models, that impedes the trafficking of glycogen to the lysosomes.

#### b. DUAL VECTOR GENE THERAPY STRATEGY

In the same publication, *Vidal and colleagues* also described a **dual vector strategy** to correct GSDIII in the *Agf<sup>-/-</sup>* mouse model<sup>69</sup>. Two rAAV9 vectors, each containing half of the expression cassette, were injected intravenously in *Agf<sup>-/-</sup>* mice at a high dose ( $5 \times 10^{13}$  vg/kg for each vector). With a CMV promoter, they obtained a **modest expression of GDE** (around 5 to 10% of control), but a **significant reduction of glycogen** content in heart and muscles (> 50% reduction) and a **rescue of the muscle strength** impairment assessed using the Wire Hang Test 3 months after injection<sup>69</sup>. They were able to show that a **GDE expression around 10%** was sufficient to achieve a reduction of more than 80% of the accumulated glycogen in muscles<sup>69</sup>, emphasizing that the restoration of a normal enzymatic activity is not a prerequisite for significant phenotype correction. The CMV promoter is known to be extinct in the liver<sup>208</sup>, and therefore no correction of the liver phenotype was observed<sup>69</sup>. The specific targeting of the liver using a capsid with liver tropism (rAAV8) and a liver promoter at the same dose could only enable a partial **correction**<sup>69</sup>. This suggests that the correction of the liver phenotype is difficult to achieve and will probably require a broad targeting of hepatocytes.

#### c. USE OF A BACTERIAL ENZYME

Another team developed an innovative gene therapy approach using the bacterial debranching enzyme **pullulanase**<sup>209</sup>. As mentioned before, the debranching activity in bacteria is carried by two different sets of enzymes and pullulanase is responsible for the  **$\alpha$ 1,6-glucosidase activity**. It is able to cleave directly the  $\alpha$ 1,6 bonds, generating shorts polymers of 4 glucose residues (maltotetraose)<sup>195,209</sup>. Its **small 2.2 kb cDNA** makes it an ideal target to be packaged in rAAV vectors. By using a dose of  $1 \times 10^{13}$  vg/kg and a muscle and heart-targeting vector, the authors demonstrated the presence of a high pullulanase activity in these tissues, leading to **glycogen clearance** and motor strength improvement (treadmill evaluation) 10 weeks after injection<sup>209</sup>. In order to correct the liver phenotype, they also used a rAAV vector with a liver promoter that could correct the glycogen accumulation, the elevation of liver enzymes and **prevent the development of fibrosis 10 weeks** after injection<sup>209</sup>. However, the main limitation of this approach, reported in their second publication, is the development of an **important immune response** to the bacterial transgene with subsequent progressive loss of correction<sup>42</sup>. The use of dual promoter with liver and muscle expression could allow an almost complete correction of both liver and muscles<sup>42</sup>. It also led to a reduction of the immune response, without fully preventing it, in line with a possible tolerance due to the transgene expression in the liver<sup>42</sup>. The presence of this strong immune response is likely to be a major drawback for this interesting approach in an eventual clinical trial in GSDIII patients.

# OBJECTIVES

## OBJECTIVES

Gene therapy using rAAVs is an interesting approach for the treatment of monogenic diseases such as inborn errors of metabolism, with **promising results for liver or muscle** targeting in clinical trials in humans. After intravenous injection, rAAVs can reach its target organs (skeletal muscles and/or liver) and achieve a long-term expression of the gene of interest, as long as there is no significant immune responses or cell proliferation. Especially in inborn errors of metabolism, a **small fraction** of the normal enzymatic activity can be sufficient to improve or even correct the phenotype of the disease. Therefore, an rAAV-based gene therapy approach **seems ideal** for the liver and muscle involvement of GSDIII.

Nonetheless, the correction of GSDIII using this approach faces **several challenges**. First, the **gene of interest is too large to be fully encapsidated** in a single expression cassette. Second, **two very different organs**, the skeletal muscles (and the heart) and the liver, must be targeted at the same time to achieve the complete correction of the phenotype. Whereas the liver is usually targeted with low doses of rAAVs, skeletal muscles often require the use of high dose of rAAV vectors, with potential liver toxicities as observed in several clinical trials such as for the MTM myopathy<sup>87</sup>. Moreover, combined targeting would require the use of a promoter with a dual specificity in order to express GDE in both organs. Third, unlike in many inborn errors of metabolism in which rAAV-mediated gene therapy has been evaluated, the **liver in GSDIII is chronically damaged** and complete correction might be more difficult to achieve<sup>143</sup>. Both fibrosis and hepatocyte proliferation are present, although at much lower levels than in other fibrotic liver diseases, and this could hamper the transduction and/or the persistence of rAAV. This would require to correct all hepatocytes at the same time in order to completely stop the progression of the liver disease. Last, GSDIII patients and mouse models develop **liver tumors**, which is another issue, both from a safety as well as from an efficacy point of view. For all these reasons, GSDIII is both an interesting and a challenging disease model in which to test a single rAAV vector approach with technological development.

Since the progressive myopathy and cardiomyopathy are **the major disease burden** in adult patients with GSDIII, we first focused on muscle and heart targeting. In the first part of this thesis, we describe the **optimization of a single rAAV strategy to target both muscles and heart while de-targeting the liver**, with the generation of a **functional truncated GDE variant** able to be encapsidated in a single rAAV vector. We then evaluated its ability to correct the muscle involvement in several models of GSDIII, namely the *Agf*<sup>-/-</sup> **mouse model**, a recently generated *Agf*<sup>-/-</sup> **rat model** and **human**

**AGL<sup>-/-</sup> induced pluripotent stem cells** (hiPSCs) generated by the team of Xavier Nissan at the I-STEM institute (Evry). This first part has been submitted to *The Journal of Clinical Investigation*.

In the second part of this thesis, we addressed the correction of the liver disease in GSDIII using a **liver-targeting single rAAV vector**. First, we characterized the hepatic phenotype of the *Agf<sup>-/-</sup>* mice and then analyzed the **efficacy and limitations of a single rAAV vector approach** for liver correction in GSDIII in both young and old *Agf<sup>-/-</sup>* mice.

PART I:  
CORRECTION OF THE MUSCLE  
DISEASE IN GSDIII

# Single AAV gene therapy with mini-GDE for glycogen storage disease type III

Antoine Gardin<sup>1,2†</sup>, Jérémy Rouillon<sup>1,2†</sup>, Valle Montalvo-Romeral<sup>1,2</sup>, Lucille Rossiaud<sup>1,2,3</sup>, Patrice Vidal<sup>1,2</sup>, Romain Launay<sup>4</sup>, Mallaury Vie<sup>1,2</sup>, Youssef Krimi Benchekroun<sup>1,2</sup>, Bérangère Bertin<sup>1,2</sup>, Guillaume Dubreuil<sup>1,2</sup>, Justine Nozi<sup>1,2</sup>, Louisa Jauze<sup>1,2</sup>, Romain Fragnoud<sup>1,2</sup>, Nathalie Daniele<sup>1</sup>, Laetitia Van Wittenberghe<sup>1</sup>, Jérémy Esque<sup>4</sup>, Isabelle André<sup>4</sup>, Xavier Nissan<sup>3</sup>, Lucile Hoch<sup>3</sup>, and Giuseppe Ronzitti<sup>1,2\*</sup>.

<sup>1</sup>Genethon, 91000 Evry, France

<sup>2</sup>Université Paris-Saclay, Univ Evry, Inserm, Genethon, Integrare research unit UMR\_S951, 91000 Evry, France

<sup>3</sup>CECS, I-STEM, Institute for Stem Cell Therapy and Exploration of Monogenic Diseases, 28 rue Henri Desbruères, 91100 Corbeil-Essonnes, France.

<sup>4</sup>Toulouse Biotechnology Institute, TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France. 135, avenue de Rangueil, F-31077 Toulouse Cedex 04, France

† Contributed equally

\*Corresponding author

## I. INTRODUCTION

Glycogen storage disease type III (GSDIII) is a rare inborn error of metabolism (incidence 1 in 100,000) caused by mutations in the *AGL* gene, encoding for the glycogen debranching enzyme (GDE, also known as amylo-alpha-1,6-glucosidase)<sup>176,177</sup>. GDE acts along with glycogen phosphorylase to degrade glycogen in the cytosol of virtually any cell. In liver and muscle, GDE function is required to maintain glycemia and to sustain fast muscle contraction, respectively<sup>176</sup>.

The earliest manifestations of GSDIII, such as fasting hypoglycemia, hepatomegaly, elevation of liver enzymes and growth retardation, are usually apparent from early childhood<sup>176,177</sup>. A severe myopathy develops during adolescence, with exercise intolerance, muscle weakness and then loss of ambulation in adult patients<sup>177,190,210</sup>. Histological analysis of muscle biopsies from GSDIII patients shows glycogen accumulation in large vacuoles which disrupts the myofibril architecture<sup>169</sup>. Most adult patients also display left ventricle hypertrophy on echocardiography, but only 15% have overt cardiomyopathy<sup>177,210</sup>. Although it was thought for long that liver involvement improved with age, recent studies report liver fibrosis from an early age with cirrhosis and liver tumor development in adult patients<sup>189,190</sup>. To date, no curative treatment is available for patients with GSDIII and therapeutic

option mostly consists in frequent meals, the use of complex carbohydrates such as uncooked cornstarch and a high-fat high-protein diet<sup>176,193</sup>. Dietary management is able to reduce the frequency of hypoglycemia and to improve in some cases the cardiac phenotype, but has little effect on the myopathy<sup>176,193</sup>.

Gene therapy using recombinant adeno-associated viral vectors (rAAV) is a promising strategy to treat inherited diseases<sup>1,31</sup> and has already been successfully used in spinal muscular atrophy<sup>211</sup>, hemophilia A and B<sup>127,212</sup> and congenital blindness<sup>213</sup> with market approval in these indications. Multiple clinical trials are currently ongoing for liver or muscle genetic diseases, with encouraging results<sup>1</sup>. One of the major constraints in the use of rAAV for gene transfer is its encapsidation size limited to 5 kb including the inverted terminal repeats (ITRs)<sup>31</sup>. In the setting of GSDIII, the 4.6 kb GDE cDNA represents a technical challenge toward the development of a single vector strategy for GSDIII<sup>143</sup>. Few years ago, our team published a proof-of-concept of GSDIII correction using a dual overlapping rAAV gene therapy strategy in which two vectors were co-injected, each encoding half of the expression cassette<sup>69</sup>. Homologous recombination occurring within the tissues allows to produce full-length GDE. By using two distinct dual rAAV vector approaches, we demonstrated complete correction of the muscle and heart phenotype, but only partial correction of the liver disease<sup>69</sup>. An improvement of this approach, based on the use of a tandem liver-muscle promoter<sup>41</sup> in combination with rapamycin treatment, was recently reported<sup>214</sup>. Despite ongoing optimization<sup>72,73</sup>, the dual vector strategy has several limitations: it requires two vectors and a consequent two-time higher dose, the recombination efficacy could be suboptimal, and the expression of truncated proteins derived from non-recombined genomes could trigger potential immune-toxicities. Another strategy has been developed using a bacterial debranching enzyme (pullulanase), whose 2.2 kb cDNA is short enough to be encapsidated in a single rAAV vector<sup>209</sup>. Although this approach allows initial liver and muscle correction<sup>209</sup>, it led to immune response toward the bacterial protein and subsequent loss of correction<sup>42</sup>. Given the limitations of the existing AAV gene transfer technology, a strategy based on a single vector expressing a shorter and active form of GDE may represent a valid solution for a safe and efficient rAAV gene therapy for GSDIII in humans.

In GSDIII patients, the liver displays liver fibrosis from early childhood<sup>177,189,190</sup> and significant, sometimes lethal, liver toxicities have been reported in clinical trials using high-dose rAAVs<sup>86,87</sup>. Based on these considerations, here we focused on the treatment of heart and muscle impairment while de-targeting the liver with a recently developed muscle-tropic AAV vector<sup>173</sup>.

Here, using *Agf* knock-out mouse and rat models, as well as a human cellular model of GSDIII, we demonstrate the possibility to engineer a single vector encoding a mini-GDE enzyme to correct the muscular and cardiac manifestations of the disease, the major disease burden in adult GSDIII patients<sup>177,190,193</sup>.

## II. MATERIAL AND METHODS

*In Vivo Studies.* The *Agl* knock-out mice ( $AgI^{-/-}$ ,  $AgI^{tm1b(EUCOMM)Wtsi}$ ) were previously described<sup>69</sup> and are of mixed background (C57BL/6J and Balb/c). The *Agl* knock-out rats ( $AgI^{-/-}$ ) were recently generated using the CRISPR tool and are of Sprague Dawley background. All animal studies were performed according to the French and European legislation on animal care and experimentation (2010/63/EU) and approved by the local institutional ethical committee. rAAV vectors encoding human GDE were either administered intravenously via the tail vein to 4-month-old male  $AgI^{-/-}$  mice and wild-type littermates ( $AgI^{+/+}$ ) or intramuscularly in the tibial anterior muscle of 4-month-old female  $AgI^{-/-}$  mice and  $AgI^{+/+}$  wild-type littermates. Intramuscular injection was performed under anesthesia by ketamine and xylazine. rAAV vectors encoding GDE were also administered in the tail vein of 6-week-old male  $AgI^{-/-}$  rats and  $AgI^{+/+}$  wild type littermates. Finally, rAAV vectors encoding mouse secreted embryonic alkaline phosphatase (mSEAP) were administered intravenously via the tail vein to 6-week-old wild type male C57BL/6J mice. All animals were randomized to receive rAAV or phosphate buffer saline (PBS) as controls. Mice were euthanized by cervical dislocation and rats by anesthesia overdosing.

*Production of rAAV Vectors.* All rAAV vectors used in this study were produced using an adenovirus-free transient transfection method and purified using a chromatographic method as described earlier<sup>85</sup>. Titers of the rAAV vector stocks were determined using a real-time qPCR using primers for the codon-optimized GDE cDNA (Forward: 5'-CTG AAG CTG TGG GAG TTC TT-3' and Reverse: 5'-CTC TTG GTC ACT CTT CTG TTC TC-3') or for inverted terminal repeats (Forward: 5'-GGA ACC CCT AGT GAT GGA GTT-3' and Reverse: 5'-CGG CCT CAG TGA GCG A-3', for vectors encoding mSEAP). The mSEAP expression cassette contains the specified promoter, the SV40 intron, the mSEAP cDNA and an SV40 poly-A. The GDE expression cassette contains the miniCMV promoter (corresponding to the nucleotides 175050\_175400 of the CMV genome, NC\_006273), the human full-length or truncated codon-optimized GDE cDNA<sup>69,215</sup> and either a bovine growth hormone polyA (bGh) or a short synthetic 58 bp polyA (polyA58)<sup>216</sup>. All cassettes were flanked by the inverted terminal repeats of AAV serotype 2 for vector packaging. The capsid used (AAV-MT) is a hybrid between AAV9 and AAVrh74, harboring a P1 peptide, as previously described<sup>173</sup>.

*Analytical ultracentrifugation (AUC).* AUC measures the sedimentation coefficient of macromolecules by following over time the optical density of a sample subjected to ultracentrifugation. The difference in the sedimentation coefficient, measured by Raleigh interference or 260-nm absorbance, depends on the content of viral genome in the capsid. AUC analysis was performed using a Proteome Lab XL-I (Beckman Coulter, Indianapolis, IN). An aliquot of 400  $\mu$ L rAAV vector and 400  $\mu$ L formulation buffer

were loaded into a two-sector velocity cell. Sedimentation velocity centrifugation was performed at 20,000 rpm and 20°C. Absorbance (260 nm) and Raleigh interference optics were used to simultaneously record the radial concentration as a function of time until the lightest sedimenting component cleared the optical window (approximately 1.5 h). Absorbance data required the use of extinction coefficients to calculate the molar concentration and the percent value of the empty and genome-containing capsids. Molar concentrations of both genome-containing and empty capsids were calculated using Beer's law, and percentages of full genome-containing and empty capsids were calculated.

*Viral genome analysis on agarose gel.* DNA was extracted using the High Pure Viral Nucleic Acid Kit (Roche, ref 11858874001). Purified viral DNA was then loaded on a 1%-Agarose gel (Eurobio Scientific) stained with SybrSafe® Gel Stain (Invitrogen) to visualize the viral DNA.

*Western Blot Analysis.* Mouse and rat tissues were homogenized in phosphate buffer saline (PBS, ThermoFisher Scientific) with cOmplete™ protease inhibitor cocktail (Roche, ref 4693132001). Protein concentration was determined using the Pierce™ BCA Protein Assay (Thermo Fisher Scientific) according to the manufacturer's instructions. A fraction of 50 µg of total proteins were loaded in each well for both PBS- and rAAV-injected *Agf*<sup>-/-</sup> mice and rats and 10 µg of total proteins per well for PBS-injected *Agf*<sup>+/+</sup> mice and rats. SDS-PAGE electrophoresis was performed in a 4–12% Bis-Tris gradient polyacrylamide gel (NuPAGE™, Invitrogen). After transfer, the membrane was blocked with Intercept Blocking buffer (LI-COR Biosciences) and incubated with an anti-GDE rabbit polyclonal antibody (16582-1-AP, Proteintech) and an anti-vinculin mouse monoclonal antibody (V9131, Sigma) for tissue lysates and with an anti-myomesin3 (Myom3) rabbit polyclonal antibody (17692-1-AP, Proteintech) for serum. The membrane was washed and incubated with the appropriate secondary antibody (LI-COR Biosciences) and visualized by Odyssey imaging system (LI-COR Biosciences). For *in vitro* experiments, normalization was performed using an anti-actin mouse monoclonal antibody (66009-1-Ig, Proteintech).

*HEK-293T cells transfection.* HEK 293T cells were transfected in 6-well plates using the Opti-MEM™ medium and Lipofectamine™ 3000 transfection reagent (ThermoFisher Scientific) according to the manufacturer's instructions. Cells were harvested 48-hours after transfection and lysed in PBS with 1% Triton X-100. Supernatants were collected after centrifugation at 11 000 x g and 50 µg of total proteins were used for Western blots.

*Measurement of Glycogen Content in tissues.* Glycogen content was measured indirectly in tissue homogenates as the glucose released after total digestion with *Aspergillus Niger* amyloglucosidase (Sigma-Aldrich, ref A1602). Samples were incubated for 10 min at 95°C and then cooled at 4°C. After the addition of amyloglucosidase (final concentration 4 U/mL) and potassium acetate pH 5.5 (final concentration 25 mM) at 37°C for 90 min, the reaction was stopped by incubating samples for 10 min at 95°C. A control reaction without amyloglucosidase was prepared for each sample and was incubated in the same conditions. The glucose released was determined using a glucose assay kit (Sigma-Aldrich) and the resulting absorbance was acquired on an EnSpire alpha plate reader (PerkinElmer) at a wavelength of 540 nm. Glucose released after amyloglucosidase was then normalized by the total protein concentration.

*Histology.* For muscle histology, heart, triceps brachii, quadriceps femoris, soleus and EDL were snap-frozen in isopentane previously chilled in liquid nitrogen. Serial 8- $\mu$ m cross sections were cut in a Leica CM3050 S cryostat (Leica Biosystems). To minimize sampling error, 2 sections of each specimen were obtained and stained with hematoxylin phloxine saffron staining (HPS) or periodic acid-Schiff (PAS) according to standard procedures. Images were digitalized using Axioscan Z1 slide scanner (Zeiss, Germany) under a Zeiss Plan-Apochromat 10X/0.45 M27 dry objectif (ZEISS, Germany). Tile scan images were reconstructed with ZEN software (ZEISS, Germany).

*Vector Genome Copy Number Determination.* Vector genome copy number was determined using a qPCR assay as previously described<sup>132</sup>. The PCR primers used in the reaction were located in the glucosyltransferase domain of the full-length and truncated codon-optimized GDE (Forward: 5'-CTG AAG CTG TGG GAG TTC TT-3' and Reverse: 5'-CTC TTG GTC ACT CTT CTG TTC TC-3') or in the inverted terminal repeats (Forward: 5'-GGA ACC CCT AGT GAT GGA GTT-3' and Reverse: 5'-CGG CCT CAG TGA GCG A-3') for mSEAP expressing cassettes. As an internal control, primers located within the mouse (Forward: 5'-AAA ACG AGC AGT GAC GTG AGC-3' and Reverse: 5'-TTC AGT CAT GCT GCT AGC GC-3') or rat (Forward: 5'-AAA ACG AGC GGT GAC ATG AGC-3' and Reverse: 5'-TTC AGT CAT GCT AGC GCT CC-3') Titin gene were used.

*RNA Expression Analysis.* Total RNAs were extracted from cell lysates using Trizol (Thermo Fisher Scientific). DNA contaminants were removed using the Free DNA kit (Thermo Fisher Scientific). Total RNAs were reverse transcribed using random hexamers and the RevertAid H minus first strand cDNA synthesis kit (Thermo Fisher Scientific). qPCR was performed with oligonucleotides specific for the codon-optimized GDE transgene (Forward: 5'-CTG AAG CTG TGG GAG TTC TT-3' and Reverse: 5'-CTC TTG GTC ACT CTT CTG TTC TC-3') and normalized by the levels of expression of the P0 ribosomal protein

*Rplp0* mRNA (Forward: 5'-CTC CAA GCA GAT GCA GCA GA-3'; Reverse: 5'-ATA GCC TTG CGC ATC ATG GT-3').

*mSEAP quantification.* mSEAP was quantified in tissues lysates using the Phospha-Light™ Kit (Applied Biosystems) following the manufacturer's instructions and was normalized by the total protein concentration.

*Muscle Function.* Forelimb wire-hang test was performed as already reported<sup>129,217</sup> at baseline and each month until euthanasia. A 4-mm-thick wire was used to record the number of falls over a period of 3 min. The average number of falls per minute was reported for each animal.

*Transduction of hiPSC-derived skeletal muscle cells.* The GSDIII<sup>CRISPR</sup> hiPSCs have been previously generated, using CRISPR-knock down of the *AGL* gene<sup>218</sup>. Controls hiPSCs were the isogenic cell line (CTRL1). GSDIII<sup>CRISPR</sup> and CTRL1 hiPSCs were differentiated into skMb as previously described<sup>218</sup>. After expansion in 96-well plate, hiPSC-derived skMb were transduced with LK03-rAAV vectors encoding either GFP or ΔNter2-GDE under the control of the miniCMV promoter, at a MOI of either 75 000 or 15 000 for 72 hours. Then, hiPSC-derived skMb were differentiated into skMt, as previously described<sup>218</sup>.

*Measurement of Glycogen Content in skMt.* After 4 days of differentiation into skMt, hiPSC-derived skMt were starved for 3 days in a no-glucose DMEM medium with 10% fetal bovine serum (ref 11966025, ThermoFisher Scientific) in order to induce glycogen degradation in CTRL1 skMt as previously described<sup>218</sup>. Cells were lysed using HCl 0.3M and Tris 450 mM pH 8.0. Glycogen was then quantified using the Glycogen-Glo™ assay kit (ref CS1823B01, Promega) and normalized using the CellTiter-Glo™ Luminescent Cell Viability Assay (ref G7570, Promega).

*Immunostaining assay.* SkMt derived from hiPSCs were fixed with 4% paraformaldehyde (Euromedex) for 10 min at room temperature. After two washes in PBS, cells were permeabilized with 0.5% Triton X-100 for 5 min and blocked in PBS solution supplemented with 1% bovine serum albumin (BSA, Sigma) for 1 h at room temperature. SkMt were stained for specific skeletal myogenic markers overnight at 4°C using primary antibodies (Desmin, ref AF3844 R&D 1:200; MF20, ref 3ea DSHB 1:50; Titin ref T5650 US Biological 1:50). After 3 washes in PBS, staining was revealed by appropriate Alexa Fluor secondary antibodies 1:1000 (Donkey anti-goat AF488 ref A11055 Invitrogen 1:1000; Donkey anti-mouse AF488 ref A21202 Invitrogen 1:1000) in the dark for 1 h at room temperature, and nuclei were visualized with

Hoechst solution 1:3000 (Invitrogen). Cell imaging was carried out with a Zen Black software-associated LSM-800 confocal microscope (Zeiss) with a 20X objective.

*Periodic Acid-Schiff (PAS) Staining on skMt.* PAS staining on hiPSC-derived skMt was performed with the PAS Staining Kit (Sigma-Aldrich) following the manufacturer's instructions. Briefly, cells were fixed with 4% paraformaldehyde for 10 min at room temperature. After 2 washes in PBS, cells were treated with PAS for 5 minutes at room temperature. After 3 washes in distilled water, cells were treated with Schiff's reagent for 15 min at room temperature. Finally, after 4 washes in tap water, staining was visualized using an EVOS XL Core microscope (Invitrogen). Images were processed and analyzed using Fiji custom-made scripts<sup>219</sup>. First, colors were split and only the green channel was kept as it was the most contrasted one. Then, images were manually thresholded into binary images where PAS signal was black and background white. The threshold was set for maximizing the difference between genotypes, and once calculated, the same threshold was applied to all images to quantify. The quantification of PAS staining was obtained using this formula : Area of PAS staining / Total area of image x 100, giving then a percentage of PAS staining within the image.

*Molecular Modelling.* Three-dimensional model of the full-length human GDE was retrieved from the AF2-database<sup>220</sup>. The 3D model of  $\Delta$ Nter2-GDE was predicted using AF2 v.2.1<sup>221</sup> and the default parameters. MD simulations were performed using Gromacs 2021.3<sup>222</sup> with CHARMM36 forcefield<sup>223</sup>. Solvation was done using explicit TIP3 water model. A cut-off of 1.2 nm and a switch function from 1.0 to 1.2 nm were used for short-range electrostatic and van der Waals interactions, respectively. Long-range electrostatic interactions were treated with particle mesh Ewald (PME, parameters by default) and periodic boundary conditions (PBC). The following protocol was used: (i) minimization of the system with 50,000 steps of steepest descent algorithm, (ii) 100ps of NVT equilibration, (iii) 100 ps of NPT equilibration, (iv) 250ns production phase without constraints and a timestep of 2 fs. MD simulations were run at 303 K in triplicates. To analyze MD trajectories, we used the MD analysis packages<sup>224</sup>. The root mean square deviation (rmsd) was computed using the C $\alpha$  atoms along the simulation, root mean square fluctuations per amino acid residue (rmsf) were calculated using the residue C $\alpha$  atoms along the simulation and the inter-domain distances were measured between the center of mass (COM) of the different domains. Structural images were prepared using Pymol2.5 (Schrodinger Inc.).

*Statistical Analysis.* All the data shown in this paper are reported as mean  $\pm$  SEM. The GraphPad Prism<sup>TM</sup> software was used for statistical analysis. p values < 0.05 were considered significant. For all the datasets, data were analyzed by parametric tests, alpha = 0.05 (one-way and two-way ANOVA with

Tukey's post hoc correction). The statistical analysis performed for each dataset is indicated in figure legend.

### III. RESULTS

#### 1. GENERATION OF A SINGLE VECTOR APPROACH FOR GSDIII BASED ON A FUNCTIONAL, TRUNCATED GDE

Our team previously demonstrated the possibility to correct the muscle and heart phenotype of a GSDIII mouse model (*Agf*<sup>-/-</sup> mice) using a dual vector approach harboring the cytomegalovirus (CMV) promoter<sup>69</sup>. A GDE expression cassette containing the human full-length GDE cDNA and the CMV promoter used in the previous study is close to 6 kb, largely beyond the packaging size limitation of a single rAAV vector. As a first step toward the optimization of the gene therapy approach for GSDIII, we carefully optimized each component of the transgene expression cassette to reduce their size at minimum.

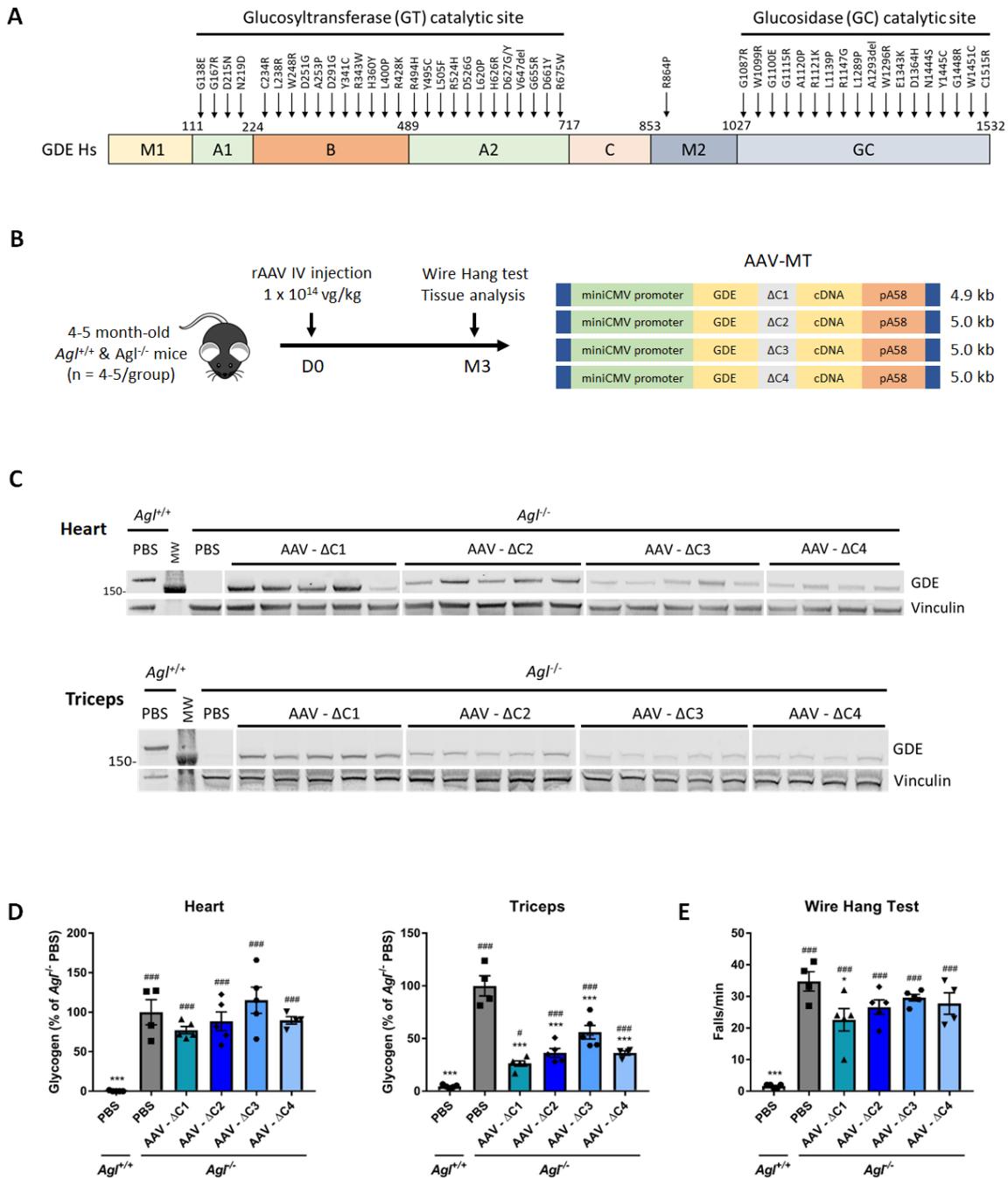
First, we compared *in vivo* the strength of promoters with proven expression in muscle and sizes spanning 334 to 1050 base pairs (bp) such as miniCMV, SPc5-12, tMCK, eSyn, Desmin or CK6<sup>38,225</sup> (**Figure S1A**). Importantly, miniCMV, one of the shortest promoters (351 bp) showed the highest mSEAP expression with similar vector genome copy numbers (VGCN) among all tested promoters (**Figure S1B-C**).

Introns are known to increase the stability of mRNA, the export to the nucleus and ultimately protein production<sup>132</sup>. To assess whether the removal of the intron can impact GDE expression, we compared the efficacy of two similar, oversized expression cassettes with or without SV40 intron in *Agf*<sup>-/-</sup> mice by intramuscular injection (**Figure S2A**). Similar vector transduction and GDE expression as measured by VGCN and Western Blot resulted in a significant decrease of glycogen accumulation in the injected muscle of mice treated with the two vectors (**Figure S2B-E**).

We then tested in *Agf*<sup>-/-</sup> mice the effect on GDE transgene expression of two poly-adenylation (poly-A) signals of 58 (pA58) and 169 (bGh) base pairs. The two expression cassettes were encapsidated in rAAV-MT, a recently described vector with efficient muscle expression and robust liver de-targeting<sup>173</sup> (**Figure S3A**). Three months post-injection, muscle transduction and GDE expression measured by VGCN and Western blot were similar in all evaluated muscles including heart (**Figure S3B-D**). Similar but partial efficacy was observed between the two expression cassettes, with 50-60% reduction of glycogen measured in heart, triceps, quadriceps, soleus, and extensor digitorum longus (EDL) muscles (**Figure S4A, Table S1**). In line with glycogen clearance, histological analysis revealed a limited normalization of the muscle histology (**Figure S4B**).

Taken together, these results allowed us to select the minimal regulatory sequences necessary to efficiently express GDE with a single vector. However, the optimized transgene expression cassette

including the native GDE cDNA had a size of 5.3 kb and showed partial efficacy in *Agf<sup>-/-</sup>* mice thus supporting efforts to reduce the size of the GDE transgene.



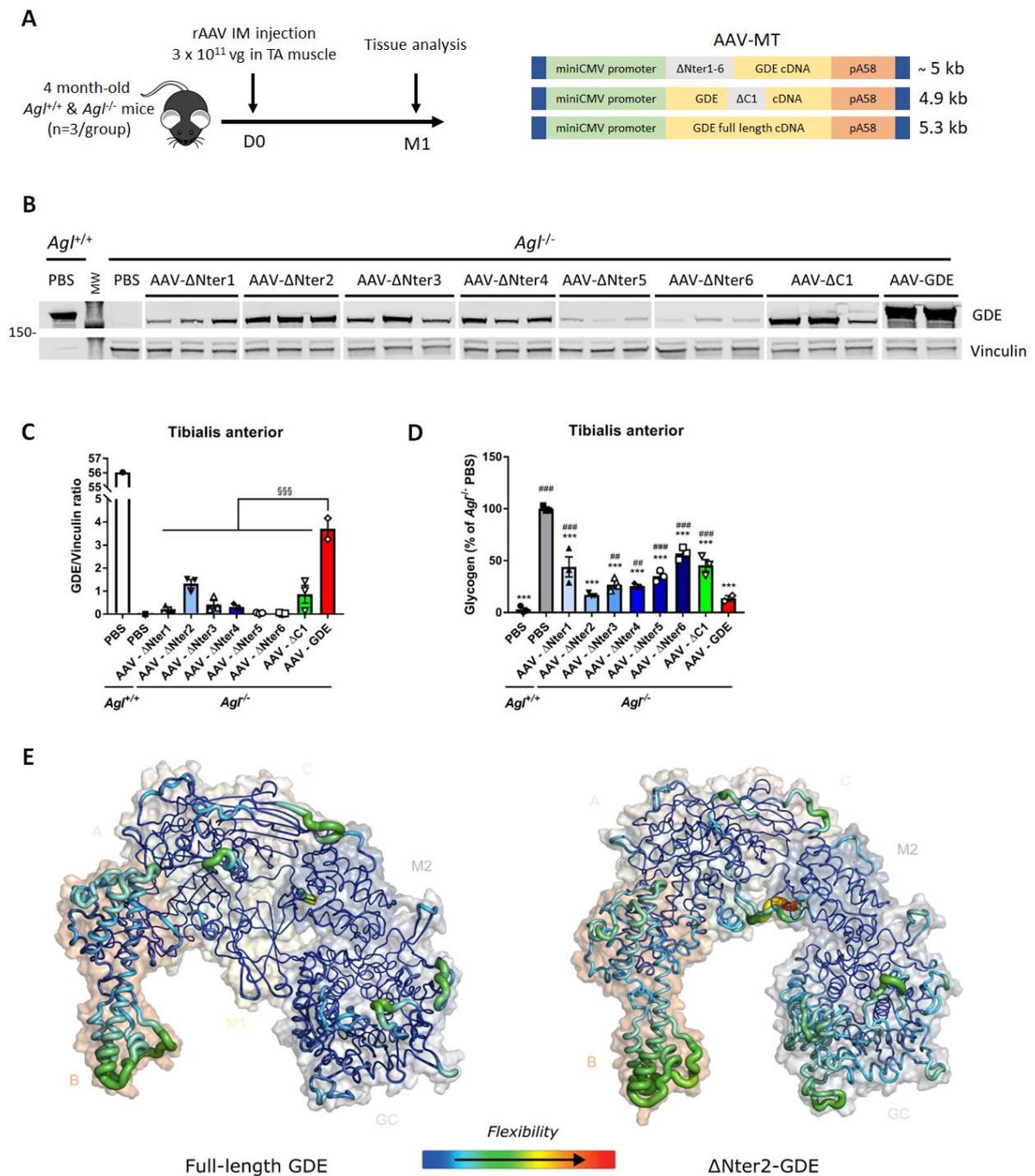
**Figure 23. Generation of truncated GDE polypeptides in the C domain.** (A) Schematic representation of the human GDE cDNA, depicting the different functional domains and the two catalytic sites<sup>196</sup>, as well as the position of missense mutations reported in GSDIII patients<sup>169,177,189,196,226-233</sup>. (B) 4-5-month-old male *Agf<sup>-/-</sup>* mice were injected in the tail vein with an rAAV-MT vector encoding ΔC1-4 GDE mutants, at the dose of  $1 \times 10^{14}$  vg/kg. PBS-injected *Agf<sup>+/+</sup>* or *Agf<sup>-/-</sup>* mice were used as controls. (C) Western Blot analysis of GDE and vinculin expression in heart and triceps at euthanasia, three months after vector injection. (D) Glycogen content measured in heart

and triceps at euthanasia. (E) Wire-hang test expressed as the number of falls per minute, performed 3 months after vectors injection. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. PBS-injected *Agf*<sup>-/-</sup> mice; # $p < 0.05$ , ## $p < 0.01$  or ### $p < 0.001$  vs. PBS-injected *Agf*<sup>+/-</sup> mice;  $n = 4-5$  mice per group). All data are shown as mean  $\pm$ SEM.

Although no three-dimensional structure of human GDE has been experimentally determined so far, the structure of a GDE homolog from *Candida glabrata* (PDB id: 5D06 and 5D0F), showing 38% of sequence identity with the human GDE, was solved using X-ray diffraction<sup>196</sup>. With the emergence of AlphaFold2, an artificial intelligence system developed by DeepMind<sup>221</sup>, high quality models are now available, notably via the AlphaFold2 database<sup>220</sup>, where a model of human GDE can be found (entry AF-P35573-F1, **Figure S5A-B**). This model was retrieved from the database and compared to the crystallographic structure of *C. glabrata* GDE, enabling the identification in human GDE of the different domains described earlier for *C. glabrata*<sup>196</sup> (**Figure S5C**), namely M1, GT (corresponding to the assembly of three subdomains: A, B and C), M2 and GC. The mapping of the missense mutations described in GSDIII patient onto the 3D model revealed the presence of most mutations in the catalytic GT and GC domains<sup>169,177,189,196,226-233</sup>, while almost no mutations were reported in the M1 and C domains (**Figure 23A**). Interestingly, the crystallographic structure of inactive *C. glabrata* GDE in complex with maltooligosaccharides did not reveal the presence of any sugar binding site in these two domains (adapted from PDB id: 5D06 and 5D0F, **Figure S5B**). Based on these considerations, it could be reasonable to assume that these two domains do not play a relevant role in catalysis or ligand binding and could be promising targets to shorten the human GDE protein while retaining activity.

We first generated four GDE mutants with deletion in the C-domain, named  $\Delta C1$  to  $\Delta C4$ , by selecting regions that were less likely to disturb formation of secondary structures, in particular the  $\beta$ -strands majorly composing the C-domain (**Figure S5D, S6A**). All expression cassettes containing the mutants were around 5 kb or lower (**Figure 23B**). Transfection in HEK-293T cells showed that all four  $\Delta C$  mutants were expressed at similar levels (**Figure S6B**). Treatment of *Agf*<sup>-/-</sup> mice with rAAV vectors encoding the  $\Delta C$  mutants showed truncated GDE expression and similar VGCN in heart and skeletal muscle of all injected mice (**Figure 23C and Figure S6C-E**). Importantly, a 50% glycogen clearance was obtained in triceps and quadriceps suggesting residual activity of the  $\Delta C$  truncated GDE enzymes, with higher correction achieved in mice treated with AAV- $\Delta C1$ -GDE (**Figure 23D, Figure S6F and Table S2**). Unexpectedly, no correction of glycogen accumulation was observed in the heart of injected *Agf*<sup>-/-</sup> mice (**Figure 23D**), despite robust tissue transduction and GDE expression. Consistently, a slight improvement of muscle strength assessed by the wire-hang test, a functional test predictive of rAAV gene therapy efficacy<sup>69</sup>, was observed only in mice treated with  $\Delta C1$ -expressing vector (**Figure 23E**). These results suggest that although C-domain-truncated GDE mutants retain some activity, the

deletions may affect GDE function in an organ-specific manner, possibly by altering its regulation in the heart.



**Figure 24. Generation of truncated GDE polypeptides in the N-terminal domain.** (A) 4-month-old female  $Agl^{-/-}$  mice were injected in the left *tibialis anterior* (TA) muscle with an rAAV-MT vector encoding either the human full-length GDE or a truncated GDE, at a dose of 3 x 10<sup>11</sup> vg/mouse. PBS-injected  $Agl^{+/+}$  or  $Agl^{-/-}$  mice were used as controls. (B) Western Blot analysis of GDE and vinculin expression in the injected TA, one month after vector injection. (C) Quantification of GDE expression on Western Blot showed in panel B, expressed as the ratio of the signal of GDE and vinculin bands. (D) Glycogen content in the injected TA measured one month after injection.

(E) Root mean square fluctuations (rmsf) computed on amino acid residue C $\alpha$  along the MD simulation and projected onto full-length human GDE and  $\Delta$ Nter2-GDE predicted structures using AF2. Warmer colors and larger ribbons indicate the most fluctuating amino acid residues. Molecular surfaces are shown in transparency with the different domains labelling for clarity purpose. Statistical analyses were performed by one-way ANOVA (<sup>s</sup>p<0.05, <sup>ss</sup>p<0.01 or <sup>sss</sup>p<0.001 vs. *Agf*<sup>-/-</sup> injected with rAAV-GDE in panel C; \*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. PBS-injected *Agf*<sup>-/-</sup> mice and #p<0.05, ##p<0.01 or ###p<0.001 vs. PBS-injected *Agf*<sup>+/+</sup> mice in panel D; n=3 mice per group). All data are shown as mean  $\pm$ SEM.

We then generated deletion mutants in the M1 domain, named  $\Delta$ Nter1 to  $\Delta$ Nter6, by selecting regions less likely to disturb the formation of  $\beta$ -strands composing the domain (**Figure S7A**). After transfection of HEK-293T cells, we observed expression of all the truncated proteins, although  $\Delta$ Nter1,  $\Delta$ Nter2 and  $\Delta$ Nter4 displayed the higher expression, comparable to the  $\Delta$ C1 mutant (**Figure S7B**). We then evaluated the efficacy of the  $\Delta$ Nter1 to 6 truncated proteins in *Agf*<sup>-/-</sup> mice by intramuscular injection in the tibialis anterior (TA) muscle, in comparison with the  $\Delta$ C1 mutant and the full-length GDE (**Figure 24A**). Western Blot analysis performed in the injected muscle one month after vector injection, revealed lower protein expression in all mutants compared to the full-length GDE despite a similar VGCN among the groups (**Figure 24B-C and Figure S7C**). Of note, the site of truncation had a significant impact on the protein expression level, since  $\Delta$ Nter2-GDE exhibited the highest protein expression while, in line with our *in vitro* data,  $\Delta$ Nter5 and  $\Delta$ Nter6 mutants had very low expression (**Figure 24B-C**). Importantly, all truncated enzymes displayed some degree of activity as reflected by the reduction of glycogen accumulation in the injected muscle with  $\Delta$ Nter2-GDE showing the highest efficacy, comparable to that of full-length GDE (**Figure 24D**). We hypothesized that variations in protein expression could be related to differences in protein stability. Indeed, the RNA levels of GDE in TA was similar in mice treated with rAAV encoding either  $\Delta$ Nter2 or full-length GDE, suggesting that the lower  $\Delta$ Nter2 protein levels were likely due to reduced protein stability (**Figure S7D**). Therefore, to further investigate the impact of the  $\Delta$ Nter2 truncation on the protein stability, we performed molecular dynamics (MD) simulations (250ns) for both the full-length GDE and the  $\Delta$ Nter2-GDE. Analysis of the conformational variations along the simulation time, monitored by the root mean square deviation (rmsd), revealed a slightly higher flexibility of the truncated protein compared to the parental one (**Figure S8A**). More detailed inspection of the amino acid fluctuations over the simulation time, described by the root mean square fluctuation (rmsf), showed main differences in the flexibility of the B and GC domains (**Figure 24E, Figure S8B**). Monitoring of the distances between domains along the simulation (**Figure S8C**), suggested that fluctuations did not appear to be directly related to conformational changes inside the domains but rather due to a relative re-orientation of the domains. Indeed, distance fluctuations between B and GC domains appeared larger in  $\Delta$ Nter2-GDE than in the

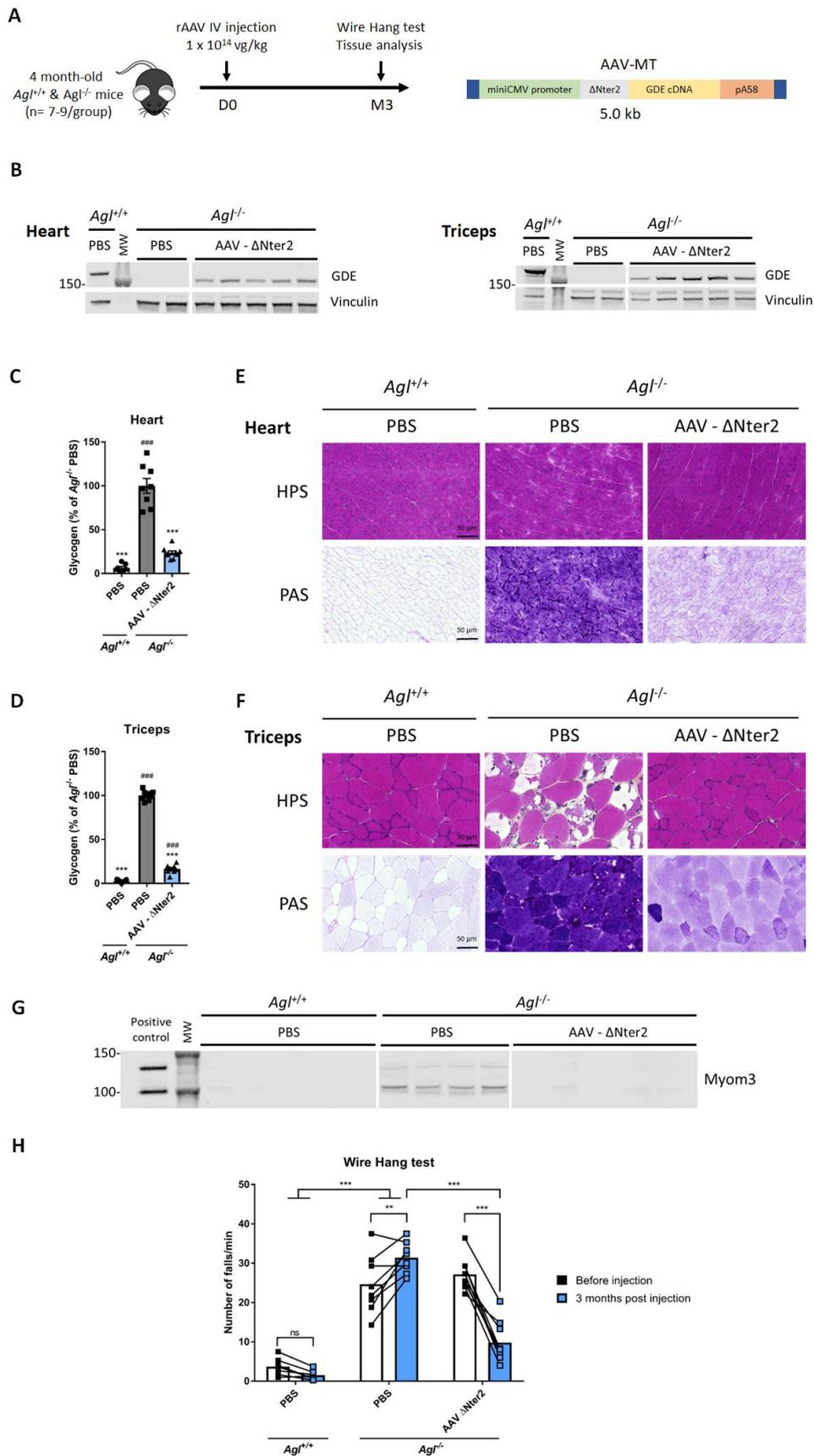
full-length protein (**Figure S8C**), likely resulting from A and M2 domains moving apart due to the absence of M1 domain in  $\Delta$ Nter2-GDE. Indeed, M1 domain is known to maintain the structural integrity of the middle region composed of A, C and M2 domains<sup>196</sup>. Despite of a slightly larger flexibility observed for  $\Delta$ Nter2-GDE compared to the full-length GDE, the structural integrity of the mutant was maintained, and catalytic sites of GC and GT were not affected, indicating that  $\Delta$ Nter2-GDE could be a good candidate for gene therapy.

Oversized rAAV vector production has a large impact on the yields and the quality of the final product, which are critical parameters for the translation to the clinic of rAAV-based gene therapies. To evaluate the impact of the use of the truncated  $\Delta$ Nter2-GDE on the production yields and the quality of the vector, three distinct oversized, expression cassettes expressing human full-length GDE were compared to the 5 kb expression cassette bearing the  $\Delta$ Nter2-GDE (**Figure S9A**). The use of the  $\Delta$ Nter2-GDE cDNA allowed for approximately ten-fold increase of rAAV production yields (**Figure S9B**) while allowing for the efficient encapsidation of a non-truncated genome (**Figure S9C**). Improved yields and genome quality resulted in dramatically improved full to empty particles ratios as measured by analytical ultracentrifugation (AUC), with up to 37% of full particles measured in AAV- $\Delta$ Nter2-GDE (**Figure S9D**).

Taken together, these data indicate that while  $\Delta$ Nter2-GDE has an *in vivo* efficacy similar to the full-size enzyme, it allows to produce rAAV vectors with higher yields and quality thus providing a potential gene therapy candidate for the treatment of GSDIII.

## **2. rAAV ENCODING $\Delta$ NTER2-GDE RESCUES THE CARDIAC AND MUSCLE PHENOTYPE OF *Agf*<sup>-/-</sup> MICE**

Next, we evaluated the efficacy of an rAAV-MT vector expressing the  $\Delta$ Nter2-GDE in *Agf*<sup>-/-</sup> mice via tail vein injection at a dose of  $1 \times 10^{14}$  vg/kg (**Figure 25A**). *Agf*<sup>-/-</sup> mice were treated at 4 months of age when they showed extensive glycogen accumulation in all muscles and functional impairment as measured by wire-hang<sup>69</sup>. Three months after vector injection, VGCN and GDE protein were detected in all the analyzed muscle tissues (**Figure 25B** and **Figure S10A-C**). Glycogen quantification showed normalized glycogen level in heart and in different skeletal muscles (**Figure 25C-D** and **S10D, Table S3**). In line with these data, histological analysis, performed on the same tissues showed complete normalization of the muscle architecture on HPS as well as an important reduction of the glycogen accumulation by PAS staining (**Figure 25E-F** and **Figure S10E-F**). We also evaluated by Western Blot the levels of myomesin3 (Myom3) fragments, a known biomarker of muscle dystrophies<sup>234</sup>. Although GSDIII is not a muscular dystrophy, we found Myom3 elevated in the serum of *Agf*<sup>-/-</sup> mice at baseline and normalized three months after treatment in mice injected with rAAV encoding for  $\Delta$ Nter2-GDE (**Figure 25G** and **Figure S10G,H**).



**Figure 25. Rescue of muscle and heart impairment in the *Agt*<sup>-/-</sup> mouse model with rAAV-ΔNter2-GDE vector.**  
**(A)** 4-month-old male *Agt*<sup>-/-</sup> mice were injected in the tail vein with an rAAV-MT vector encoding ΔNter2-GDE, at

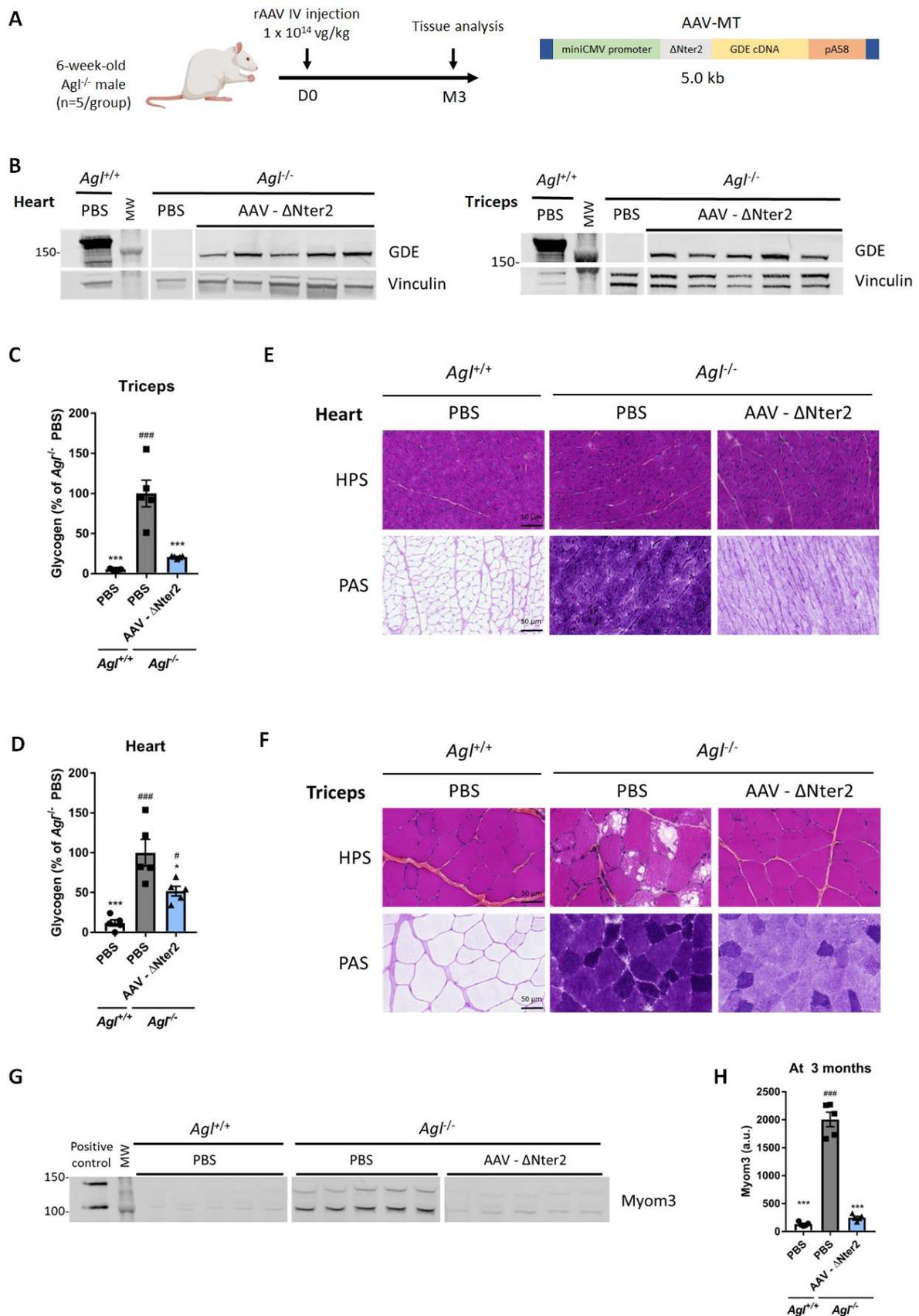
the dose of  $1 \times 10^{14}$  vg/kg. PBS-injected *Agf<sup>+/+</sup>* and *Agf<sup>-/-</sup>* mice were used as controls. (B) Western Blot analysis of GDE and vinculin expression in heart and triceps at euthanasia, three months after vector injection. (C-D) Glycogen content measured in heart (C) and triceps (D) at euthanasia. (E-F) Histological analysis of heart (E) and triceps (F) using HPS and PAS staining. Representative images are shown. (G) Western blot analysis of myomesin3 fragments (Myom3) in serum of mice at euthanasia. Serum from mdx mice was used as positive control. (H) Wire-hang test expressed as number of falls per minute, performed before and 3 months after vector injection. Statistical analyses were performed by one-way ANOVA in C-D and two-way ANOVA in H (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. PBS-injected *Agf<sup>-/-</sup>* mice; # $p < 0.05$ , ## $p < 0.01$  or ### $p < 0.001$  vs. PBS-injected *Agf<sup>+/+</sup>* mice;  $n = 7-9$  mice per group coming from two independent experiments). All data are shown as mean  $\pm$  SEM.

Importantly, treatment with rAAV- $\Delta$ Nter2-GDE rescued muscle strength assessed by wire-hang. Before injection, *Agf<sup>-/-</sup>* animals exhibited a high frequency of falls (26.0 falls/min  $\pm$  1.4), compared to *Agf<sup>+/+</sup>* mice (3.7 falls/min  $\pm$  0.8) (Figure 25H). Three months post-injection rAAV- $\Delta$ Nter2-GDE injected mice showed less frequency of falls compared to the PBS-injected group (9.8 falls/min  $\pm$  1.8 vs. 31.4 falls/min  $\pm$  1.4), almost reaching wild-type levels (Figure 25H). These data clearly demonstrate that treatment with rAAV- $\Delta$ Nter2-GDE reverses the muscle impairment in adult, symptomatic GSDIII mice at both biochemical and functional level.

### 3. rAAV ENCODING $\Delta$ NTER2-GDE CORRECTS GLYCOGEN ACCUMULATION IN AN *Agf<sup>-/-</sup>* RAT MODEL OF GSDIII

To confirm the efficacy of the optimized rAAV- $\Delta$ Nter2-GDE in a larger animal model, we administered the vector to 6-week-old *Agf<sup>-/-</sup>* rats. *Agf<sup>-/-</sup>* rats, treated with  $1 \times 10^{14}$  vg/kg of rAAV- $\Delta$ Nter2-GDE via tail vein injection, were analyzed three months post-injection (Figure 26A). Vector genomes and  $\Delta$ Nter2-GDE protein were detected in all evaluated muscles (Figure 26B, Figure S11A-C). Treatment with rAAV- $\Delta$ Nter2-GDE reduced glycogen content in different skeletal muscles to levels similar to those measured in *Agf<sup>+/+</sup>* rats. (Figure 26C and Figure S11D-F, Table S4). A significant, 50% reduction in glycogen accumulation was observed in the heart (Figure 26D) that was even more evident in the PAS staining (Figure 26E). Histological analysis also showed normalization of muscle architecture, as well as an important glycogen reduction in all skeletal muscles tested (Figure 26F and Figure S11G-I). Analysis of the Myom3 levels before treatment showed moderate elevation in the serum of *Agf<sup>-/-</sup>* rats (Figure S12A-B). Importantly, *Agf<sup>-/-</sup>* rats treated with rAAV- $\Delta$ Nter2-GDE showed reduced Myom3 levels, similar to *Agf<sup>+/+</sup>* rats, three months after vector injection (Figure 26G-H).

In conclusion, the data obtained by the treatment of *Agf<sup>-/-</sup>* rats with an rAAV vector encoding for  $\Delta$ Nter2-GDE confirm those obtained in the mouse model of the disease and further support the clinical translation of rAAV- $\Delta$ Nter2-GDE to treat the muscle disease in GSDIII patients.

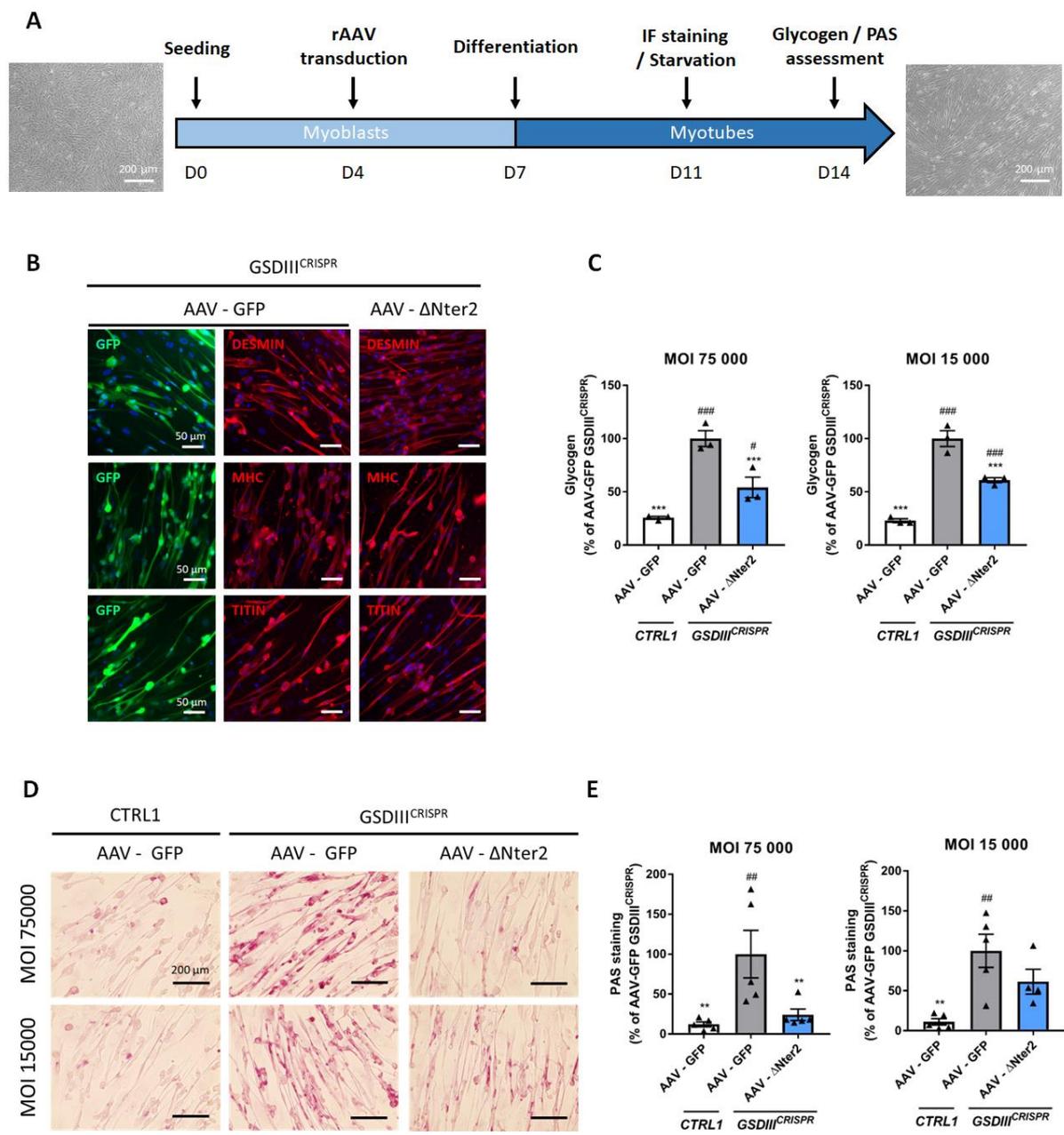


**Figure 26. Correction of muscle and heart impairment in the *Agl*<sup>-/-</sup> rat model with rAAV-ΔNter2-GDE vector. (A)** 6-week-old male *Agl*<sup>-/-</sup> rats were injected in the tail vein with an rAAV-MT vector encoding the ΔNter2-GDE, at

the dose of  $1 \times 10^{14}$  vg/kg. PBS-injected *Agf*<sup>-/-</sup> and *Agf*<sup>+/+</sup> rats were used as controls. **(B)** Western Blot analysis of GDE and vinculin expression in heart and triceps at euthanasia, three months after vector injection. **(C-D)** Glycogen content measured in triceps **(C)** and coeur **(D)** at euthanasia. **(E-F)** Histological analysis of heart **(E)** and triceps **(F)** with HPS and PAS staining. Representative images are shown. **(G)** Western blot analysis of myomesin 3 fragments (Myom3) in the serum of rats 3 months post-injection. Serum from mdx mice was used as positive control **(H)** Quantification of Myom3 expression on the Western Blot showed in Panel G (arbitrary units). Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. PBS-injected *Agf*<sup>-/-</sup> rats; #p<0.05, ##p<0.01 or ###p<0.001 vs. PBS-injected *Agf*<sup>+/+</sup> rats; n=5 rats per group). All data are shown as mean ±SEM.

#### **4. ΔNter2-GDE CLEARS GLYCOGEN IN A HUMAN SKELETAL MUSCLE CELL MODEL OF GSDIII WITHOUT ANY SPECIFIC TOXICITY**

To evaluate the activity of the truncated ΔNter2-GDE in a human pathological context, we took advantage of a recently reported *in vitro* human skeletal muscle model of GSDIII derived from human induced pluripotent stem cells (hiPSCs) edited by CRISPR/Cas9 technology (*GSDIII*<sup>CRISPR</sup>)<sup>218</sup>. Skeletal muscle cells derived from the isogenic control hiPSC line (CTRL1) were used as unaffected control<sup>218</sup>. We produced an rAAV vector expressing ΔNter2-GDE or GFP suitable for *in vitro* use and we treated *GSDIII*<sup>CRISPR</sup> and CTRL1 hiPSC-derived skeletal myoblasts (skMb) following a protocol recently reported<sup>218</sup> (**Figure 27A**). After transduction with rAAV, skMb were differentiated into skeletal myotubes (skMt). Transduction with rAAV expressing GFP or ΔNter2-GDE did not alter the differentiation of skMb, that showed similar expression of skeletal myogenic markers by immunostaining analysis (**Figure 27B and Figure S13A**). Cell viability was reduced in *GSDIII*<sup>CRISPR</sup> skMt compared to CTRL1 skMt (**Figure S13B**). Similar viability was observed between GFP- and ΔNter2-GDE-transduced cells, suggesting an absence of specific toxicity of the mutant GDE in this human cellular model of GSDIII (**Figure S13B**). *GSDIII*<sup>CRISPR</sup> skMt transduced with rAAV expressing ΔNter2-GDE exhibited a significant reduction of glycogen content when compared to *GSDIII*<sup>CRISPR</sup> skMt transduced with the control vector (**Figure 27C**). PAS staining performed on transduced-skMt confirmed these results (**Figure 27D-E**). These data demonstrate that the truncated ΔNter2-GDE clears glycogen not only in mouse and rat muscles, but also in a human skeletal muscle model of GSDIII without any specific toxicity.



**Figure 27. Clearance of glycogen in human induced pluripotent stem cells (hiPSC)-derived muscle cells treated with rAAV-ΔNter2-GDE.** (A) Transduction protocol of GSDIII<sup>CRISPR</sup> hiPSC-derived skMb and isogenic controls with an rAAV-LK03 vector expressing either GFP or ΔNter2-GDE at a MOI of either 15 000 or 75 000. (B) Expression of specific myogenic markers (desmin, myosin heavy chain (MHC) and titin) and GFP by immunofluorescence analysis, after rAAV transduction at MOI 75 000 in GSDIII<sup>CRISPR</sup> hiPSC-derived skMt. Representative images are shown (C) Glycogen content normalized to the number of viable cells measured in GSDIII<sup>CRISPR</sup> hiPSC-derived skMt and isogenic controls (CTRL1) transduced as described above. (D) PAS staining on GSDIII<sup>CRISPR</sup> hiPSC-derived skMt and isogenic controls (CTRL1) transduced as described above. Representative images are shown. (E) Quantification of PAS staining. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. GSDIII<sup>CRISPR</sup> cells transduced with rAAV-GFP; #p<0.05, ##p<0.01 or ###p<0.001 vs. CTRL1 cells, three

replicates for experiments in panel B-C and 5 replicates for experiments in panel D-E). All data are shown as mean  $\pm$ SEM.

#### IV. DISCUSSION

In GSDIII, inactivating mutations on the *AGL* gene result in the absence of a functional GDE and accumulation of limit dextrin in liver, heart, and skeletal muscles. The hepatic and metabolic phenotypes (hepatomegaly and hypoglycemia) appear early, in patients of around one year of age. Moreover, recent publications report that children exhibit significant liver fibrosis as early as one year of age<sup>177,189</sup>. Although hypoglycemia and liver enzymes usually improve with age, the liver disease progresses and adult patients present with cirrhosis and tumors and may require liver transplantation<sup>177,190</sup>. Nonetheless, the major disease burden at adulthood remains the heart and muscle phenotype<sup>177,190</sup>. Patients benefit from dietary treatment to control the hypoglycemic episodes<sup>176,193</sup>, but there is no curative treatment for GSDIII. Gene therapy with rAAV vectors appears as an attractive option to correct the muscle and heart phenotype of GSDIII<sup>143</sup>, but the size of the GDE cDNA (4.6 kb) along with the required regulatory sequences hampers its efficient encapsidation in a single rAAV. Two different strategies have been published, each with their own drawbacks: a dual vector approach<sup>69</sup> and a pullulanase-based approach<sup>42,209</sup>. While the dual vector approach was efficient for muscle and heart correction<sup>69</sup>, it required two vectors and might expose patients to toxicities. The second strategy relied on a single vector encoding the bacterial enzyme pullulanase and induces immune response toward the transgene<sup>42</sup>. In the present work, we used rational engineering to generate a truncated version of the human GDE efficiently packaged in a single rAAV vector. Given the recent reports of liver-associated toxicity after high-dose gene therapy for neuromuscular disease<sup>86,87</sup> and the presence of an underlying liver disease in GSDIII patients<sup>177,189,190</sup>, we took advantage of a recently developed AAV capsid with improved muscle targeting and liver de-targeting<sup>173</sup>. Such an approach may provide additional safety while ensuring an efficient muscle targeting.

A translationally viable strategy to achieve efficient rAAV packaging and broader muscle cell transduction with large transgenes is the generation of shorter proteins<sup>164,216</sup>. A similar approach has also been successfully used for liver disease<sup>77</sup>. GDE is a complex enzyme that involves two distinct enzymatic activities in a single polypeptide chain<sup>196</sup>. As such, even slight modifications of the sequence may induce dramatic reduction of the activity. To reduce the size of GDE, we took advantage of the large number of reported missense variants in GSDIII patients. We identified two putative regions for which we derived different truncated candidates. Intriguingly, deletion mutants within the C-domain displayed activity in skeletal muscles, but not in the heart. Further characterization of the underlying mechanisms is ongoing although the working hypothesis is that the C-domain participate in the

regulation of GDE activity in a tissue-specific manner. Differently from the C-domain mutants, the N-terminal truncated  $\Delta$ Nter2-GDE was active and decreased glycogen accumulation in both heart and skeletal muscle. This truncation was predicted by molecular modelling to have limited impact on the structural integrity of the protein and its catalytic sites. However, molecular dynamics simulations indicated a higher flexibility of  $\Delta$ Nter2-GDE compared to the full-length GDE which is coherent with the lower protein stability observed experimentally. The use of MD simulations is therefore a valuable tool for the design of truncated proteins and enable AAV gene therapy in diseases that involve large transgenes.

The use of  $\Delta$ Nter2-GDE allowed us to generate a potent, high-quality product for GSDIII. rAAV- $\Delta$ Nter2-GDE corrected muscle and heart phenotype in both mouse and rat models of GSDIII with extensive glycogen clearance, reflecting broader cell transduction. Treatment of *Agf*<sup>-/-</sup> mice with rAAV- $\Delta$ Nter2-GDE vector promoted the recovery of muscle strength 3 months post-injection. This suggests that in GSDIII, the myopathy, which displays low fibrosis and muscle fibers renewal in contrast to what has been observed in muscular dystrophies<sup>158,169</sup>, can be reversed at least in rodent models.

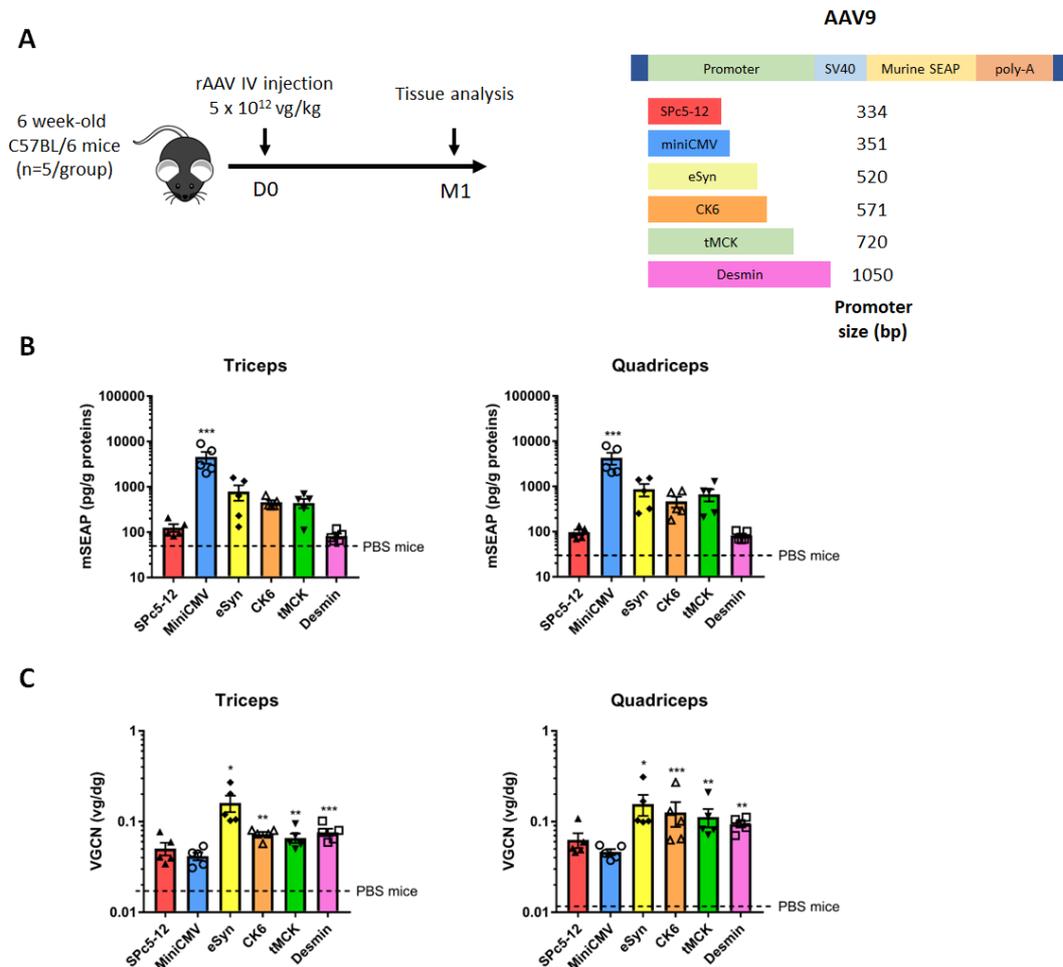
A recently developed human skeletal muscle model of GSDIII derived from hiPSCs edited by CRISPR/Cas9 technology<sup>218</sup> offers a unique opportunity to evaluate the efficacy and safety of the expression of the truncated  $\Delta$ Nter2-GDE in a human pathological context. These cells, in contrast to explanted cells, have self-renewal capacities and the possibility of differentiation in a variety of cell types<sup>235</sup>. In the context of GSDIII, the limited access to liver and muscle biopsies makes *in vitro* human models of the affected tissues essential to evaluate safety and efficacy of therapeutic approaches. In our study, we demonstrated that  $\Delta$ Nter2-GDE clears the accumulated glycogen, suggesting that the truncation on the N-terminal region does not affect the enzyme function and regulation in a human cellular context. Furthermore, although the transduction and starvation protocol required for glycogen accumulation in this *in vitro* system induced some cellular toxicity, cell viability and myogenic markers were similar between rAAV-GFP- and rAAV- $\Delta$ Nter2-transduced cells, suggesting the absence of specific toxicity related to the expression of the truncated mutant in human muscle cells. These data are extremely valuable in the prospect of translating this gene therapy approach into humans.

Beyond the correction of the muscle disease, further optimizations could improve single rAAV vector gene transfer for GSDIII. Targeting of the liver may prove complex due to the underlying liver disease<sup>177,189,190</sup> and the complexity of developing AAV capsids able to target the liver without overloading at the doses used for muscle gene transfer. However, the use of a muscle and liver tropic capsid combined with a short promoter able to achieve both liver and muscle transgene expression could potentially allow complete correction of GSDIII phenotype. Indeed, tandem promoters with both efficient liver and muscle expression have been described, but are too large to be efficiently encapsidated along with the  $\Delta$ Nter2-truncated GDE in a single rAAV vector<sup>41,42</sup>. Combination of smaller

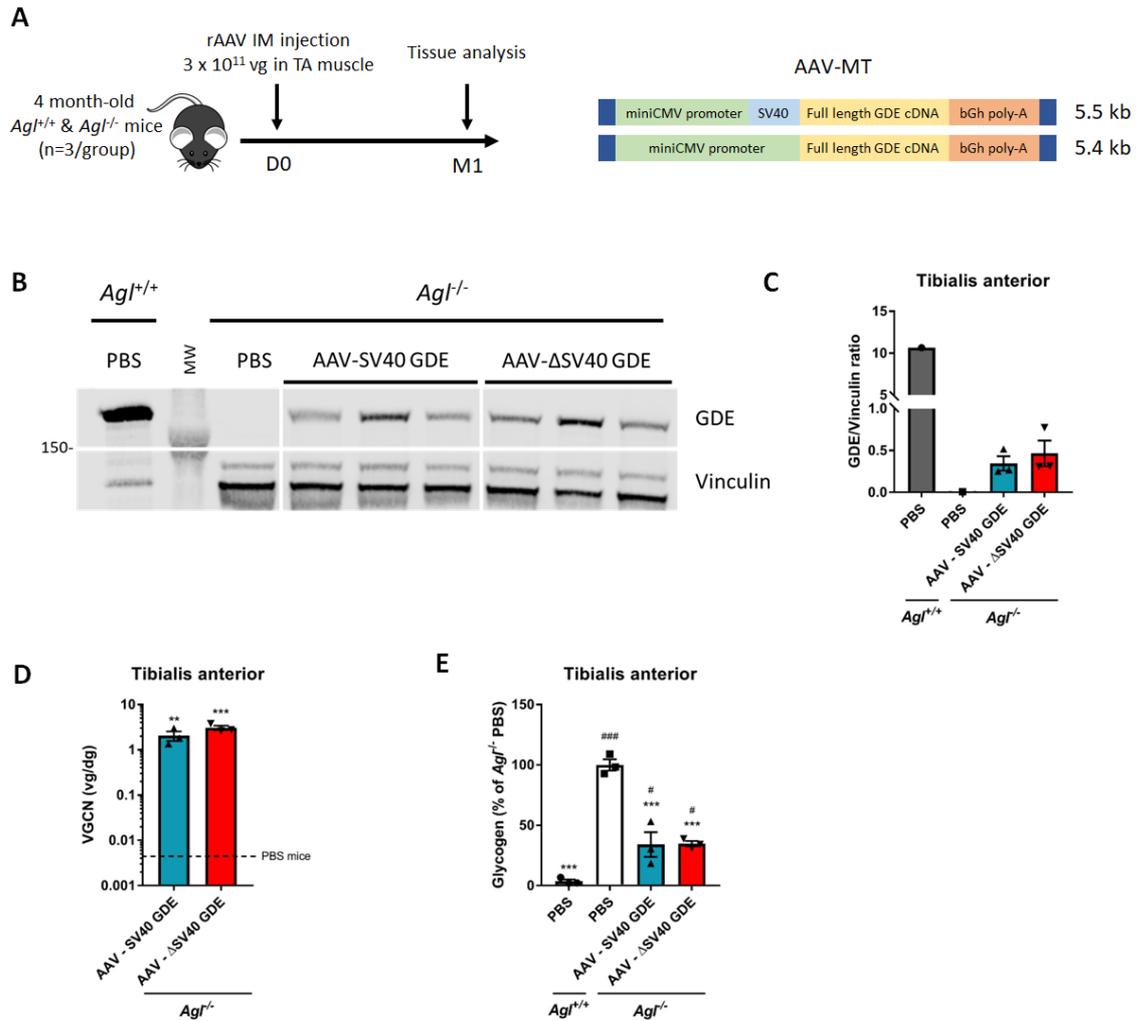
promoters and novel rAAV capsids with a liver-muscle targeting tailored to GSDIII may be used in the next future to improve safety and efficacy of the approach.

To conclude, our work provides proof-of-concept of the use of a single rAAV vector expressing a truncated form of GDE for the rescue of the muscle and heart impairment in two GSDIII rodent models as well as in a novel human muscle cellular model of GSDIII, and support the clinical translation of this approach to provide a one-shot, definitive treatment for this burdensome disease.

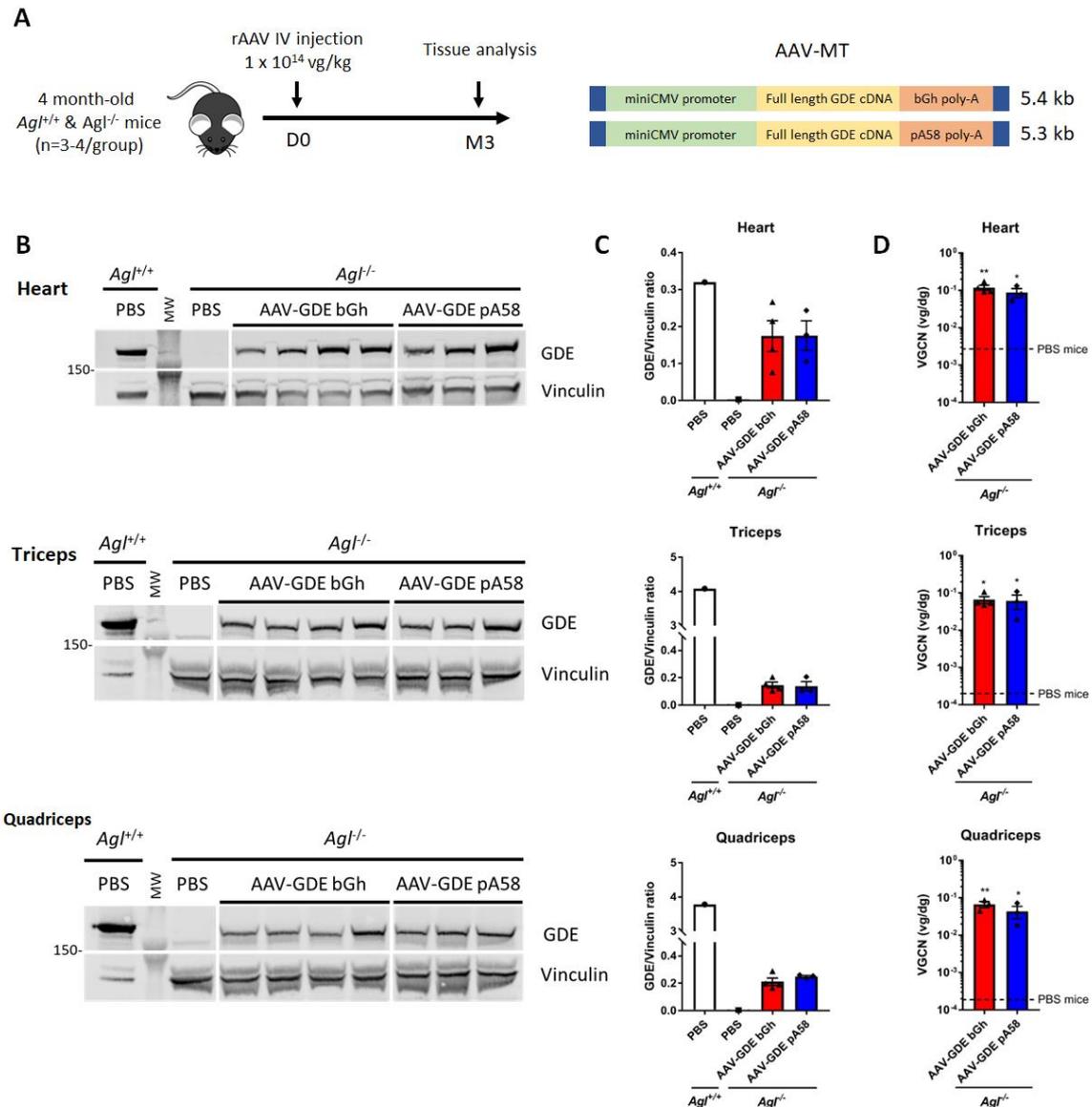
## V. SUPPLEMENTARY MATERIALS



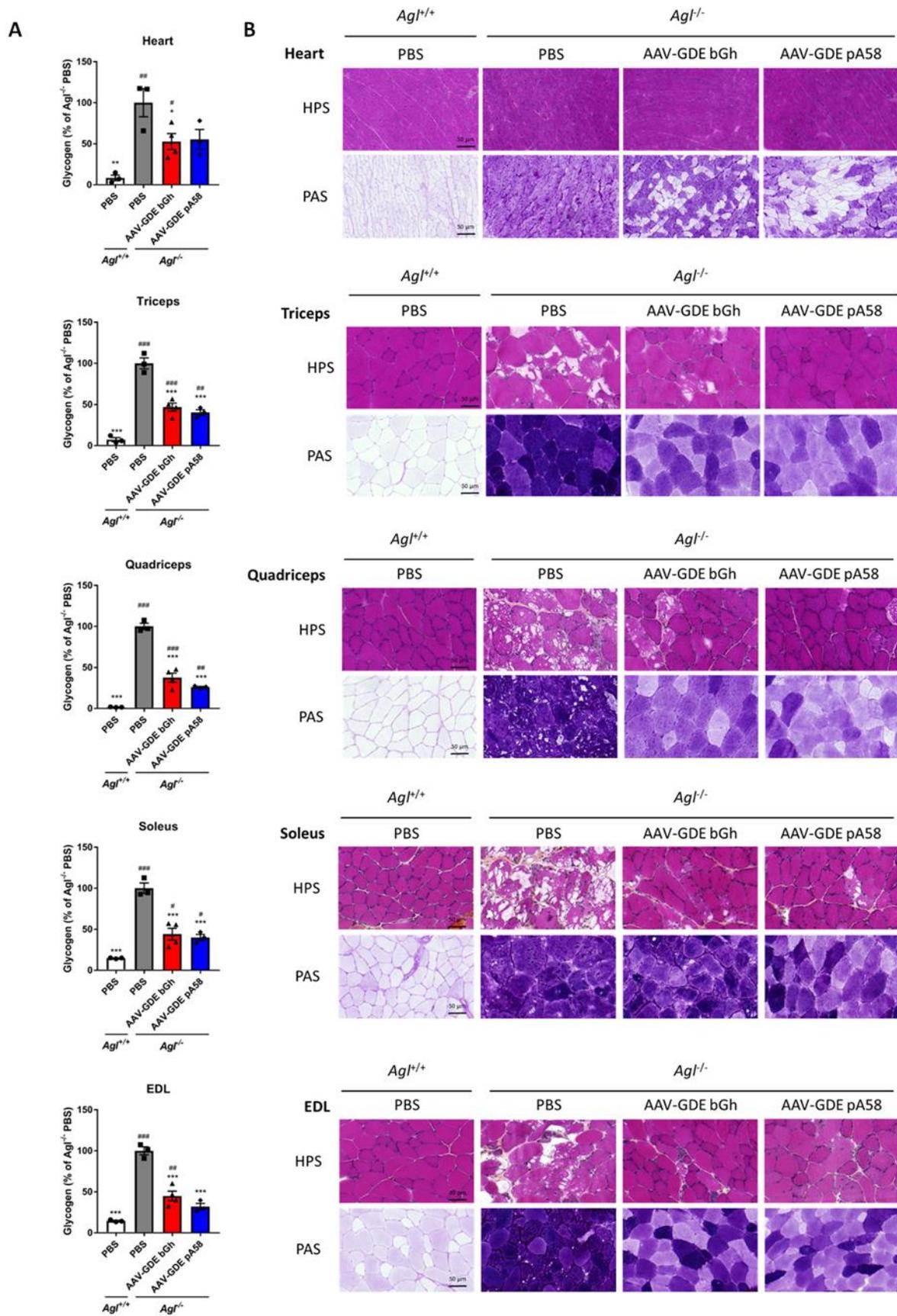
**Figure S1. Selection of a short promoter able to efficiently express a reporter gene in muscle.** (A) 6-week-old male C57BL/6 mice were injected in the tail vein with PBS or with 5 x 10<sup>12</sup> vg/kg of rAAV9 encoding the mSEAP gene reporter under the control of various promoters including SpC512, miniCMV, eSyn, CK6, truncated muscle creatine kinase (tMCK) or Desmin<sup>38,225</sup>. (B-C) mSEAP activity quantification (B) and vector genome copy number (VGCN) (C) in triceps and quadriceps one month after injection. The dashed line represents the background level measured in PBS-injected mice. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. PBS-injected mice, n=5 mice per group). All data are shown as mean ±SEM.



**Figure S2. Evaluation of the requirement for an intron in rAAV vectors expressing GDE.** (A) 4-month-old female *Agl*<sup>-/-</sup> mice were injected in the left *tibialis anterior* (TA) muscle with an rAAV-MT vector encoding the human full-length GDE with or without SV40 intron, at a dose of 3 x 10<sup>11</sup> vg/mouse. PBS-injected *Agl*<sup>-/-</sup> and *Agl*<sup>+/+</sup> mice were used as controls. (B) Western Blot analysis of GDE and vinculin expression in TA muscle at euthanasia one month after injection. (C) Quantification of GDE expression on Western Blot in Figure S2B, expressed as the ratio of the signal of GDE and vinculin bands. (D) VGCN in the left TA. The dashed line represents the background level measured in PBS-injected mice. (E) Glycogen content in the injected TA at euthanasia one month after injection. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. PBS-injected *Agl*<sup>-/-</sup> mice; #p<0.05, ##p<0.01 or ###p<0.001 vs. PBS-injected *Agl*<sup>+/+</sup>, n=3 mice per group). All data are shown as mean ±SEM.

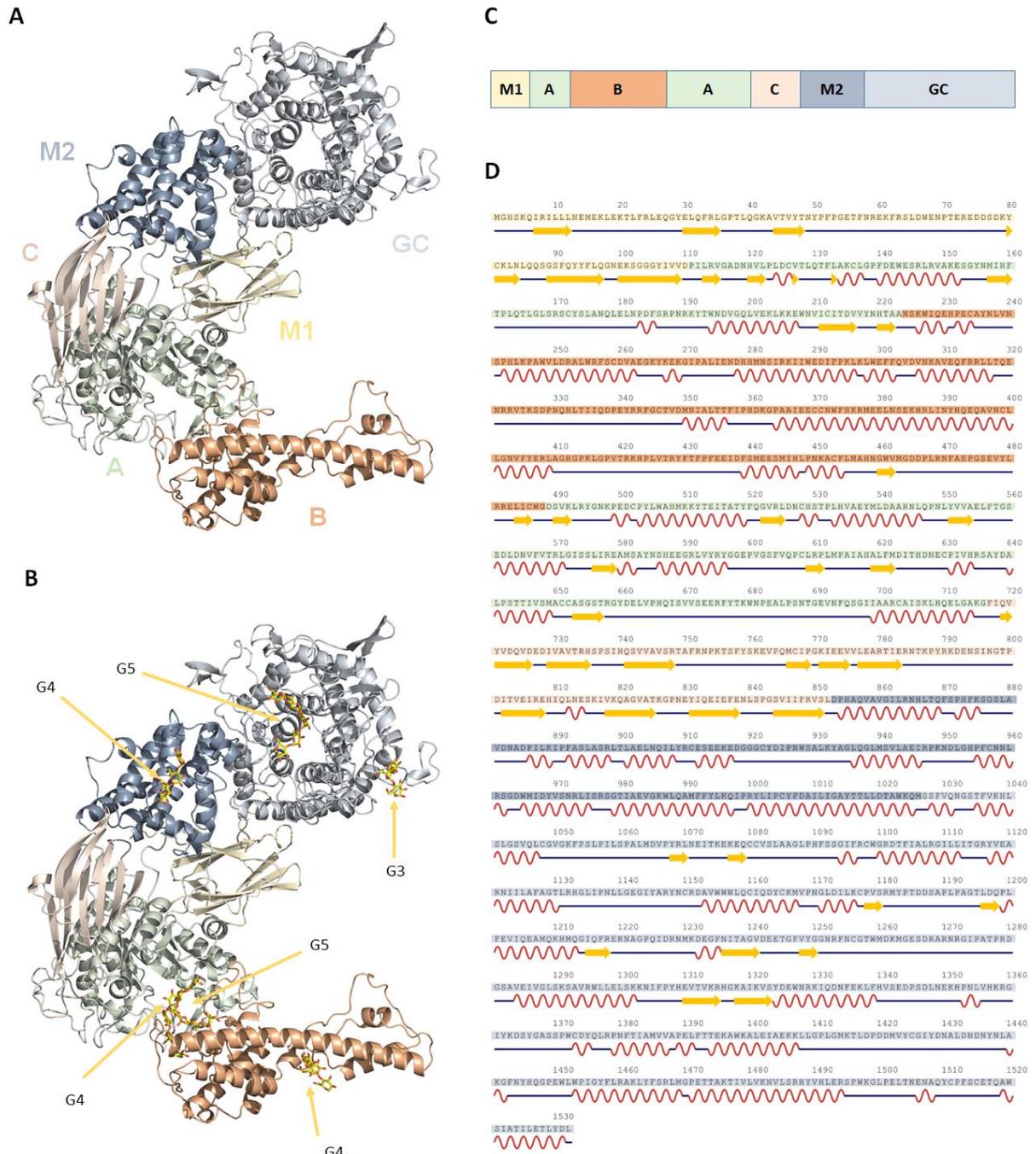


**Figure S3. Expression of the full-length GDE in heart and muscle of *Agl*<sup>-/-</sup> mice using an optimized rAAV vector.** (A) 4-month-old *Agl*<sup>-/-</sup> male mice were injected in the tail vein with an rAAV-MT vector encoding human full-length GDE with either the bGh or the pA58 poly-A signal, at a dose of 1 x 10<sup>14</sup> vg/kg. PBS-injected *Agl*<sup>-/-</sup> and *Agl*<sup>+/+</sup> mice were used as controls. (B) Western Blot analysis of GDE and Vinculin in the heart, triceps and quadriceps at euthanasia 3 months after injection. (C) Quantification of GDE expression on Western Blot in Figure S3B, expressed as the ratio of the signal of GDE and vinculin bands. (D) VGCCN in heart, triceps and quadriceps of injected mice. The dashed line represents the background level measured in PBS-injected mice. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. PBS-injected *Agl*<sup>-/-</sup> mice; #p<0.05, ##p<0.01 or ###p<0.001 vs. PBS-injected *Agl*<sup>+/+</sup> mice, n=3-4 mice per group). All data are shown as mean ±SEM.

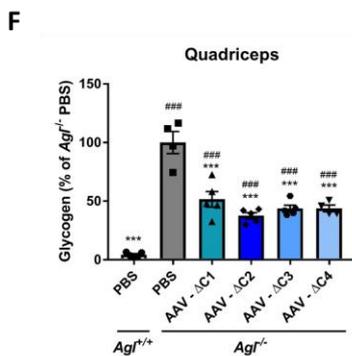
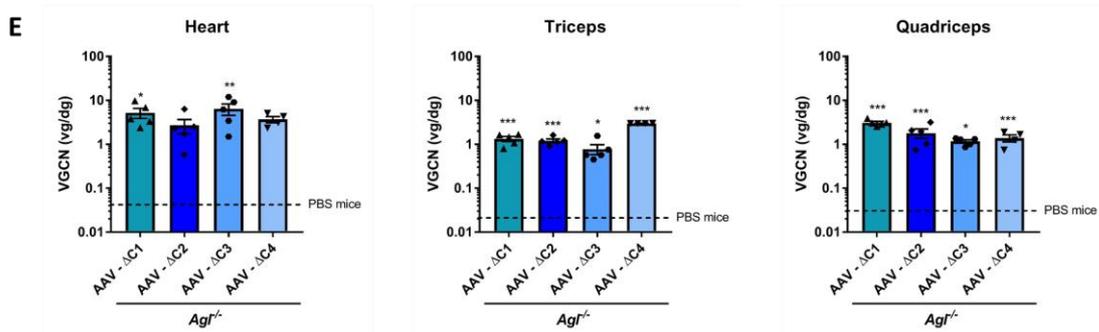
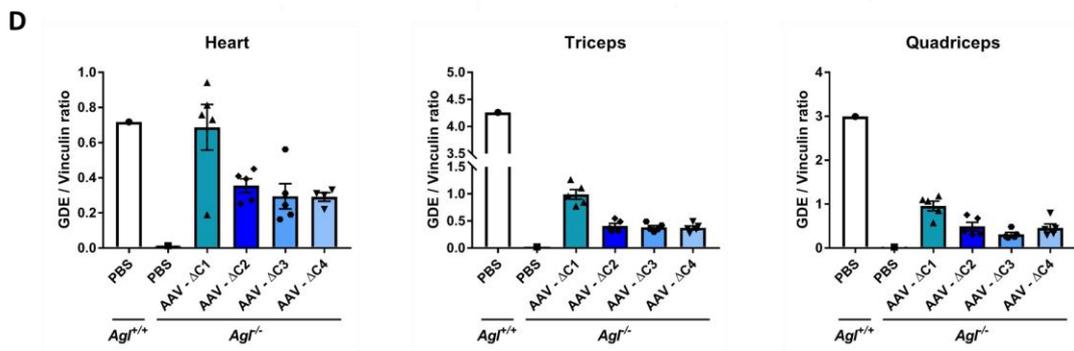
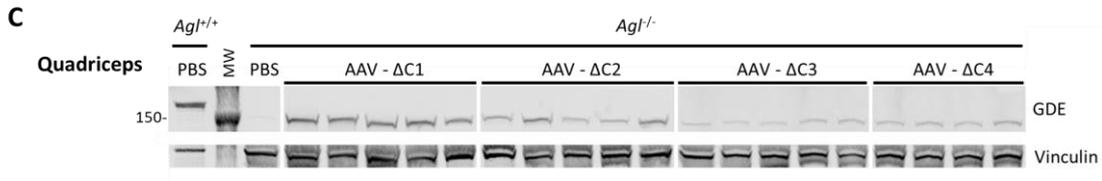
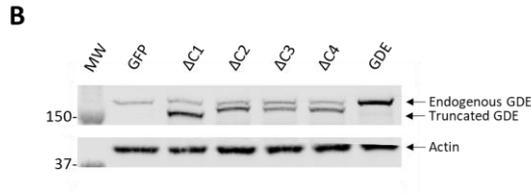
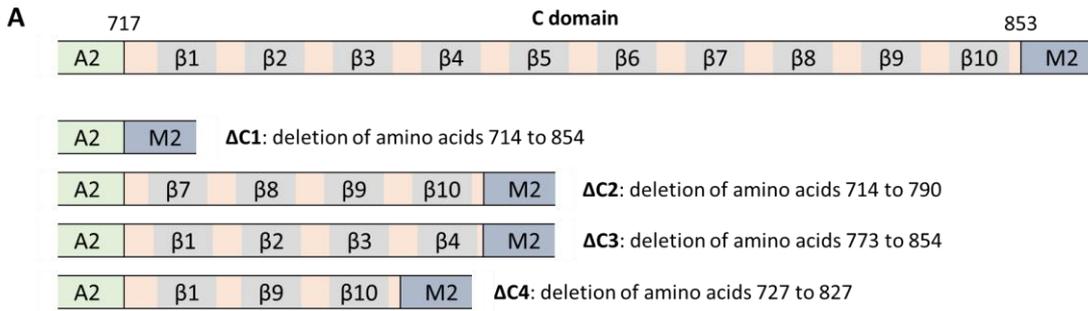


**Figure S4. Partial correction of muscle glycogen accumulation in *Agt*<sup>-/-</sup> mice treated with a rAAV vector encoding full-length GDE. (A) Glycogen content in heart, triceps, quadriceps, soleus and EDL muscles at**

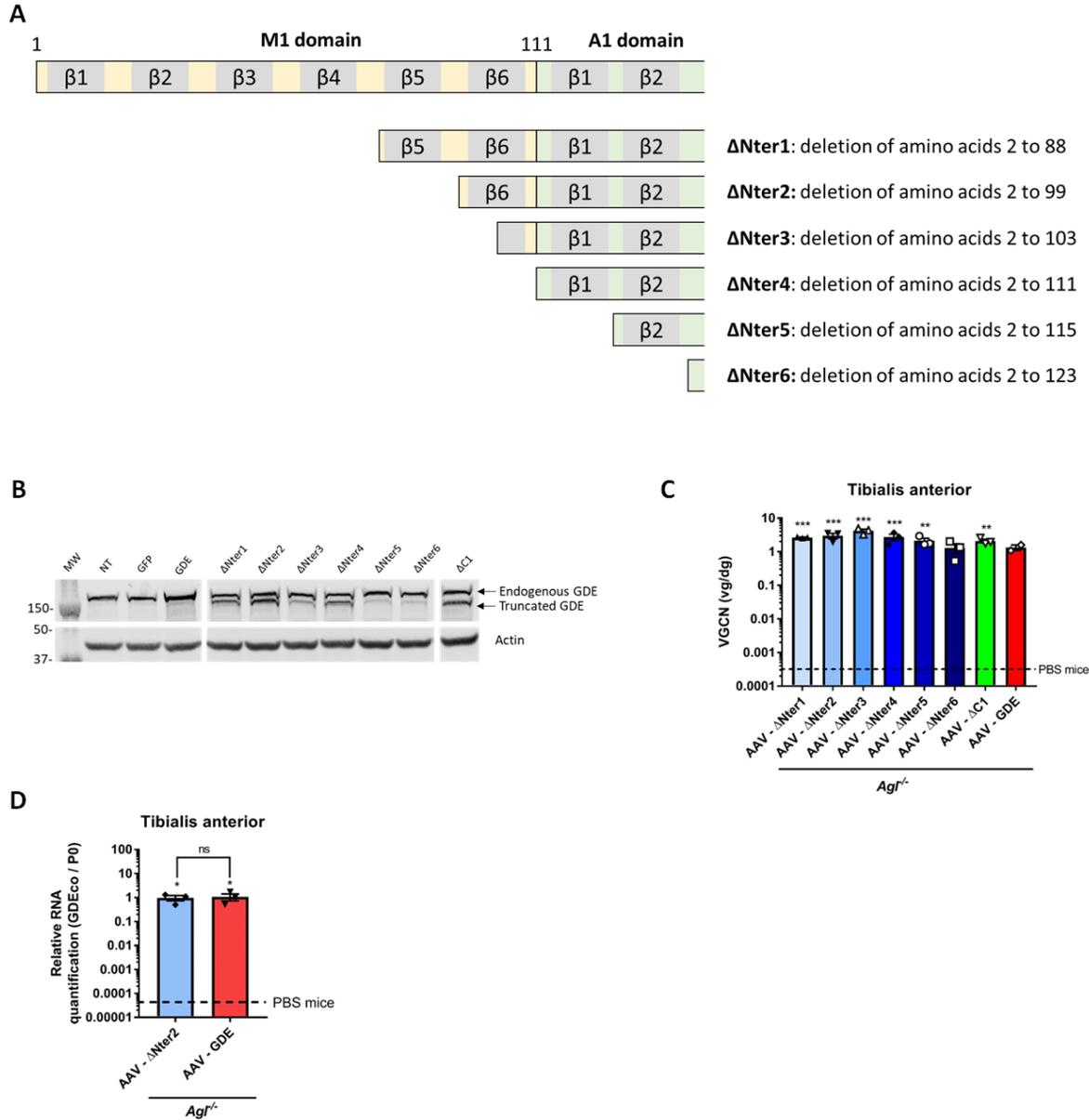
euthanasia 3 months after injection. **(B)** Histological analysis in heart, triceps, quadriceps, soleus and EDL using HPS and PAS staining. Representative images are shown. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. PBS-injected *Agf<sup>-/-</sup>* mice; # $p < 0.05$ , ## $p < 0.01$  or ### $p < 0.001$  vs. PBS-injected *Agf<sup>+/+</sup>* mice, n=3-4 mice per group). All data are shown as mean  $\pm$ SEM.



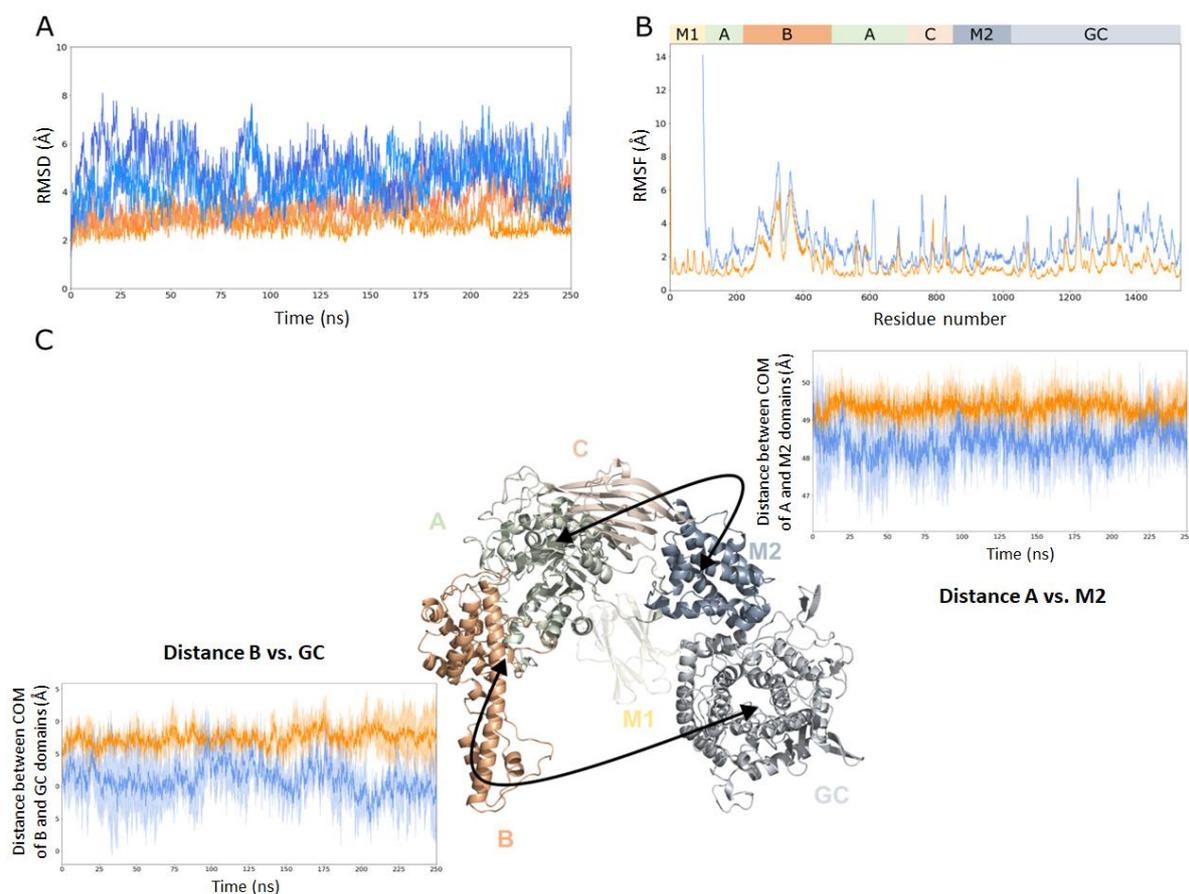
**Figure S5. Structural organization of the human GDE protein.** (A) human GDE model retrieved from the AlphaFold2 database and colored according to the protein domain (B) Mapping onto human GDE of the sugar ligands bound in *C. glabrata* GDE crystallographic structure (PDB id: 5D0F). G3: maltotriose, G4: maltotetraose and G5: maltopentaose. (C) Simplified representation of the sequence of full-length GDE colored according to its domains (D) Amino acid sequence and secondary structures of human GDE (wave:  $\alpha$ -helix; arrow:  $\beta$ -strand), with amino acids colored according to the domain they belong to (same color code as in panel A).



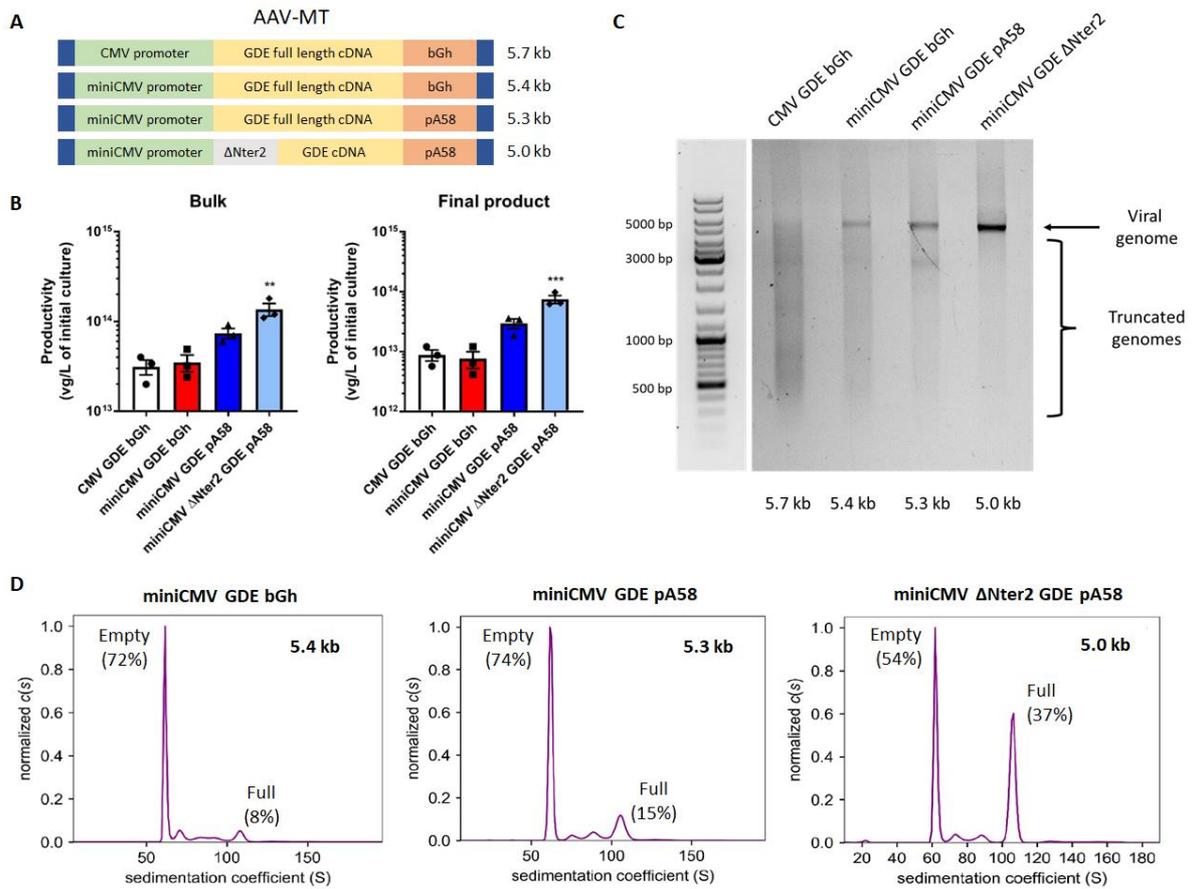
**Figure S6. Generation of functional truncated GDE polypeptides in the C domain.** (A) Positions of the truncation sites in the C-domain of the human GDE for mutants  $\Delta C1$  to  $\Delta C4$ . (B) Western blot analysis of GDE and actin in HEK 293T transiently transfected with plasmids expressing green fluorescent protein (GFP), full-length GDE or  $\Delta C1$  to  $\Delta C4$  GDE deletion mutants. (C) Western blot analysis of GDE and Vinculin in the quadriceps at euthanasia three months after injection. (D) Quantification of GDE expression on Western Blot showed in Figure 23C and S6C, expressed as the ratio of the signal of GDE and vinculin bands. (E) VGCN in heart, triceps and quadriceps of injected mice. The dashed line represents the background level measured in PBS-injected mice. (F) Glycogen content in quadriceps at euthanasia 3 months after injection. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. PBS-injected  $Ag1^{-/-}$  mice; # $p < 0.05$ , ## $p < 0.01$  or ### $p < 0.001$  vs. PBS-injected  $Ag1^{+/+}$  mice,  $n=4-5$  mice per group). All data are shown as mean  $\pm$ SEM.



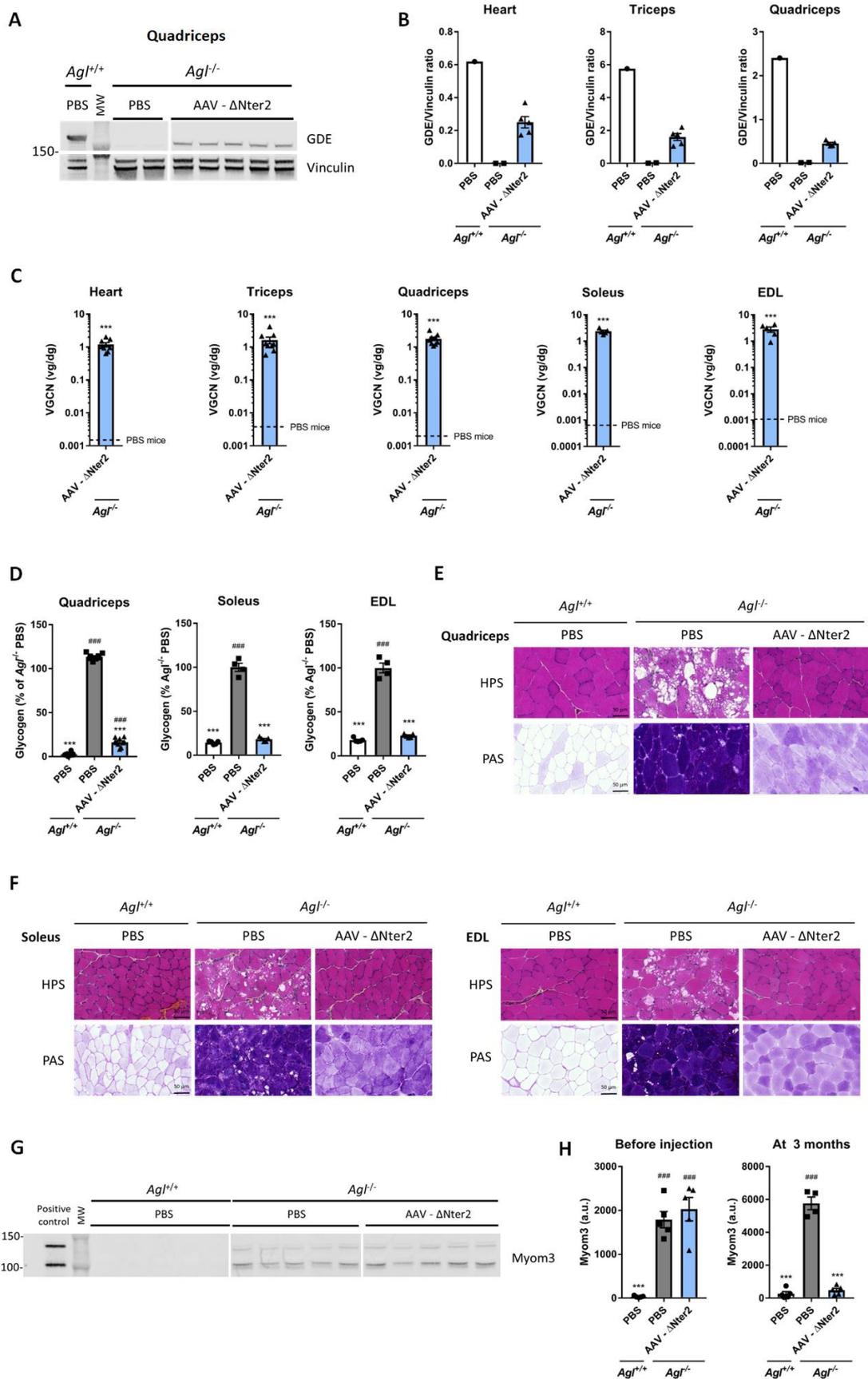
**Figure S7. Generation of functional truncated GDE polypeptides in the N-terminal domain.** (A) Positions of the truncation sites in the N-terminal domain of the human GDE for mutants  $\Delta$ Nter1 to  $\Delta$ Nter6. (B) Western blot analysis of GDE and Actin in HEK 293T transiently transfected with plasmids expressing GFP, full-length GDE, a  $\Delta$ Nter1 to  $\Delta$ Nter6 or  $\Delta$ C1 GDE deletion mutants. NT= not transfected. (C) VGCN measured in *tibialis anterior* of mice injected with rAAV encoding the full-length or truncated GDE cDNA. The dashed line represents the background levels measured in PBS-injected mice. (D) RNA expression levels of full-length or  $\Delta$ Nter2-truncated GDE. The dashed line represents the background levels measured in PBS-injected mice. Statistical analyses were performed by one-way ANOVA (\* $p$ <0.05, \*\* $p$ <0.01 or \*\*\* $p$ <0.001 vs. PBS-injected  $Agf^{-/-}$  mice,  $n=3$  mice per group). All data are shown as mean  $\pm$ SEM.



**Figure S8. Molecular dynamics (MD) of human GDE and  $\Delta$ Nter2-GDE showing conformational changes occurring along the simulation.** (A) Root mean square deviation (rmsd) computed on C $\alpha$  atoms along the 250 ns of the simulation. In orange are shown the simulation replicates of full-length human GDE and in blue those of  $\Delta$ Nter2-GDE. (B) Root mean square fluctuations (rmsf) computed on amino acid residue C $\alpha$  atoms along the 250 ns simulations and averaged over the triplicates of full-length human GDE (in orange) and  $\Delta$ Nter2-GDE (in blue). (C) 3D-model representation of the full-length human GDE, with in transparency the M1 domain that has been truncated in  $\Delta$ Nter2-GDE. The simulation of the distances between the A and M2 domains and between the B and GC domains depicted by arrows, are shown in the top right graph and the bottom left graph, respectively. Average distances were calculated on simulation triplicates at each time point between centers of mass (COM) of corresponding domains for the full-length human GDE (in orange) and  $\Delta$ Nter2-GDE (in blue). The standard deviation is shown as a semi-transparent area.

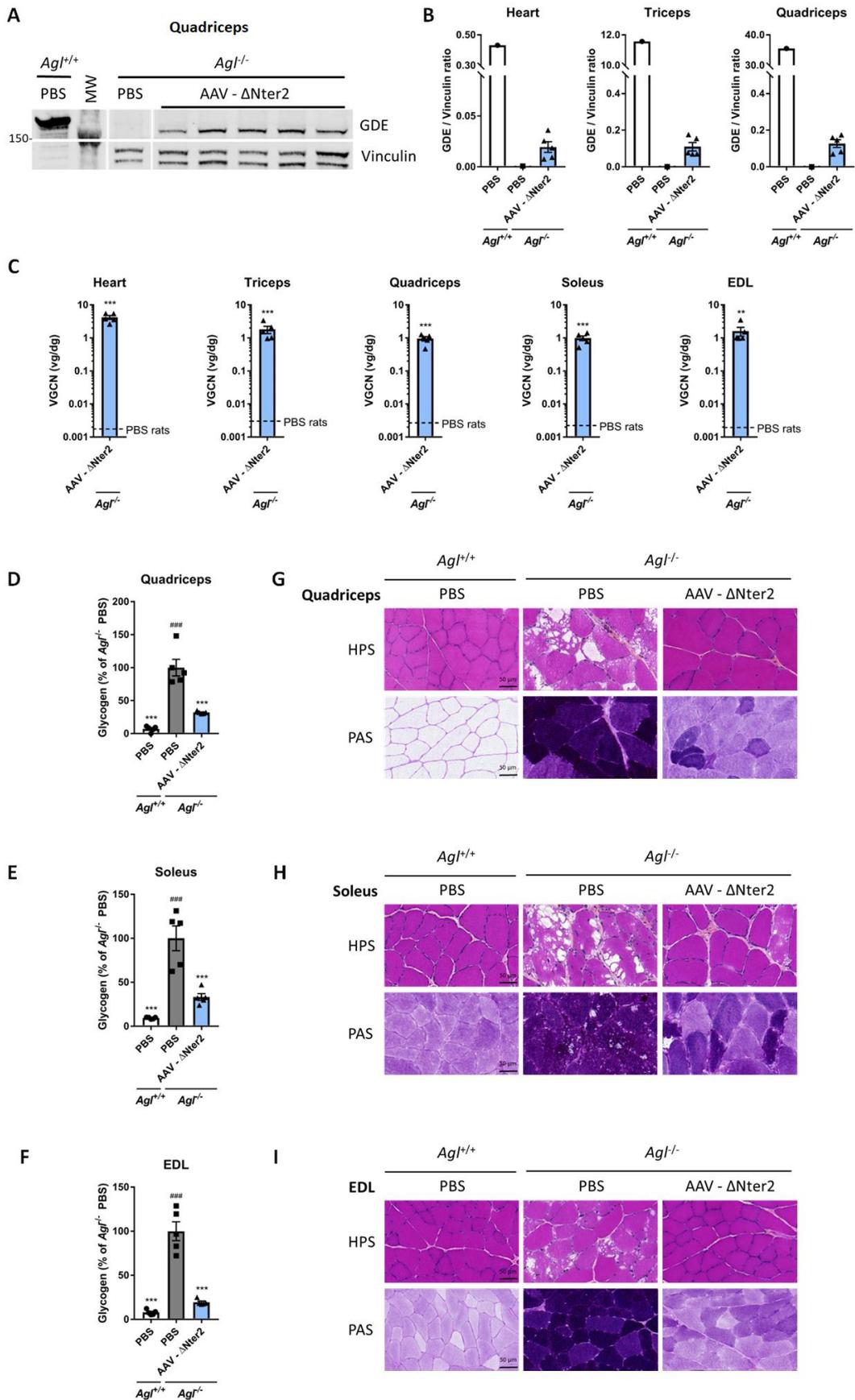


**Figure S9. Improved productivity and quality of vectors bearing truncated  $\Delta$ Nter2-GDE.** (A) Schematic representation of the four expression cassettes evaluated, with their respective size. (B) Small scale rAAV productions (50 mL culture) were performed in triplicate for each vector and viral titers were measured both before (bulk) and after purification (final product) for each triplicate. (C) rAAV vector DNA was extracted and loaded on a 1% Agarose gel to assess genome integrity. Expected genome size range from 5.0 to 5.7 kb. (D) Analytical ultracentrifugation analysis of the proportion of full (right peak in the magenta curve) and empty (left peak in the magenta curve) particles. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. CMV GDE bGh, experiment performed in triplicate). All data are shown as mean  $\pm$  SEM.



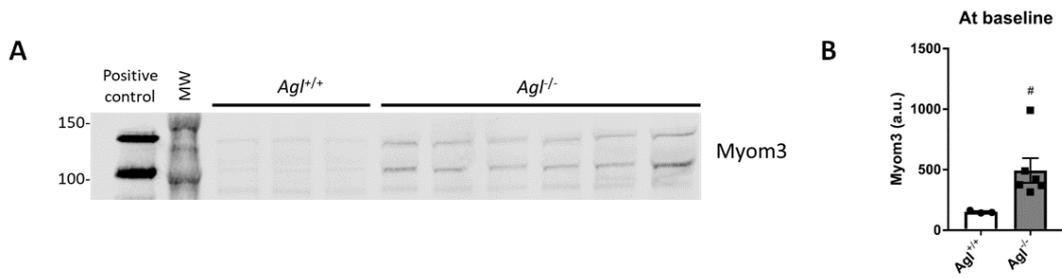
**Figure S10. Rescue of the muscle and heart phenotype of the *Ag1<sup>-/-</sup>* mouse model using rAAV vector encoding for ΔNter2-GDE. (A) Western blot analysis of GDE and Vinculin in the quadriceps three months after injection.**

(B) Quantification of GDE expression on Western Blot showed in Figure 25B and in Figure S10A, expressed as the ratio of the signal of GDE and vinculin bands. (C) VGCN in heart, triceps, quadriceps, soleus and EDL muscles of injected mice. The dashed line represents the background level measured in PBS-injected mice. Only 5 mice per group for soleus and EDL. (D) Glycogen content 3 months after injection in quadriceps, soleus and EDL muscles. Only 5 mice per group for *soleus* and EDL. (E) Histological analysis of quadriceps using HPS and PAS staining. Representative images are shown. (F) Histological analysis of soleus and EDL muscles using HPS and PAS staining. Representative images are shown. (G) Western blot analysis of myomesin3 fragments (Myom3) in serum of mice before injection. Serum from mdx mice was used as positive control. (H) Quantification of Myom3 expression on Western Blot showed in Figure 25G and in Figure S10G (arbitrary units). Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. PBS-injected *Ag1<sup>-/-</sup>* mice; #p<0.05, ##p<0.01 or ###p<0.001 vs. PBS-injected *Ag1<sup>+/+</sup>* mice, n=7-9 mice per group coming from two independent experiments). All data are shown as mean ±SEM.

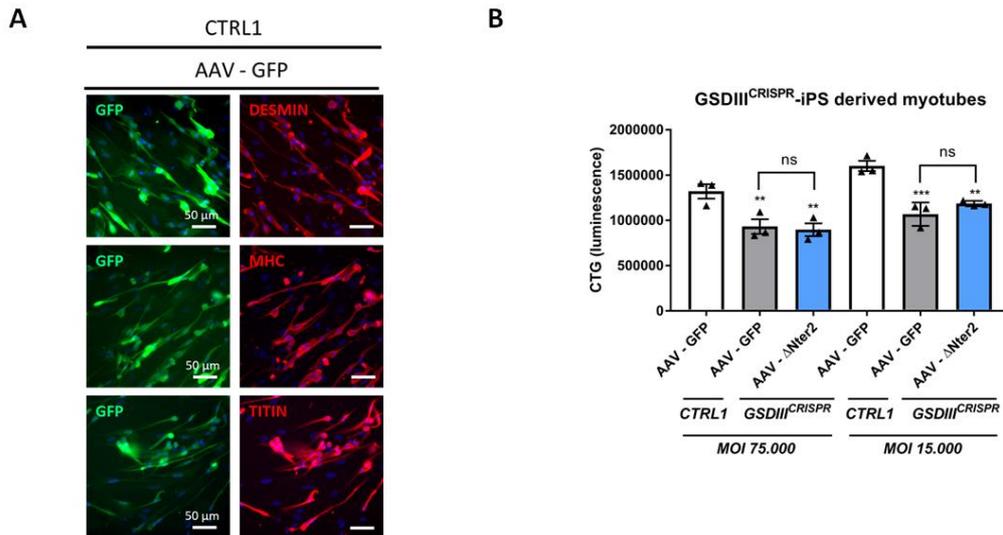


**Figure S11. Correction of the muscle and heart phenotype of the *Ag1<sup>-/-</sup>* rat model using rAAV vector encoding for  $\Delta$ Nter2-GDE. (A) Western blot analysis of GDE and Vinculin in the quadriceps three months after injection.**

**(B)** Quantification of GDE expression on Western Blot showed in Figure 26B and in Figure S11A, expressed as the ratio of the signal of GDE and vinculin bands. **(C)** VGCN in heart, triceps, quadriceps, soleus, and EDL muscles of injected rats. The dashed line represents the background level measured in PBS-injected rats. **(D-F)** Glycogen content in quadriceps **(D)**, soleus **(E)** and EDL **(F)** muscles 3 months after injection. **(G-I)** Histological analysis of quadriceps **(G)**, *soleus* **(H)** and EDL **(I)** muscles using HPS and PAS staining. Representative images are shown. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. PBS-injected *Agf*<sup>-/-</sup> rats; # $p < 0.05$ , ## $p < 0.01$  or ### $p < 0.001$  vs. PBS-injected *Agf*<sup>+/+</sup> rats, n=5 rats per group). All data are shown as mean  $\pm$ SEM.



**Figure S12. Elevation of Myom3 level in the plasma of 6-week-old *Agl*<sup>-/-</sup> rats.** (A) Western blot analysis of myomesin3 fragments (Myom3) in serum of 6-week-old male *Agl*<sup>+/+</sup> and *Agl*<sup>-/-</sup> rats. Serum from mdx mice was used as positive control (B) Quantification of Myom3 expression on Western Blot showed in Panel A (arbitrary units). Statistical analyses were performed by Welch T-test (#p<0.05, ##p<0.01 or ###p<0.001, n=3 *Agl*<sup>+/+</sup> rats and 6 *Agl*<sup>-/-</sup> rats). All data are shown as mean ±SEM.



**Figure S13. Absence of specific toxicity in human induced pluripotent stem cells (hiPSC)-derived muscle cells treated with rAAV-ΔNter2-GDE. (A)** Expression of specific myogenic markers (desmin, myosin heavy chain (MHC) and titin) and green fluorescent protein (GFP) by immuno-fluorescence analysis, after rAAV transduction at MOI 75 000, in CTRL1 hiPSC-derived skMb. Representative images are shown **(B)** Cell viability assessed by evaluating ATP levels in GSDIII<sup>CRISPR</sup> hiPSC-derived skMt and isogenic controls (CTRL1) transduced as described in **Figure 27**. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$  or \*\*\* $p < 0.001$  vs. AAV-GFP-transduced CTRL1 cells). All data are shown as mean  $\pm$  SEM.

**Table S1. Glycogen levels expressed as mmol glucose/g proteins from Figure S4**

		Glycogen (mmol glucose/g proteins)				
Genotype	Treatment (n)	Heart	Triceps	Quadriceps	Soleus	EDL
Agl <sup>+/+</sup>	PBS (n=3)	0.04 ± 0.02	0.15 ± 0.05	0.06 ± 0.01	0.38 ± 0.01	0.33 ± 0.02
Agl <sup>-/-</sup>	PBS (n=3)	0.45 ± 0.08	2.11 ± 0.14	3.49 ± 0.14	2.55 ± 0.16	2.27 ± 0.11
	AAV-GDE bGh (n=4)	0.24 ± 0.04	0.99 ± 0.1	1.32 ± 0.19	1.12 ± 0.18	1.02 ± 0.13
	AAV-GDE pA58 (n=3)	0.25 ± 0.06	0.85 ± 0.08	0.91 ± 0.04	1.02 ± 0.1	0.72 ± 0.09

Values are shown as Mean ± SEM

**Table S2. Glycogen levels expressed as mmol glucose/g proteins from Figure 23**

Genotype	Treatment (n)	Glycogen (mmol glucose/g proteins)		
		Heart	Triceps	Quadriceps
Agl <sup>+/+</sup>	PBS (n=5)	0.002 ± 0.002	0.12 ± 0.02	0.11 ± 0.02
Agl <sup>-/-</sup>	PBS (n=4)	0.85 ± 0.14	2.71 ± 0.26	2.57 ± 0.24
	AAV-ΔC1 GDE (n=5)	0.66 ± 0.04	0.71 ± 0.07	1.33 ± 0.17
	AAV-ΔC2 GDE (n=5)	0.75 ± 0.1	0.99 ± 0.11	0.96 ± 0.07
	AAV-ΔC3 GDE (n=5)	0.98 ± 0.14	1.52 ± 0.17	1.12 ± 0.07
	AAV-ΔC4 GDE (n=4)	0.76 ± 0.04	0.98 ± 0.05	1.13 ± 0.07

Values are shown as Mean ± SEM

**Table S3. Glycogen levels expressed as mmol glucose/g proteins from Figure 25**

Genotype	Treatment (n)	Glycogen (mmol glucose/g proteins)				
		Heart	Triceps	Quadriceps	Soleus <sup>#</sup>	EDL <sup>#</sup>
Agl <sup>+/+</sup>	PBS (n=7)	0.07 ± 0.02	0.09 ± 0.02	0.09 ± 0.02	0.89 ± 0.05	0.71 ± 0.04
Agl <sup>-/-</sup>	PBS (n=8)	1.06 ± 0.09	3.18 ± 0.06	3.35 ± 0.04	6.19 ± 0.28	3.97 ± 0.22
	AAV-ΔNter2 GDE (n=9)	0.25 ± 0.02	0.52 ± 0.05	0.48 ± 0.05	1.13 ± 0.07	0.91 ± 0.04

Values are shown as Mean ± SEM

<sup>#</sup>Only five mice per group had sampling of soleus and EDL.

**Table S4. Glycogen levels expressed as mmol glucose/g proteins from Figure 26**

Genotype	Treatment (n)	Glycogen (mmol glucose/g proteins)				
		Heart	Triceps	Quadriceps	Soleus	EDL
Agl <sup>+/+</sup>	PBS (n=5)	0.15 ± 0.05	0.1 ± 0.01	0.1 ± 0.03	0.46 ± 0.02	0.2 ± 0.03
Agl <sup>-/-</sup>	PBS (n=5)	1.25 ± 0.21	1.99 ± 0.33	1.42 ± 0.18	4.87 ± 0.69	2.52 ± 0.27
	AAV-ΔNter2 GDE (n=5)	0.65 ± 0.07	0.41 ± 0.01	0.46 ± 0.01	1.62 ± 0.19	0.49 ± 0.04

Values are shown as Mean ± SEM

PART II:  
CORRECTION OF THE LIVER DISEASE  
IN GSDIII

## I. INTRODUCTION

In the first section of this thesis, we have demonstrated the ability of an optimized single rAAV vector approach expressing a truncated version of GDE ( $\Delta$ Nter2) to correct the muscle and heart involvement of several models of the GSDIII disease.

As described in the *Introduction* part, the liver and metabolic phenotypes are prominent in young children, but is often overlooked in adult patient which can present with cirrhosis and liver tumors<sup>177,190</sup>. Previous studies have demonstrated that the CMV promoter is not suited for efficient liver expression<sup>69,208</sup> and therefore, the rAAV vector optimized for muscle and heart targeting in the first part of this work might not be used to correct the liver phenotype. The previous dual rAAV vector approach developed by our team allowed a partial correction of the liver phenotype in terms of glycogen content and glycemia using a liver promoter (ApoE-hAAT) and a liver tropic capsid (AAV8)<sup>69</sup>. This approach has recently been optimized<sup>214</sup>. Similarly, the use of a strong liver specific promoter (LSP) was required to achieve pullulanase expression and therefore liver phenotype correction in another studies<sup>42,209</sup>.

In this part, we asked whether a single rAAV vector approach optimized for liver targeting is able to achieve complete and long-lasting hepatic phenotype correction in *AgI*<sup>-/-</sup> mice. First, we described the liver phenotype of the *AgI*<sup>-/-</sup> mouse model. Then, we confirmed the absence of rescue of the liver phenotype by an rAAV vector encoding the GDE transgene under the miniCMV promoter. Then, we assessed the efficacy and limitations of a gene therapy approach using an rAAV vector optimized for transgene expression in the liver, in both old and young *AgI*<sup>-/-</sup> mice.

## II. MATERIAL AND METHODS

*In Vivo Studies.* The *AgI* knockout mice (*AgI*<sup>-/-</sup>, *AgI*<sup>tm1b(EUCOMM)Wtsi</sup>) were previously described<sup>69</sup> and are of mixed background (C57BL/6J and Balb/c). All animal studies were performed according to the French and European legislation on animal care and experimentation (2010/63/EU) and approved by the local institutional ethical committee. rAAV vectors encoding GDE were administered intravenously *via* the tail vein to either 6-week-old or 4-5-month-old male *AgI*<sup>-/-</sup> mice and wild-type littermates (*AgI*<sup>+/+</sup>). Finally, rAAV vectors encoding mouse secreted embryonic alkaline phosphatase (mSEAP) were administered intravenously *via* the tail vein to 6-week-old wild type male C57BL/6J mice. All animals were randomized to receive either rAAVs or phosphate buffer saline (PBS) as controls. Mice were euthanized by cervical dislocation. Liver to body weight ratio was measured at euthanasia.

*Production of rAAV Vectors.* All rAAV vectors used in this study were produced using an adenovirus-free transient transfection method and purified using a chromatographic method as described earlier<sup>85</sup>. Titers of the rAAV vector stocks were determined using a real-time qPCR using primers for the GDE transgene (Forward: 5'-CTG AAG CTG TGG GAG TTC TT-3' and Reverse: 5'-CTC TTG GTC ACT CTT CTG TTC TC-3') or the SV40 poly-A signal (Forward: 5'-AGC AAT AGC ATC ACA AAT TTC ACA A-3' and Reverse: 5'-CCA GAC ATG ATA AGA TAC ATT GAT GAG TT-3') for rAAVs containing mSEAP. The GDE expression cassette contained the full-length human codon-optimized GDE cDNA<sup>69,215</sup> under the control of either the miniCMV promoter (corresponding to the nucleotides 175050\_175400 of the CMV genome, NC\_006273), a murine liver specific promoter (mLSP) or a recently developed human liver specific promoter (hLSP). The poly-A used was the short synthetic 58 bp polyA (polyA58)<sup>216</sup>. The mSEAP expression cassette contained the specified promoter, the SV40 intron, the mSEAP cDNA and an SV40 poly-A. All cassettes were flanked by the inverted terminal repeats of AAV serotype 2 for vector packaging. The capsid used was either the AAV-MT (hybrid between AAV9 and AAVrh74, harboring a P1 peptide)<sup>173</sup> or a chimeric capsid close to AAV8 (Mut, "Mutant5" in the patent n° WO2022003211A1). This Mut capsid was obtained by rational shuffling, with replacement of the hypervariable region n°5 (amino-acids 446-484) of AAV8 by the corresponding region of the #704 AAV2-13 isolate, generating a capsid with similar liver targeting to AAV8 but more efficient heart and muscle transduction.

*Western Blot Analysis.* Mouse tissues were homogenized in phosphate buffer saline (PBS) with cOmplete™ protease inhibitor cocktail (Roche). Protein concentration was determined using the Pierce™ BCA Protein Assay (Thermo Fisher Scientific) according to the manufacturer's instructions. A fraction of 50 µg of total proteins was loaded in each well for both PBS- and rAAV-injected *Agf<sup>-/-</sup>* mice and 10 µg of total proteins per well for PBS-injected *Agf<sup>+/+</sup>* mice. SDS-PAGE electrophoresis was performed in a 4–12% Bis-Tris gradient polyacrylamide gel (NuPAGE™, Invitrogen). After transfer, the membrane was blocked with Intercept Blocking buffer (LI-COR Biosciences) and incubated with an anti-GDE rabbit polyclonal antibody (16582-1-AP, Proteintech) for muscles, an anti-GDE rabbit polyclonal antibody (AS09454, Agrisera) for liver and either an anti-vinculin mouse monoclonal antibody (V9131, Sigma) or anti-actin mouse monoclonal antibody (66009-1-Ig, Proteintech) for normalization. The membrane was washed and incubated with the appropriate secondary antibody (LI-COR Biosciences) and visualized by Odyssey imaging system (LI-COR Biosciences).

*Measurement of Glycogen Content.* Glycogen content was measured indirectly in tissue homogenates as the glucose released after total digestion with *Aspergillus Niger* amyloglucosidase (Sigma-Aldrich, ref A1602). Samples were incubated for 10 min at 95°C and then cooled at 4°C. After the addition of amyloglucosidase (final concentration 4 U/mL) and potassium acetate pH 5.5 (final concentration 25

mM) at 37°C for 90 min, the reaction was stopped by incubating samples for 10 min at 95°C. A control reaction without amyloglucosidase was prepared for each sample and was incubated in the same conditions. The glucose released was determined using a glucose assay kit (Sigma-Aldrich) and the resulting absorbance was acquired on an EnSpire alpha plate reader (PerkinElmer) at a wavelength of 540 nm. Glucose released after amyloglucosidase was then normalized by the total protein concentration.

*Measurement of glycemia and serum Aspartate aminotransferase (AST) and Alanine aminotransferase (ALT) levels.* Blood was sampled in fed conditions in mice anesthetized with isoflurane. Glycemia was measured using a glucometer (Accu-Chek). AST and ALT levels were measured on plasma using microchip DRI-CHEM SLIDE (Fujifilm, AST-P III: ref 3150, ALT-P III, ref 3250) and DRI-CHEM NX500 spectrophotometer (Fujifilm) following the manufacturer's instructions.

*Histology.* For histology, livers were snap-frozen in isopentane previously chilled in liquid nitrogen. Serial 8- $\mu$ m cross sections were cut in a Leica CM3050 S cryostat (Leica Biosystems). To minimize sampling error, 2 sections of each specimen were obtained and stained with Hematoxylin Eosin (HES), Sirius Red or Periodic Acid Schiff (PAS) according to standard procedures. Images were digitalized using Axioscan Z1 slide scanner (Zeiss, Germany) under a Zeiss Plan-Apochromat 10X/0.45 M27 dry objective (ZEISS, Germany). Tile scan images were reconstructed with ZEN software (ZEISS, Germany). For the quantification of fibrosis, the pixel classification feature from QuPath<sup>236</sup> 0.4.3 was used by creating two classifiers, using each time 2 images for ground truth. The first pixel classifier identifies pixels belonging to the tissue slice, excluding veins, fold, dust, bubbles or any artifact encountered, and draw an annotation of the analysable tissue. The second classifier was trained to identify fibrosis and healthy tissue, based on manual annotation of Picrosirius Red staining. The fibrosis classifier was then applied in the annotation created by the first classifier. The total surface of PSR staining was divided by the total surface area of the muscle slice resulting then in a Fibrosis Ratio (% of fibrotic tissue) for each tissue slice.

*Vector Genome Copy Number Determination.* Vector genome copy number was determined using a qPCR assay as previously described<sup>132</sup>. The PCR primers used in the reaction were located in the glucosyltransferase domain of the full-length and truncated human codon-optimized GDE (Forward: 5'-CTG AAG CTG TGG GAG TTC TT-3' and Reverse: 5'-CTC TTG GTC ACT CTT CTG TTC TC-3') or within the SV40 poly-A for mSEAP containing rAAV (Forward: 5'-AGC AAT AGC ATC ACA AAT TTC ACA A-3' and Reverse: 5'-CCA GAC ATG ATA AGA TAC ATT GAT GAG TT-3'). As an internal control, primers located

within the mouse *Titin* gene were used (Forward: 5'-AAA ACG AGC AGT GAC GTG AGC-3' and Reverse: 5'-TTC AGT CAT GCT GCT AGC GC-3').

*Huh7 cells transfection.* Huh7 cells were transfected in 24-well plates using the Opti-MEM™ medium and Lipofectamine™ 3000 transfection reagent (ThermoFisher Scientific) according to the manufacturer's instructions. Cells were harvested 48-hours after transfection, media were collected for mSEAP quantification.

*mSEAP quantification.* mSEAP was quantified in tissue lysates using the Phospha-Light™ Kit (Applied Biosystems) following the manufacturer's instructions and was normalized by the total proteins concentration.

*Muscle Function.* Forelimb wire-hang test was performed as already reported<sup>129,217</sup> at baseline and each month until euthanasia. A 4-mm-thick wire was used to record the number of falls over a period of 3 min. The average number of falls per minute was reported for each animal.

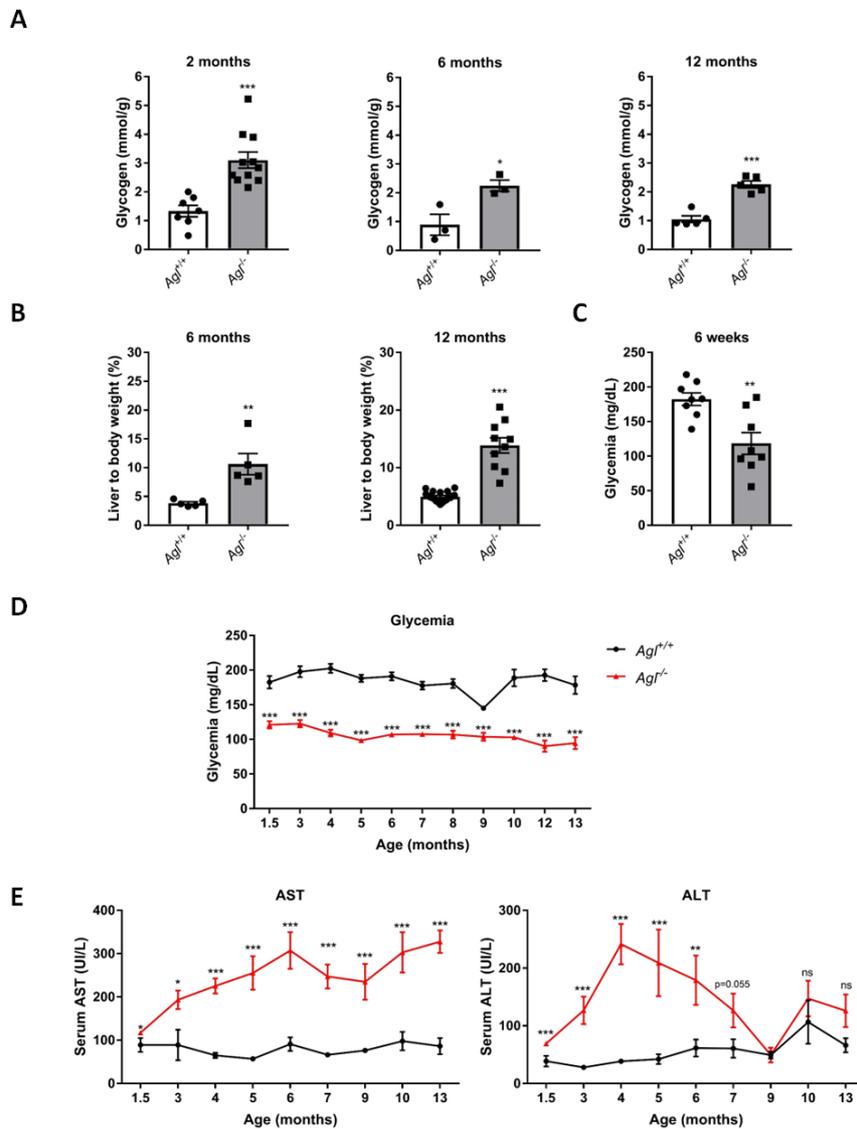
*Statistical Analysis.* All the data shown in this work are reported as mean ± SEM. The GraphPad Prism™ software was used for statistical analysis. p values < 0.05 were considered significant. For all the datasets, data were analyzed by parametric tests, alpha = 0.05 (one-way ANOVA with Tukey's post hoc correction). The statistical analysis performed for each dataset is indicated in each figure legend.

### III. RESULTS

#### 1. CHARACTERIZATION OF THE LIVER INVOLVEMENT IN *AGL*<sup>-/-</sup> MICE

The hepatic phenotype of *Agf*<sup>+/+</sup> and *Agf*<sup>-/-</sup> was analyzed at several ages in order to characterize the progression of the liver disease (Figure 28). First, *Agf*<sup>-/-</sup> mice accumulated two to three times more glycogen than *Agf*<sup>+/+</sup> mice, as early as 2 months of age (Figure 28A). In line with this, *Agf*<sup>-/-</sup> mice displayed a progressive enlargement of the liver at all ages analyzed (Figure 28B). Furthermore, *Agf*<sup>-/-</sup> mice displayed lower glycemia measured under anesthesia in the fed state, at all ages evaluated (Figure 28C-D), although some mice still had normal glycemia at 6 weeks (Figure 28C). No symptoms of hypoglycemia were present neither in the fed or fasted states (no death, no hypoactivity or change in behavior). Elevation of the plasma AST and ALT levels was present from 6 weeks of age (Figure 28E). Although AST levels remained elevated in older mice, likely because of muscle damage, ALT, which are more specific for the liver involvement, improved with age and normalized in older animals (Figure

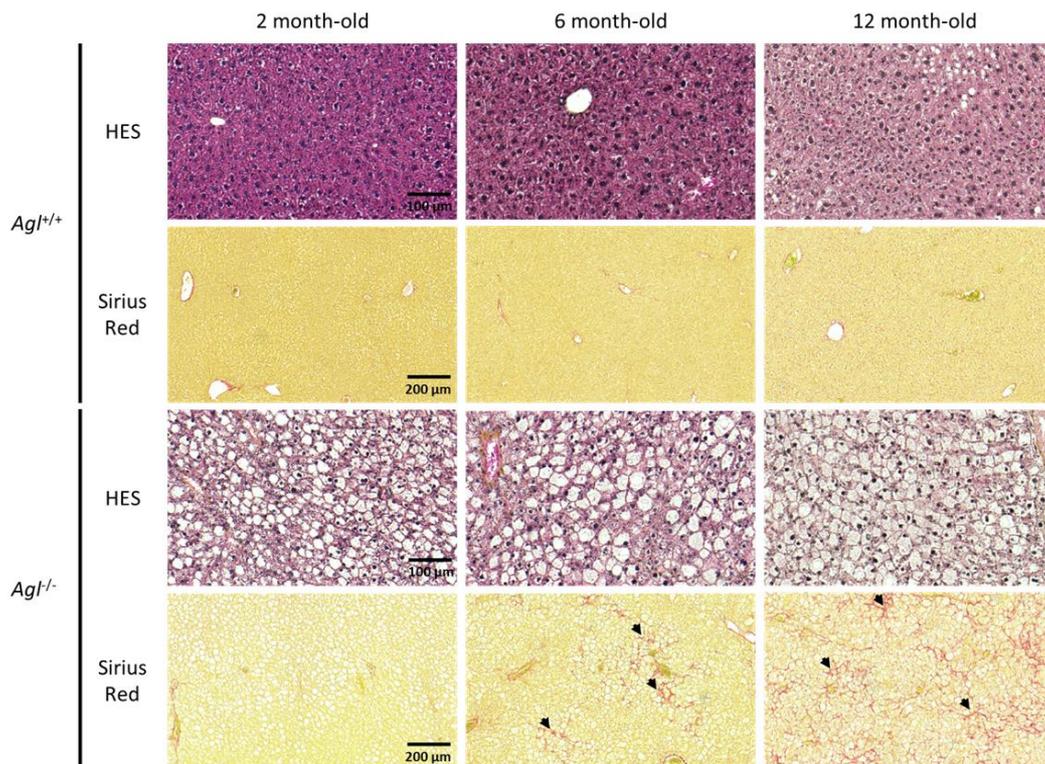
28E). These elements recapitulate the human phenotype and demonstrate the presence of liver involvement early in the course of the disease in mice, starting from 6 weeks of age .



**Figure 28. Characterization of the liver biochemical phenotype of *Agt1*<sup>-/-</sup> mice.** (A) Glycogen content in liver of 2-month, 6-month and 12-month-old *Agt1*<sup>+/+</sup> and *Agt1*<sup>-/-</sup> mice. (B) Hepatomegaly, assessed by the liver to body weight ratio, in 6-month and 12-month-old *Agt1*<sup>+/+</sup> and *Agt1*<sup>-/-</sup> mice. (C-D) Glycemia measured in plasma of *Agt1*<sup>+/+</sup> and *Agt1*<sup>-/-</sup> at 6 weeks of age (C) and its evolution over time (D). (E) AST and ALT levels in plasma of *Agt1*<sup>+/+</sup> and *Agt1*<sup>-/-</sup> at different ages. Statistical analyses were performed by student T-test (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001). All data are shown as mean ± SEM.

We then analyzed the histologic features of the liver involvement in 2-month, 6-month and 12-month-old *Agt1*<sup>+/+</sup> and *Agt1*<sup>-/-</sup> (Figure 29). We observed an intense hepatocyte vacuolization due to glycogen accumulation in *Agt1*<sup>-/-</sup> mice at all ages evaluated, whereas *Agt1*<sup>+/+</sup> mice had normal histology

besides steatosis in older animals (Figure 29). Since the analysis was performed in fed mice, glycogen was present in both  $Agf^{-/-}$  and  $Agf^{+/+}$  mice on PAS staining (Figure S14). To assess the presence of fibrosis, we used a Sirius red staining, which labels in red the collagen fibers. Although  $Agf^{+/+}$  mice did not display fibrosis even at an old age, Sirius red staining revealed a progressive fibrosis in  $Agf^{-/-}$  mice starting from 6 months of life (Figure 29, arrows), especially in 12-month old animals. Of note, young  $Agf^{-/-}$  mice did not present liver fibrosis. To conclude, the liver phenotype is present in young  $Agf^{-/-}$  mice and worsens with age. It can be partially monitored longitudinally in plasma (glycemia, AST, ALT).



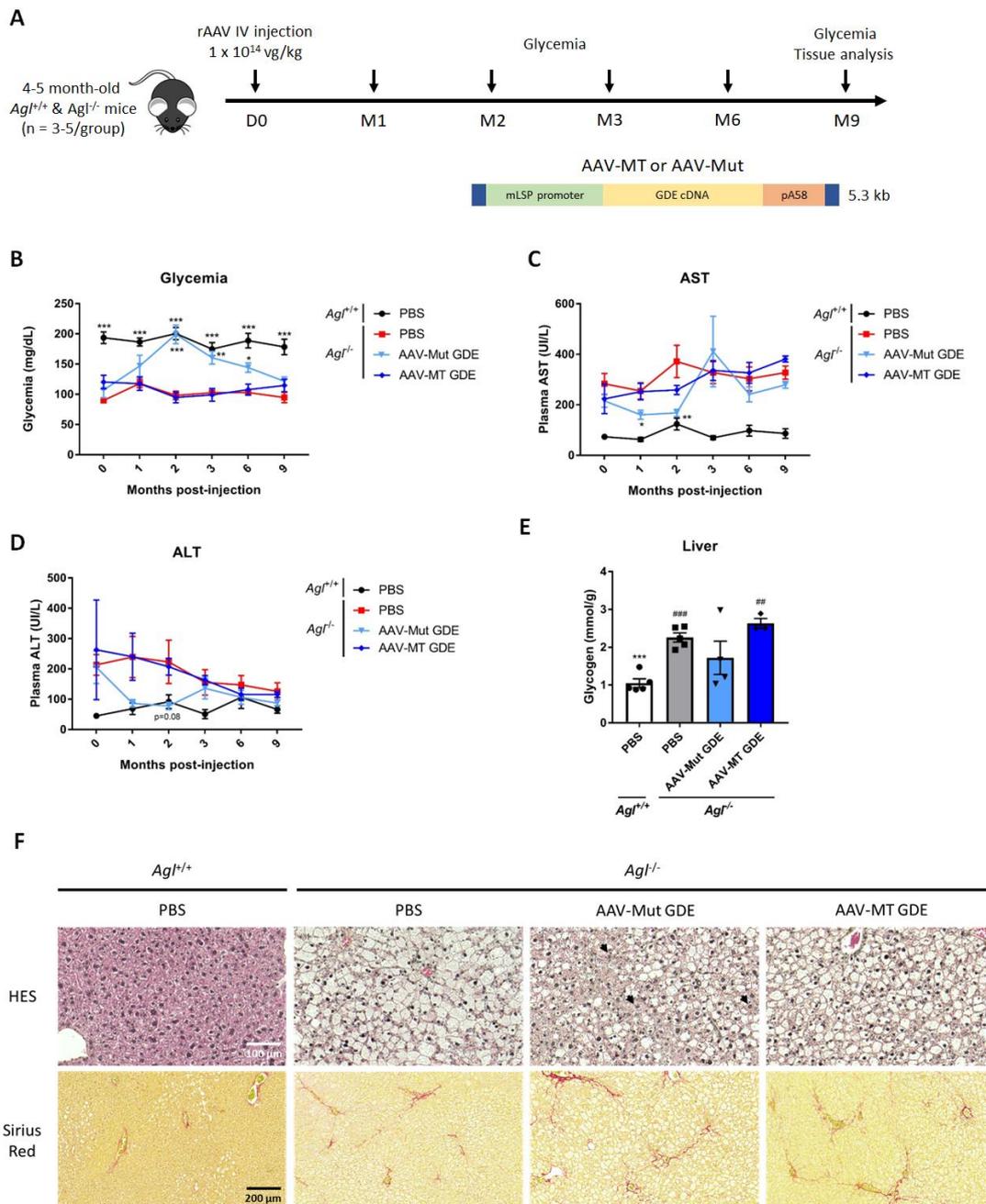
**Figure 29. Characterization of liver histology in  $Agf^{-/-}$  mice.** Histological analysis of the liver of 2-month, 6-month and 12-month-old  $Agf^{+/+}$  and  $Agf^{-/-}$  mice using HES and Sirius red staining. Representative images are shown.

## 2. TRANSIENT CORRECTION OF THE LIVER INVOLVEMENT IN 4-5-MONTH-OLD $AGL^{-/-}$ MICE

Having characterized the course of the liver disease progression in  $Agf^{-/-}$  mice, we evaluated a single vector rAAV approach in 4-5 month-old animals, in whom the whole liver phenotype, including liver fibrosis, is already present. This mimics the treatment of adult GSDIII patients, who already display liver fibrosis.

A rAAV vector encoding the human full-length codon-optimized GDE under the control of the miniCMV promoter was generated with either the liver detargeting AAV-MT capsid (the vector used in Part I, Figure S3) or a chimeric capsid displaying liver targeting (AAV-Mut). High dose rAAV vectors ( $1 \times 10^{14}$  vg/kg) were administered intravenously in adult 4-5-month-old  $Agf^{-/-}$  (Figure S15A). As expected,

analysis of the VGCN in the liver three months after injection revealed a strong targeting of the liver using the AAV-Mut capsid, whereas the AAV-MT almost completely detargeted the liver (**Figure S15B**). No expression of GDE was present in the liver (**Figure S15C**), even when the mice were injected with the AAV-Mut, confirming the absence of expression of CMV-derived promoters in the mouse liver<sup>69,208</sup>. Accordingly, no reduction of glycogen accumulation and no correction of the glycemia was observed (**Figure S15D-E**). Therefore, as expected, the miniCMV promoter is not suited for liver expression and the correction of the liver phenotype will likely require the use of a liver specific promoter.



**Figure 30. Transient correction of the liver phenotype of  $Agt^{-/-}$  mice injected at 5 months of age with a rAAV containing an hepatic promoter.** (A) 4 to 5-month-old male  $Agt^{-/-}$  mice were injected intravenously with a rAAV

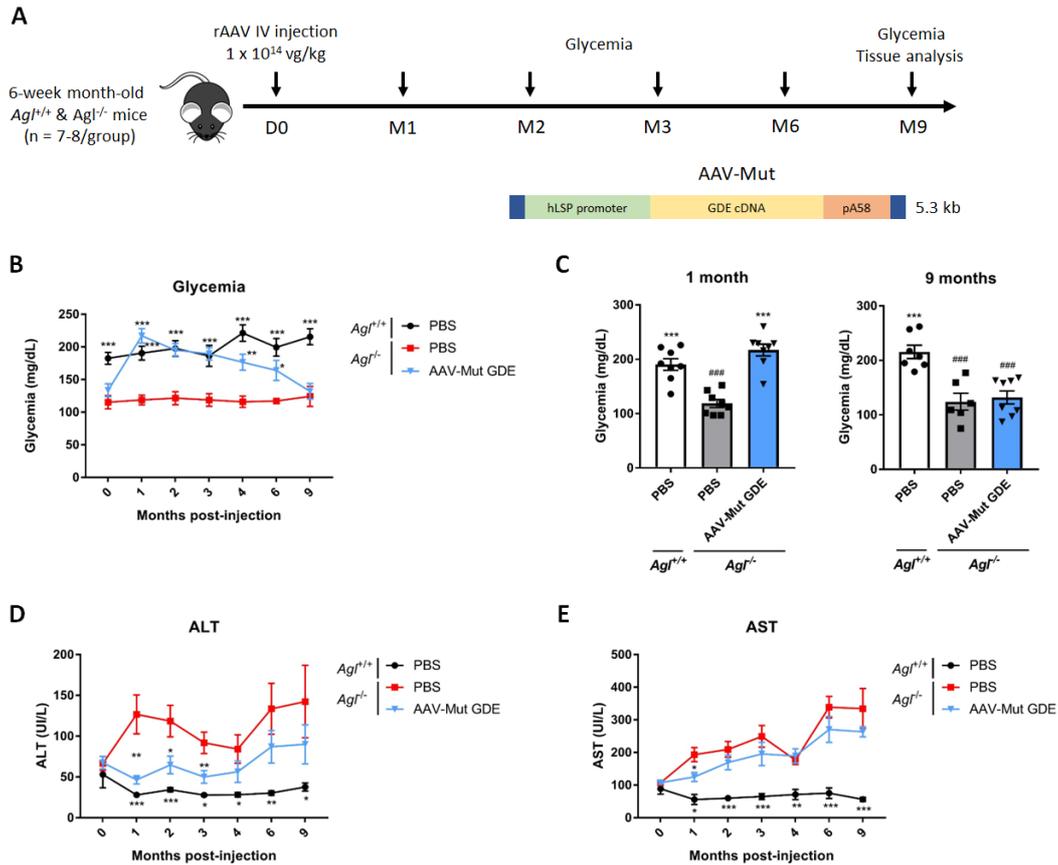
encoding the human full-length codon-optimized GDE cDNA under the control of a mouse liver specific promoter (mLSP) and either the AAV-MT or AAV-Mut capsids at a dose of  $1 \times 10^{14}$  vg/kg. PBS-injected *Agf<sup>+/+</sup>* and *Agf<sup>-/-</sup>* mice were used as controls. (B) Evaluation of glycemia after 1, 2, 3, 6 and 9-months post-injection, measured in the fed state under anesthesia. (C-D) Evaluation of plasma AST (C) and ALT (D) levels at 1, 2, 3, 6 and 9-months post-injection, measured in the fed state under anesthesia. (E) Glycogen content in the liver 9 months after injection. (F) Histological analysis in the liver using HES and Sirius Red staining. Representative images are shown. Arrowheads indicate corrected hepatocytes (no vacuoles). Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  vs. PBS-injected *Agf<sup>-/-</sup>* mice, # $p < 0.05$ , ## $p < 0.01$ , ### $p < 0.001$  vs. PBS-injected *Agf<sup>+/+</sup>* mice,  $n = 3-5$  mice per group). All data are shown as mean  $\pm$  SEM.

Due to the size constraints of the expression cassette, we used a short murine liver specific promoter (mLSP) to generate rAAV vectors able to encapsidate the human full-length GDE cDNA. rAAV-Mut or rAAV-MT vectors were administered intravenously to 4-5-month-old *Agf<sup>-/-</sup>* at the dose of  $1 \times 10^{14}$  vg/kg (**Figure 30A**). Mice injected with the AAV-Mut rAAV showed glycemia correction at 2 months post-injection, but the correction was progressively lost between 3 and 6 months post-injection (**Figure 30B**). No correction was observed with the AAV-MT capsid, suggesting that the liver targeting was too low to achieve sufficient GDE expression (**Figure 30B**). ALT and AST course paralleled the one of the glycemia: plasma AST levels improved in the group of mice injected with the AAV-Mut capsid, whereas it did not in the group injected with AAV-MT (**Figure 30C**); the same tendency was present for serum ALT levels (**Figure 30D**). At 9 months post-injection, VGCN and a faint GDE expression were present in the liver of mice injected with AAV-Mut, in contrast to mice injected with AAV-MT (**Figure S16A-B**). In line with this, no significant reduction of glycogen was observed with either vector (**Figure 30E**), although glycogen levels decreased in three out of 4 of mice injected with AAV-Mut. Analysis of liver histology revealed the persistence of hepatocytes vacuolization in the liver of injected mice, although groups of hepatocytes appeared to be morphologically normal in the liver of mice injected with AAV-Mut (**Figure 30F**). Fibrosis was similar between PBS-injected and rAAV-treated *Agf<sup>-/-</sup>* mice (**Figure 30F**). Only one mouse developed a liver tumor in the PBS-injected *Agf<sup>-/-</sup>* mice group. As expected, due to the absence of GDE expression in muscle (not shown), no correction of the muscle phenotype was observed with either rAAV vector (**Figure S17A-B**).

To conclude, the use of a liver specific promoter and liver-tropic capsid allowed a significant but transient correction of the metabolic phenotype, that was lost between 3- and 6-months post-injection. The identification of the mechanisms behind the transgene expression loss are still under characterization.

### 3. LONG-TERM CORRECTION OF THE LIVER INVOLVEMENT IN 6-WEEK-OLD *AGL*<sup>-/-</sup> MICE

We hypothesized that the loss of correction could be due to incomplete liver correction, with persistence of uncorrected proliferative hepatocytes. Importantly, the liver is already damaged at 6 months of age with significant liver fibrosis (Figure 29), which might either impede hepatocyte transduction and/or reduce the potential for liver correction.

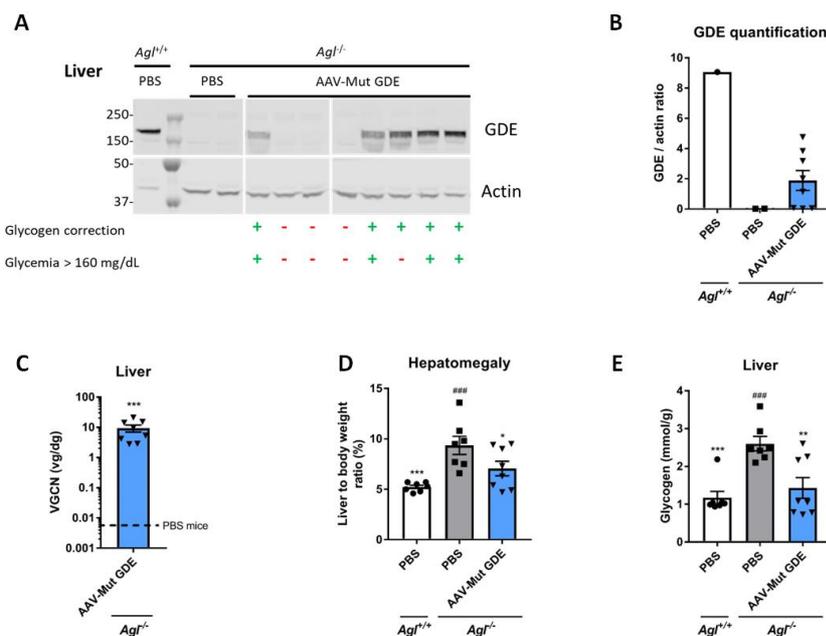


**Figure 31. Long term correction of the hepatic phenotype in some *Agl*<sup>-/-</sup> mice injected at 6 weeks of age with an rAAV vector optimized for liver targeting.** (A) 6-week-old male *Agl*<sup>-/-</sup> mice were injected intravenously with a rAAV encoding the full-length human codon-optimized GDE cDNA under the control of the hLSP promoter and the liver targeting AAV-Mut capsid at a dose of 1 x 10<sup>14</sup> vg/kg. PBS-injected *Agl*<sup>+/+</sup> and *Agl*<sup>-/-</sup> mice were used as controls. (B) Evolution of glycemia 1, 2, 3, 4, 6 and 9-months after injection, measured in the fed state under anesthesia. (C) Glycemia at 1 month and 9 months after injection. (D-E) Evaluation of plasma ALT (D) and AST (E) levels at 1, 2, 3, 4, 6 and 9-months post-injection, measured in the fed state under anesthesia. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001 vs. PBS-injected *Agl*<sup>-/-</sup> mice; #p<0.05, ##p<0.01, ###p<0.001 vs. PBS-injected *Agl*<sup>+/+</sup> mice, n=7-8 mice per group). All data are shown as mean ± SEM.

Therefore, we hypothesized that better correction of the liver phenotype could be obtained by selecting a stronger liver promoter and by treating *Agl*<sup>-/-</sup> mice at an earlier age. First, we evaluated a recently developed short human liver specific promoter (hLSP) *in vitro* and *in vivo* using the mSEAP

reporter gene system (Figure S18). We first performed transient transfection of human hepatoma cells (Huh7) with plasmids encoding the mSEAP reporter gene under the control of selected promoters and quantified mSEAP secreted in the cell media. Similar mSEAP levels were achieved with both the mLSP and hLSP promoters (Figure S18A), comparable to the level achieved using the much larger ApoE-hAAT promoter (723 bp). Then, we generated rAAVs expressing mSEAP under the control of the mLSP, the hLSP or the ApoE-hAAT promoters and administered these rAAVs *in vivo* to 6-week-old wild type C57BL/6 mice (Figure S18B). Higher mSEAP expression was observed in mice after 1 month of treatment with rAAVs bearing hLSP-mSEAP compared to mLSP-mSEAP (Figure S18C), despite similar VGCN (Figure S18D), suggesting higher strength of the hLSP promoter. Therefore, we selected this promoter for further experiments.

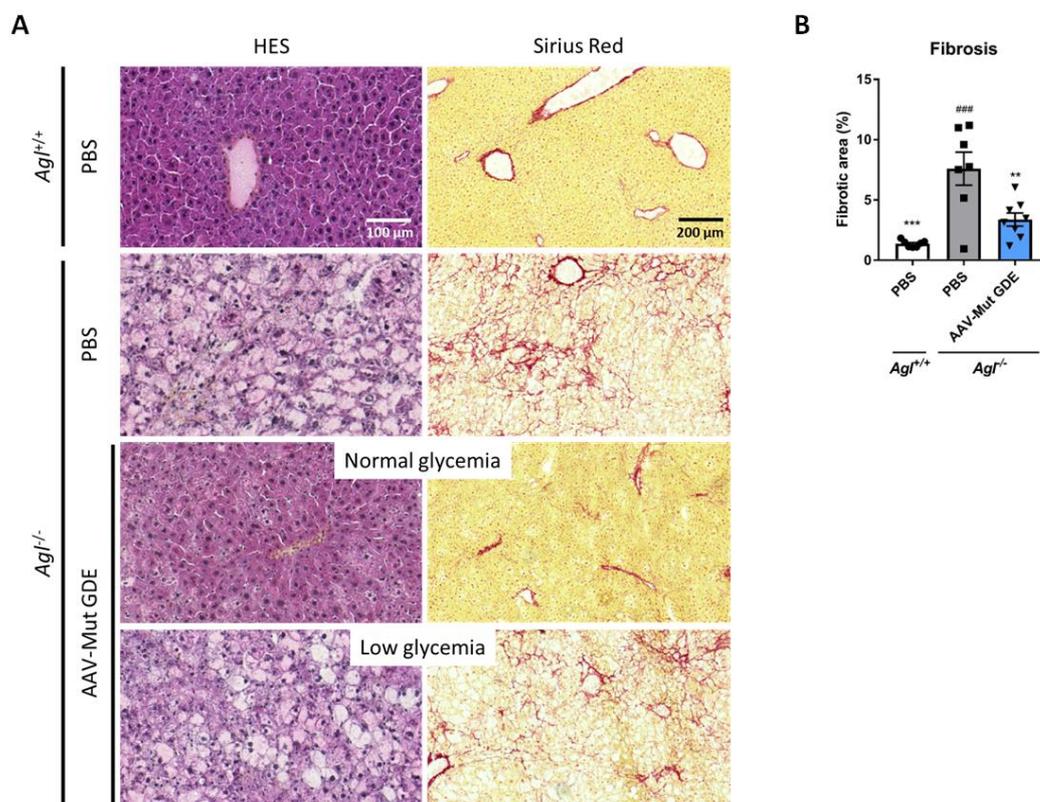
We generated a rAAV vector containing the human full-length codon-optimized GDE cDNA under the control of the hLSP promoter and with the AAV-Mut capsid (Figure 31A). rAAV vectors were administered intravenously to 6-week-old *Agf<sup>-/-</sup>* mice and regular blood sampling was performed up to 9 months post-injection to analyze both glycemia and liver enzymes (Figure 31A). Mice injected with rAAV vectors displayed complete liver correction at 1 month post-injection, with normal glycemia (Figure 31B-C) and serum ALT levels (Figure 31D), whereas serum AST levels were not corrected (Figure 31E). Correction of the glycemia was maintained for 6 months but was subsequently lost in most of the mice injected with the rAAV vector at 9 months after injection (Figure 31B-C). Although not significant, there was a tendency for lower ALT levels in these animals (Figure 31D), with all animals with normal glycemia at euthanasia (above 160 mg/dL) harboring normal ALT level (Figure S19A).



**Figure 32.** Expression of GDE in the liver of treated mice and normalization of glycogen content in some *Agf<sup>-/-</sup>* mice injected at 6 weeks of age with an rAAV optimized for liver targeting. (A) Western blot analysis of GDE

and vinculin in liver 9 months after injection. Below are provided the status of each mice for the correction of glycogen or glycemia at euthanasia (+ indicates correction, - indicates absence of correction). Normal values for glycemia were defined as > 160 mg/dL, this threshold corresponding to -2 standard deviations of the distribution of the glycemia values observed in PBS-injected *Agf<sup>+/+</sup>* mice. (B) Quantification of GDE expression on Western Blot in panel A, expressed as the ratio of the signal of the GDE and vinculin bands. (C) VGCN in the liver 9 months after injection. The dashed line represents the level measured in PBS-injected mice. (D) Liver to body weight ratio 9 months after injection. (E) Glycogen content in the liver 9 months after injection. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  vs. PBS-injected *Agf<sup>-/-</sup>* mice; # $p < 0.05$ , ## $p < 0.01$ , ### $p < 0.001$  vs. PBS-injected *Agf<sup>+/+</sup>* mice,  $n = 7-8$  mice per group). All data are shown as mean  $\pm$  SEM.

Analysis of the liver 9 months after injection revealed GDE expression in half of the mice in the treated group (Figure 32A-B), most of these mice displaying a normal glycemia (Figure 32A). Vector genomes were detected in the liver with some variability between the mice (Figure 32C). Similarly, half of the treated mice displayed a normalization of the liver to body weight ratio, reflecting correction of hepatomegaly (Figure 32D), as well as a significant reduction of glycogen content in the liver (Figure 32E). All animals with high glycogen content also had low glycemia (Figure S19B) and absence of GDE expression and corresponds to the mice with loss of correction of the phenotype over time.



**Figure 33: Correction of liver histology in some *Agf<sup>-/-</sup>* mice injected at 6 weeks of age with an optimized rAAV for liver targeting.** (A) Histological analysis in the liver using HES and Sirius Red staining. Representative images

are shown. Mice were classified as having either normal ( $\geq 160$  mg/dL,  $n = 4$ ) or low ( $< 160$  mg/dL,  $n = 4$ ) glycemia. This threshold corresponds to -2 standard deviations of the glycemia distribution of the PBS-injected *Agf<sup>+/+</sup>* group. (B) Quantification of the fibrotic tissue area in percentage of the total tissue. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  vs. PBS-injected *Agf<sup>-/-</sup>* mice; # $p < 0.05$ , ## $p < 0.01$ , ### $p < 0.001$  vs. PBS-injected *Agf<sup>+/+</sup>* mice,  $n = 7-8$  mice per group). All data are shown as mean  $\pm$  SEM.

Finally, histological analysis of the liver revealed the presence of two different groups of mice according to the correction of the liver phenotype (**Figure 33A**). In mice who displayed normal glycemia ( $> 160$  mg/dl,  $n = 4$ ), most hepatocytes were normal with only few cells with either clarification of the cytoplasm or vacuolization (**Figure 33A**). Fibrosis was either absent or slight in these mice, close to the level observed in *Agf<sup>+/+</sup>* mice (**Figure 33B**). In contrast, mice with loss of correction over time (glycemia  $< 160$  mg/dL,  $n = 4$ ) showed a recurrence of the GSDIII phenotype, with intense vacuolization of hepatocytes and massive fibrosis (**Figure 33A**). Finally, we observed the occurrence of liver tumor in few mice, one in the PBS-injected group and two in the rAAV-injected group. Interestingly, the two mice injected with rAAV that developed tumors lost the correction of the liver phenotype over time. The potential link with the injection of the vector is yet to be elucidated.

To conclude, long-term complete correction was achieved up to 9 months post-injection in half of *Agf<sup>-/-</sup>* mice with the use of an optimized rAAV vector intended for liver targeting injected at a young age. The reason behind the loss of correction observed in half of our cohort is still unclear and is currently being evaluated. It could involve either an immune response exacerbated by the underlying liver disease and/or the persistence of hepatocyte proliferation with progressive rAAV vector dilution.

#### **IV. DISCUSSION**

Although the main disease burden in adult patients with GSDIII is the muscle and heart impairment, the liver and metabolic involvement is preeminent in children and often overlooked in adult patients<sup>177,190</sup>. Hypoglycemia and hepatomegaly are the first symptoms to appear, with increase of serum AST and ALT levels<sup>176,177</sup>. Then, progressive liver fibrosis develops often from an early age, affecting most of the patients at adulthood<sup>189,190</sup>. Although decompensated cirrhosis is a rare event, these patients can develop liver tumors, including HCC, and can require liver transplantation<sup>177,190</sup>. The mouse model used in this work recapitulates the major clinical and biochemical features of the liver involvement observed in GSDIII patients, with hepatomegaly, low glycemia and elevation of liver enzymes from an early age. Liver fibrosis developed progressively starting from 6 months of age, and worsened with age, along with hepatomegaly. Tumors are present in up to 30% of the animals at 14 months of age (*Montalvo-Romeral et al.*, manuscript in preparation).

Previous studies using either a dual vector approach<sup>69</sup> or a bacterial enzyme<sup>42,209</sup> suggested that the correction of the liver phenotype with rAAV was difficult to achieve in GSDIII mice. Indeed, in this study, we demonstrated that even high doses of rAAV vectors ( $1 \times 10^{14}$  vg/kg) optimized for liver targeting could only achieve a partial and transient correction of the phenotype in symptomatic 4-5 month-old *Agf<sup>-/-</sup>* mice, in which the liver fibrosis is already present at baseline. In contrast, the phenotype was corrected more easily in *Agf<sup>-/-</sup>* mice injected at 6 weeks of age although half of the treated mice still displayed progressive loss of correction over time. The difficulty to achieve correction in this mouse model and the mechanisms behind the loss of correction over time deserve further discussion.

First, the requirement for a high rAAV dose and the delayed correction (2 months) observed in *Agf<sup>-/-</sup>* mice injected at 4-5 months of age was unexpected based on previous gene therapy studies performed in animal models of Crigler-Najjar disease<sup>132</sup> or GSD Ia<sup>237</sup>. In these models, biochemical correction is observed as early as 15 days after rAAV injection with a much lower rAAV dose ( $5 \times 10^{11}$  to  $1 \times 10^{13}$  vg/kg)<sup>132,237</sup>. In both diseases, even few corrected hepatocytes can be sufficient to detoxify the bilirubin accumulation (in Crigler-Najjar disease) or increase glucose production to normalize glycemia (in GSD Ia), since, in contrast to GSDIII, corrected hepatocytes can compensate for uncorrected cells. Nonetheless, in GSD Ia, the restoration of enzymatic activity in a larger number of hepatocytes is required to achieve the complete correction of the metabolic phenotype, beyond the sole prevention of hypoglycemia, and especially to prevent the occurrence of tumors<sup>143</sup>.

In GSD Ia, the enzymatic defect affects the gluconeogenesis pathway, which allows glucose production from glucose-6-phosphate (and therefore glycogen), lactate and amino-acids<sup>180</sup>. It leads to the accumulation of all these metabolites in hepatocytes and in the bloodstream<sup>176,180,187,238</sup>. In GSD Ia, glycogen breakdown is functional but inhibited by the high amount of glucose-6-phosphate accumulated secondary to the enzymatic blockade<sup>180,187</sup>. Since many of these metabolites (lactate, amino-acids...) can be easily transferred between hepatocytes, the restoration of G6Pase in some hepatocytes would reduce the global accumulation of these metabolites and progressively restore glycogen breakdown both in corrected but also uncorrected hepatocytes ("sink hypothesis", as described in other diseases<sup>77</sup>). This allows normalization of the glycemia despite the correction of only few hepatocytes<sup>237</sup>. In contrast, in GSDIII, the enzymatic defect affects the breakdown of glycogen, and therefore the restoration of GDE would allow the production of glucose only from the glycogen stores of the corrected hepatocytes, as long as these stores are still available. Since neither the accumulated cytosolic glycogen, nor the GDE protein, can be transferred between hepatocytes in GSDIII, long-term rescue of the hepatic phenotype requires the broad transduction and stable correction of most of the hepatocytes. This would imply the existence of a "threshold" of corrected cells required to restore the normal regulation of glycemia, since corrected hepatocytes cannot accumulate more glycogen in order

to compensate for uncorrected ones. Interestingly, the AAV gene transfer in *Agf<sup>-/-</sup>* mice at an earlier age (and with a stronger promoter) was associated with faster correction, as early as 1 month of age. These mice did not display fibrosis yet (**Figure 29**) in contrast to older 4-5-month-old *Agf<sup>-/-</sup>* mice. Although the presence of fibrosis has been shown to reduce the transduction of the liver in mouse models of severe fibrotic liver disease such as PFIC<sup>147</sup>, in GSDIII mice fibrosis is nonetheless much milder and cannot explain alone this difference of timing of correction. Since GSDIII is a progressive storage disease, one hypothesis could be the presence of a more advanced disease in older animals, requiring therefore more time to be corrected by the delivered GDE.

Second, another interesting element is the loss of correction observed in both *Agf<sup>-/-</sup>* animals treated either at a young or old age. In both experiments, the loss of correction follows the same pattern, with a progressive decrease of the glycemia between 3 and 6-months post-injection and complete loss of correction at 9 months post-injection in all *Agf<sup>-/-</sup>* mice treated at 4-5 months of age and in one half of our cohort of younger *Agf<sup>-/-</sup>* treated mice. Several hypotheses can be made to explain this phenomenon: either the transduced cells are lost and/or the vector is still present but silenced<sup>239</sup>. For the first hypothesis, the mechanism could either be: 1/an insufficient transduction, with persisting proliferation and progressive dilution of the vector, and/or 2/ a toxic effect of the vector and/or of the amount of GDE expressed in cells and/or 3/ the presence of an immune response directed against GDE and/or the rAAV capsid. Whatever the mechanism involved, this would lead to the loss of transduced cells and to the progressive reappearance of the phenotype. Although several of these factors could be involved at the same time, the leading hypothesis would be that we achieved only partial transduction of the liver. In GSDIII, proliferation is known to occur, as demonstrated by others<sup>202</sup>, and is associated with activation of stellate cells and deposition of fibrosis<sup>202</sup>. After rAAV injection, it is possible that the loss of uncorrected hepatocytes with time perpetuates an environment of inflammation and mitogenic signals in the liver, therefore inducing proliferation of transduced hepatocytes and progressive vector dilution. This hypothesis is compatible with the slow timing of the loss of correction that we observed, but does not explain why only half of the cohort of mice lost the correction. Although the liver histology in corrected animals suggest that most hepatocytes are corrected (without vacuoles), we have no precise estimation of the fraction of corrected cells. The development of an immunostaining for GDE on liver samples would allow us to obtain such an estimate and is currently ongoing. Similarly, evaluation of proliferation using either immunostaining on liver histology (Ki67) and/or RNA quantification of genes associated with proliferation in the liver could provide insights for the proliferation hypothesis.

Another possible explanation is the presence of toxicity or immune response. A toxic effect of GDE itself is unlikely to manifest only 3 to 6 months post-injection and has not been observed in other studies targeting muscles with rAAV vectors expressing full-length or truncated GDE. Nonetheless, the

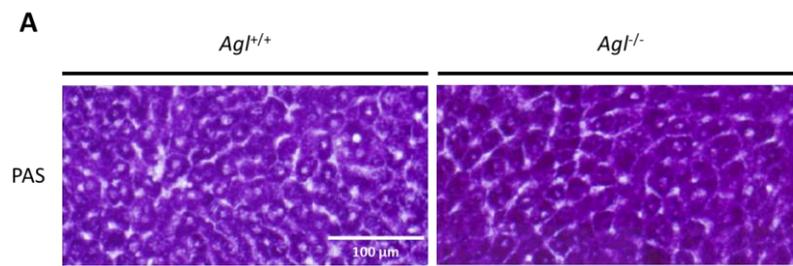
rAAV vectors used in this study are largely oversized, likely containing a significant proportion of truncated genomes (see **Figure S9** for muscle-targeting rAAV), potentially leading to liver-specific toxicities or enhanced immunogenicity. Long-term persistence has been observed in the heart and muscles of *Agf<sup>-/-</sup>* mice treated with AAV-MT vectors expressing full-length GDE up to 9 months after injection (data not shown), which is not in favor of a cytotoxic T-cell mediated rejection of transduced cells. However, we cannot exclude the possibility that an immune response occurred specifically in the liver, especially given the inflammatory context with multiple hepatocyte losses. We did not observe rises of serum AST and ALT levels following rAAV injection, in contrast to human clinical trials<sup>81</sup>. To confirm the presence of an immune response, the analysis of immune response using an ELISPOT assay with GDE peptides could be valuable to understand the potential contribution of an immune response. Finally, a last hypothesis could be an extinction of the expression cassette despite the vector still being present in the transduced cells, as recently described in mice transduced with rAAV vectors expressing factor FVIII<sup>239</sup>. Progressive epigenetic silencing is also described for the CMV promoter<sup>208</sup> and in case of high CpG content within the expression cassette<sup>93</sup>. High CpG content was shown to play a role in sudden loss of transgene expression in human clinical trials for hemophilia<sup>81,93</sup>, through the activation of TLR9, and to lead to cytokines secretion and loss of transduced cells<sup>31</sup>. The evaluation of a CpG-depleted cassette could allow us to test this hypothesis.

Despite the progresses made through this work, several challenges remain for the treatment of the liver disease in GSDIII. First, a better understanding of the mechanisms behind the loss of correction observed in liver is required in order to adapt our gene therapy strategy. Second, the potential role of the rAAV cassette in the development of the tumors observed in *Agf<sup>-/-</sup>* mice injected at a young age needs to be further investigated. Although the frequency of tumors was similar to PBS-injected mice and was restricted to *Agf<sup>-/-</sup>* mice displaying complete loss of correction, an analysis of potential integrations in the tumors could be valuable to understand the potential contribution of the rAAV. The occurrence of tumors in this disease is not only a challenge in terms of safety but also in terms of efficacy. Indeed, in order to prevent their development, broad transduction of hepatocytes is likely to be required if the tumor triggering mechanism is cell autonomous, since HCC could still arise from focal areas of untransduced hepatocytes. Furthermore, some studies performed in a dog model of GSDIa<sup>240</sup> and in a mouse model of PFIC3<sup>147</sup> have suggested that the prevention of tumors could require a higher vector dose compared to the dose required to correct the biochemical phenotype. Third, we have shown in the first part that the use of a truncated N-terminal mutant ( $\Delta$ Nter2) is better encapsidated in rAAV, reaches more cells than the oversized vectors and therefore allows a better correction of the heart and muscle tissues. In line with this, we will evaluate whether  $\Delta$ Nter2-expressing rAAV vectors could also improve the correction of the liver phenotype and prevent the loss of correction. Lastly, the goal is to achieve correction of liver, heart and muscles tissues. Due to size

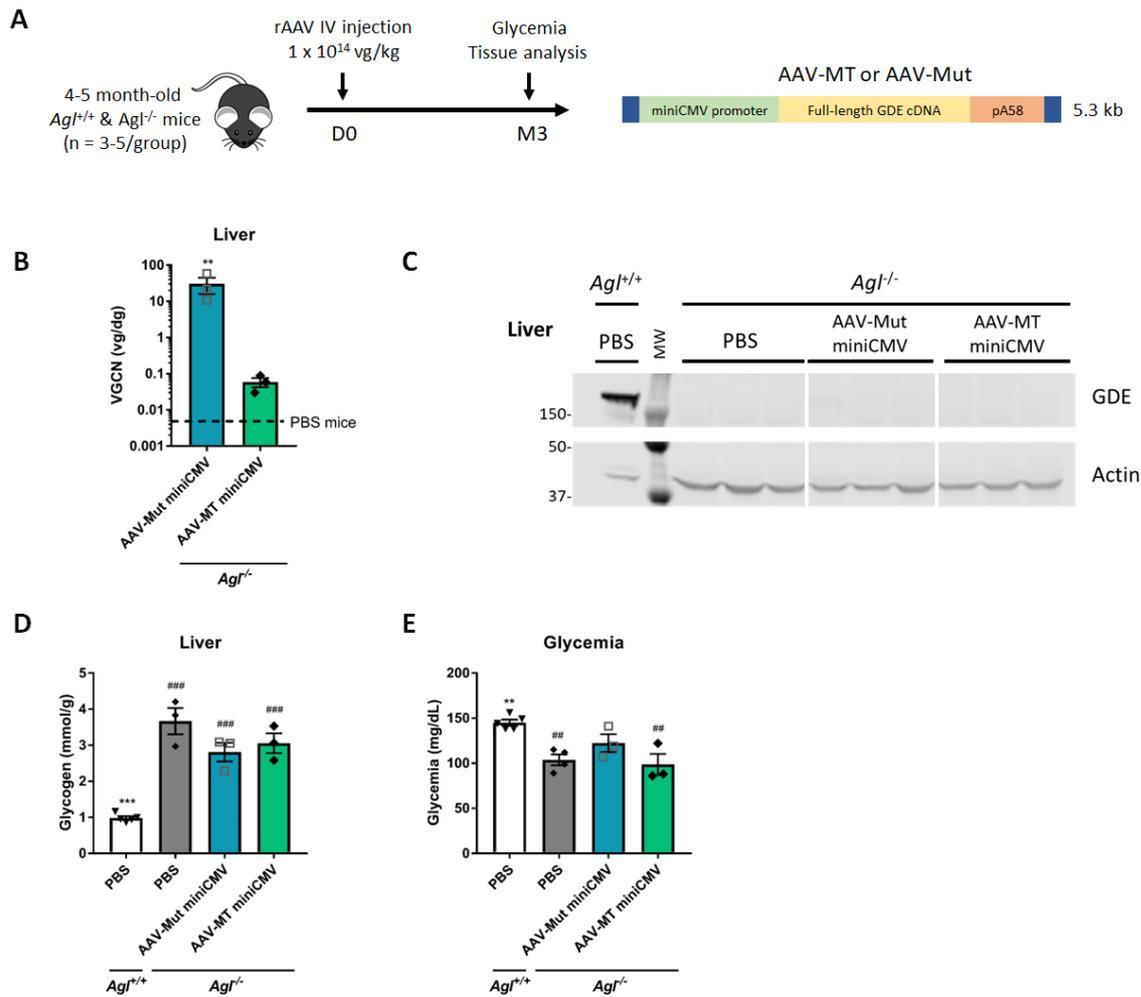
constraints, efforts are needed to identify a small promoter with dual expression abilities in both liver and muscles, in order to be able to cure GSDIII with a unique rAAV injection, if possible.

To conclude, we have optimized an rAAV for liver targeting and shown its limitations with progressive loss of correction in both young and old *Agf<sup>-/-</sup>* mice. The causative mechanisms are still under characterization, and might involve hepatocyte proliferation and/or immune responses. Further optimizations are required especially concerning the prevention of liver tumor development. Finally, the *Agf<sup>-/-</sup>* mouse model could be a valuable tool to understand the determinants of long-term persistence in the liver and to evaluate different integrative or non-integrative approaches to treat genetic fibrotic diseases of the liver.

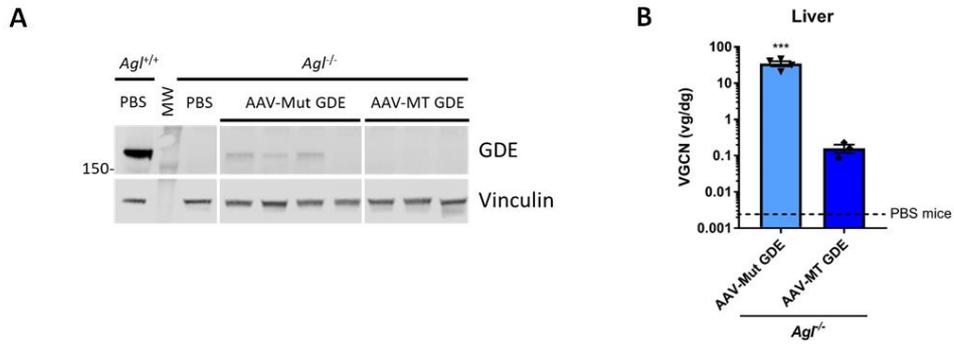
V. SUPPLEMENTARY MATERIALS



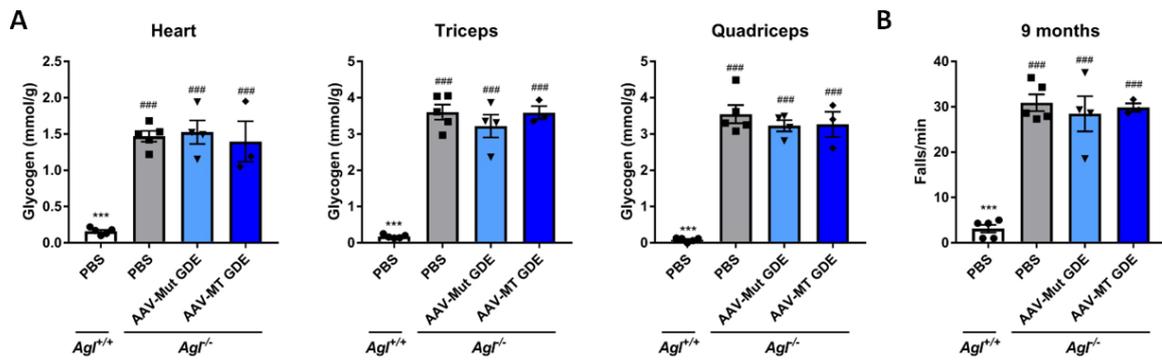
**Figure S14. PAS staining in the liver of *Ag1<sup>+/+</sup>* and *Ag1<sup>-/-</sup>* mice.** Histological analysis in the liver using PAS staining. Representative images are shown.



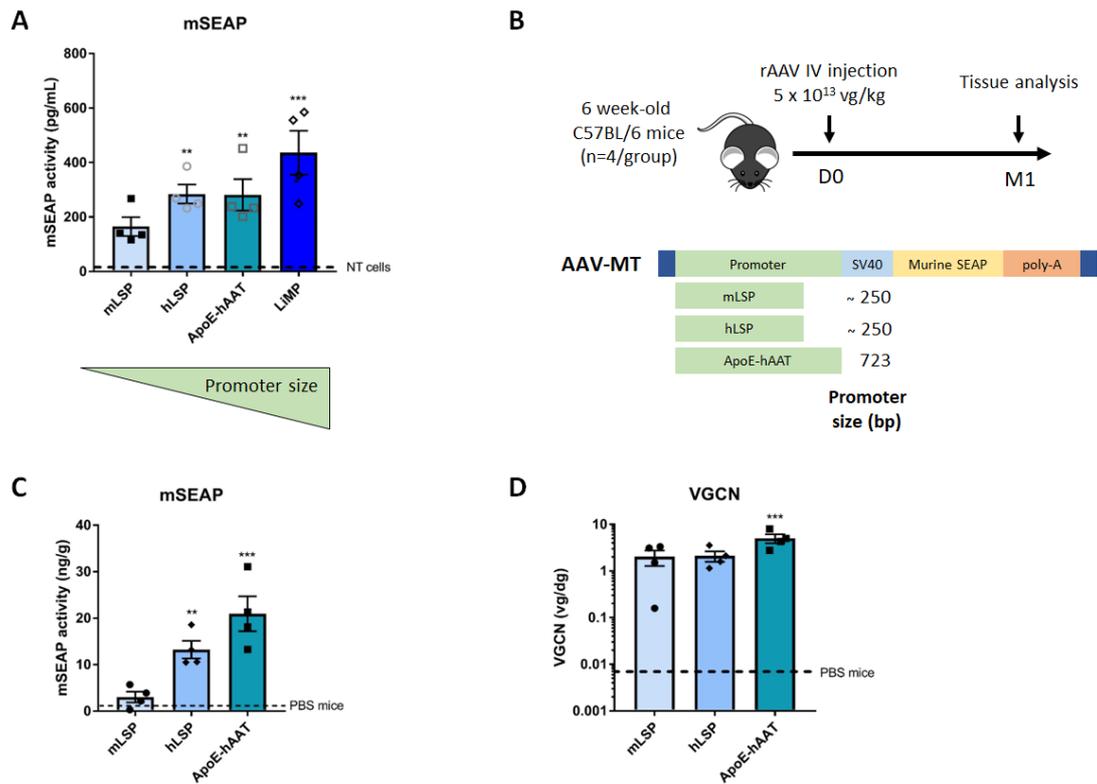
**Figure S15. Absence of correction of the liver phenotype in  $Agl^{-/-}$  injected with rAAV containing the miniCMV promoter.** (A) 4 to 5-month-old male  $Agl^{-/-}$  mice were injected intravenously with a rAAV encoding the human full-length codon-optimized GDE cDNA, the miniCMV promoter and either the AAV-MT or AAV-Mut capsids at a dose of  $1 \times 10^{14}$  vg/kg. PBS-injected  $Agl^{+/+}$  and  $Agl^{-/-}$  mice were used as controls. (B) VGCN in the liver 3 months after injection. The dashed line represents the level measured in PBS-injected mice. (C) Western blot analysis of GDE and actin in livers 3 months after injection. (D) Glycogen content in the liver 3 months after injection. (E) Glycemia three months after injection, measured in the fed state under anesthesia. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  vs. PBS-injected  $Agl^{-/-}$  mice; # $p < 0.05$ , ## $p < 0.01$ , ### $p < 0.001$  vs. PBS-injected  $Agl^{+/+}$  mice, n=3-5 mice per group). All data are shown as mean  $\pm$  SEM.



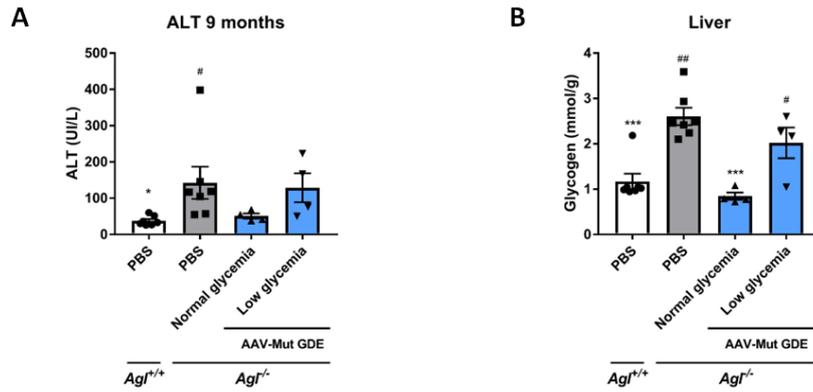
**Figure S16. Absence of GDE expression in the liver of *Agl*<sup>-/-</sup> mice injected with rAAV containing the mLSP promoter.** (A) Western blot analysis of GDE and vinculin in the liver 9 months post-injection. (B) VGCN in the liver 9 months after injection. The dashed line represents the level measured in PBS-injected mice. Statistical analyses were performed by one-way ANOVA (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  vs. PBS-injected *Agl*<sup>-/-</sup> mice, # $p < 0.05$ , ## $p < 0.01$ , ### $p < 0.001$  vs. PBS-injected *Agl*<sup>+/+</sup> mice,  $n = 3-5$  mice per group). All data are shown as mean  $\pm$  SEM.



**Figure S17. Absence of correction of the muscle phenotype in *Agr1*<sup>-/-</sup> mice injected with rAAV containing the mLSP promoter.** (A) Glycogen content measured in heart, triceps and quadriceps at sacrifice 9 months post-injection. (B) Wire-hang test expressed as number of falls per minute, performed 9 months after vector injection. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001 vs. PBS-injected *Agr1*<sup>-/-</sup> mice, #p<0.05, ##p<0.01, ###p<0.001 vs. PBS-injected *Agr1*<sup>+/+</sup> mice, n=3-5 mice per group). All data are shown as mean ± SEM.



**Figure S18. Selection of a liver promoter using the mSEAP reporter gene.** (A) Huh7 cells were transfected with plasmids encoding the mSEAP reporter gene under the control of the murine liver specific promoter (mLSP), the human liver specific promoter (hLSP), the ApoE-hAAT promoter or the liver-muscle tandem promoter (LiMP)<sup>41</sup>. mSEAP activity quantification in the media collected 48-hours post-transfection. The dashed line represents the level measured in non-transfected (NT) cells. (B) 6-week-old male C57BL/6 mice were injected via tail vein with 5 x 10<sup>13</sup> vg/kg of rAAV-MT encoding the mSEAP gene reporter under the control of the mLSP, hLSP or ApoE-hAAT promoters. PBS-injected mice were used as controls (C) mSEAP activity quantification in liver one month after injection. The dashed line represents the level measured in PBS-injected mice. (D) VGCN in liver one month after injection. The dashed line represents the level measured in PBS-injected mice. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01 or \*\*\*p<0.001 vs. NT cells in panel A, n = 4 replicates, and vs. PBS-injected mice in panels B and C, n=4 mice per group). All data are shown as mean ± SEM.



**Figure S19: Correction of the liver phenotype according to the correction of the glycemia.** Evaluation of serum ALT level (A) and glycogen content (B), as shown in Figure 31D and 32E respectively, function of the correction of glycemia at sacrifice. The threshold of 160 mg/dL corresponds to -2 standard deviations of the distribution of the glycemia values observed in PBS-injected *Agl*<sup>+/+</sup> mice. Statistical analyses were performed by one-way ANOVA (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001 vs. PBS-injected *Agl*<sup>-/-</sup> mice, #p<0.05, ##p<0.01, ###p<0.001 vs. PBS-injected *Agl*<sup>+/+</sup> mice, n=7-8 mice per group). All data are shown as mean ± SEM.

# CONCLUSION

## CONCLUSION

The treatment of GSDIII using an rAAV-based gene therapy approach exhibit many challenges: the size of the GDE cDNA exceeding the encapsidation limit of AAV, the requirement for simultaneous correction of several different organs, such as liver and muscles, and the transduction and correction of a chronically-damaged liver. The development of an rAAV-based gene therapy approach for this disease requires several optimizations. In this thesis, we first designed a **single vector rAAV approach to express a N-terminal truncated functional version of the GDE enzyme** in muscle and heart. We showed that the use of this rAAV vector corrected the muscle and heart phenotype in a mouse model of GSDIII, in a recently generated rat model of GSDIII and also corrected the accumulation of glycogen in absence of specific toxicity in human GSDIII myotubes derived from hiPSC. In the second part of this thesis, we characterized the hepatic phenotype of the *Agf<sup>-/-</sup>* mice and analyzed the **efficacy and limitations of a single rAAV vector approach expressing the full-length GDE** in liver of both young and old *Agf<sup>-/-</sup>* mice. We demonstrated complete but transient correction of the metabolic phenotype after intravenous injection in *Agf<sup>-/-</sup>* mice injected at an old age, with progressive loss of correction over time potentially related to immune response and/or persisting hepatocyte proliferation with vector dilution. Treatment of younger mice led to faster correction and long-term complete correction could be achieved in half of our cohort, whereas the remaining animals still displayed a progressive loss of correction starting from 6 months post-injection.

**Further optimizations are required to improve liver targeting and GDE expression**, especially the design of non-oversized vectors, such as one expressing our N-terminal truncated GDE mutant, as well as for the complete correction of both the liver and muscle disease with a single rAAV vector. On the other hand, the treatment of genetic diseases is not limited to gene therapy approaches. Many combinatorial therapies could be valuable, such as the use of orally given small soluble molecules either reducing the glycogen accumulated or said to improve tissues transduction by rAAV, such as rapamycin, recently evaluated in combination with a dual vector approach in GSDIII by our team<sup>214</sup>. Another interesting strategy would be the combination of a mRNA therapy targeting the liver, in order to correct the metabolic phenotype in childhood and maybe prevent the development of fibrosis, and a liver/muscle-targeted gene therapy approach at adulthood, once the liver has completed growth. Such combinations of treatment, although very complex to translate into the clinic, might allow the correction of the full phenotype of GSDIII.

# REFERENCES

## REFERENCES

1. Gardin, A., Remih, K., Gonzales, E., Andersson, E. R. & Strnad, P. Modern therapeutic approaches to liver-related disorders. *J. Hepatol.* **76**, 1392–1409 (2022).
2. Maestro, S., Weber, N. D., Zabaleta, N., Aldabe, R. & Gonzalez-Aseguinolaza, G. Novel vectors and approaches for gene therapy in liver diseases. *JHEP Rep.* **3**, 100300 (2021).
3. Kunz, J. B. & Kulozik, A. E. Gene Therapy of the Hemoglobinopathies. *HemaSphere* **4**, e479 (2020).
4. Fischer, A. Gene therapy for inborn errors of immunity: past, present and future. *Nat. Rev. Immunol.* (2022).
5. Biffi, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* **341**, 1233158–1233158 (2013).
6. Kuzmin, D. A. *et al.* The clinical landscape for AAV gene therapies. *Nat. Rev. Drug Discov.* **20**, 173–174 (2021).
7. Gillmore, J. D. *et al.* CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. *N. Engl. J. Med.* **385**, 493–502 (2021).
8. Kotterman, M. A., Chalberg, T. W. & Schaffer, D. V. Viral Vectors for Gene Therapy: Translational and Clinical Outlook. *Annu. Rev. Biomed. Eng.* **17**, 63–89 (2015).
9. Milani, M. *et al.* Phagocytosis-shielded lentiviral vectors improve liver gene therapy in nonhuman primates. *Sci. Transl. Med.* **11**, eaav7325 (2019).
10. Milani, M. *et al.* Liver-directed lentiviral gene therapy corrects hemophilia A mice and achieves normal-range factor VIII activity in non-human primates. *Nat. Commun.* **13**, 2454 (2022).
11. Clar, J. *et al.* Hepatic lentiviral gene transfer prevents the long-term onset of hepatic tumours of glycogen storage disease type 1a in mice. *Hum. Mol. Genet.* **24**, 2287–2296 (2015).
12. Raper, S. E. *et al.* A Pilot Study of In Vivo Liver-Directed Gene Transfer with an Adenoviral Vector in Partial Ornithine Transcarbamylase Deficiency. *Hum. Gene Ther.* **13**, 163–175 (2002).
13. Hastie, E. & Samulski, R. J. Adeno-Associated Virus at 50: A Golden Anniversary of Discovery, Research, and Gene Therapy Success—A Personal Perspective. *Hum. Gene Ther.* **26**, 257–265 (2015).
14. Wörner, T. P. *et al.* Adeno-associated virus capsid assembly is divergent and stochastic. *Nat. Commun.* **12**, 1642 (2021).
15. Cao, M., You, H. & Hermonat, P. L. The X Gene of Adeno-Associated Virus 2 (AAV2) Is Involved in Viral DNA Replication. *PLoS ONE* **9**, e104596 (2014).
16. Elmore, Z. C. *et al.* The membrane associated accessory protein is an adeno-associated viral egress factor. *Nat. Commun.* **12**, 6239 (2021).
17. Ogden, P. J., Kelsic, E. D., Sinai, S. & Church, G. M. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science* **366**, 1139–1143 (2019).
18. Pupo, A. *et al.* AAV vectors: The Rubik’s cube of human gene therapy. *Mol. Ther.* **30**, 3515–3541 (2022).
19. Govindasamy, L. *et al.* Structurally Mapping the Diverse Phenotype of Adeno-Associated Virus Serotype 4. *J. Virol.* **80**, 11556–11570 (2006).
20. Meyer, N. L. & Chapman, M. S. Adeno-associated virus (AAV) cell entry: structural insights. *Trends Microbiol.* **30**, 432–451 (2022).
21. Tseng, Y.-S. & Agbandje-McKenna, M. Mapping the AAV Capsid Host Antibody Response toward the Development of Second Generation Gene Delivery Vectors. *Front. Immunol.* **5**, (2014).
22. Wang, L. *et al.* Comparative Study of Liver Gene Transfer With AAV Vectors Based on Natural

- and Engineered AAV Capsids. *Mol. Ther.* **23**, 1877–1887 (2015).
23. Lisowski, L. *et al.* Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature* **506**, 382–386 (2014).
  24. Kondratov, O. *et al.* A comprehensive study of a 29-capsid AAV library in a non-human primate central nervous system. *Mol. Ther.* **29**, 2806–2820 (2021).
  25. Grimm, D. *et al.* In Vitro and In Vivo Gene Therapy Vector Evolution via Multispecies Interbreeding and Retargeting of Adeno-Associated Viruses. *J. Virol.* **82**, 5887–5911 (2008).
  26. Weinmann, J. *et al.* Identification of a myotropic AAV by massively parallel in vivo evaluation of barcoded capsid variants. *Nat. Commun.* **11**, 5432 (2020).
  27. Tabebordbar, M. *et al.* Directed evolution of a family of AAV capsid variants enabling potent muscle-directed gene delivery across species. *Cell* **184**, 4919–4938.e22 (2021).
  28. Lisowski, L., Tay, S. S. & Alexander, I. E. Adeno-associated virus serotypes for gene therapeutics. *Curr. Opin. Pharmacol.* **24**, 59–67 (2015).
  29. Li, C. & Samulski, R. J. Engineering adeno-associated virus vectors for gene therapy. *Nat. Rev. Genet.* **21**, 255–272 (2020).
  30. Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature* **530**, 108–112 (2016).
  31. Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* **18**, 358–378 (2019).
  32. Nicolson, S. C. & Samulski, R. J. Recombinant Adeno-Associated Virus Utilizes Host Cell Nuclear Import Machinery To Enter the Nucleus. *J. Virol.* **88**, 4132–4144 (2014).
  33. Hösel, M. *et al.* Autophagy determines efficiency of liver-directed gene therapy with adeno-associated viral vectors. *Hepatology* **66**, 252–265 (2017).
  34. Fisher, K. J. *et al.* Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *J. Virol.* **70**, 520–532 (1996).
  35. La Bella, T. *et al.* Adeno-associated virus in the liver: natural history and consequences in tumour development. *Gut* **69**, 737–747 (2020).
  36. Gil-Farina, I. *et al.* Recombinant AAV Integration Is Not Associated With Hepatic Genotoxicity in Nonhuman Primates and Patients. *Mol. Ther.* **24**, 1100–1105 (2016).
  37. McCarty, D. M. *et al.* Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther.* **10**, 2112–2118 (2003).
  38. Skopenkova, V. V., Egorova, T. V. & Bardina, M. V. Muscle-Specific Promoters for Gene Therapy. *Acta Naturae* **13**, 47–58 (2021).
  39. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
  40. Cabrera-Pérez, R. *et al.* Alpha-1-Antitrypsin Promoter Improves the Efficacy of an Adeno-Associated Virus Vector for the Treatment of Mitochondrial Neurogastrointestinal Encephalomyopathy. *Hum. Gene Ther.* **30**, 985–998 (2019).
  41. Colella, P. *et al.* AAV Gene Transfer with Tandem Promoter Design Prevents Anti-transgene Immunity and Provides Persistent Efficacy in Neonate Pompe Mice. *Mol. Ther. - Methods Clin. Dev.* **12**, 85–101 (2019).
  42. Lim, J.-A., Kishnani, P. S. & Sun, B. Suppression of pullulanase-induced cytotoxic T cell response with a dual promoter in GSD IIIa mice. *JCI Insight* **7**, e152970 (2022).
  43. Hanson, G. & Collier, J. Codon optimality, bias and usage in translation and mRNA decay. *Nat.*

*Rev. Mol. Cell Biol.* **19**, 20–30 (2018).

44. Liguore, W. A. *et al.* AAV-PHP.B Administration Results in a Differential Pattern of CNS Biodistribution in Non-human Primates Compared with Mice. *Mol. Ther.* **27**, 2018–2037 (2019).
45. Gao, G. *et al.* Clades of Adeno-Associated Viruses Are Widely Disseminated in Human Tissues. *J. Virol.* **78**, 6381–6388 (2004).
46. Gao, G.-P. *et al.* Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci.* **99**, 11854–11859 (2002).
47. Calcedo, R. *et al.* Adeno-Associated Virus Antibody Profiles in Newborns, Children, and Adolescents. *Clin. Vaccine Immunol.* **18**, 1586–1588 (2011).
48. Boutin, S. *et al.* Prevalence of Serum IgG and Neutralizing Factors Against Adeno-Associated Virus (AAV) Types 1, 2, 5, 6, 8, and 9 in the Healthy Population: Implications for Gene Therapy Using AAV Vectors. *Hum. Gene Ther.* **21**, 704–712 (2010).
49. Bello, A. *et al.* Isolation and evaluation of novel adeno-associated virus sequences from porcine tissues. *Gene Ther.* **16**, 1320–1328 (2009).
50. Smith, L. J. *et al.* Stem cell-derived clade F AAVs mediate high-efficiency homologous recombination-based genome editing. *Proc. Natl. Acad. Sci.* **115**, E7379–E7388 (2018).
51. Büning, H. & Srivastava, A. Capsid Modifications for Targeting and Improving the Efficacy of AAV Vectors. *Mol. Ther. - Methods Clin. Dev.* **12**, 248–265 (2019).
52. Zhong, L. *et al.* Next generation of adeno-associated virus 2 vectors: Point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc. Natl. Acad. Sci.* **105**, 7827–7832 (2008).
53. Zinn, E. *et al.* In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. *Cell Rep.* **12**, 1056–1068 (2015).
54. Maheshri, N., Koerber, J. T., Kaspar, B. K. & Schaffer, D. V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat. Biotechnol.* **24**, 198–204 (2006).
55. Colella, P., Ronzitti, G. & Mingozzi, F. Emerging Issues in AAV-Mediated In Vivo Gene Therapy. *Mol. Ther. - Methods Clin. Dev.* **8**, 87–104 (2018).
56. Perocheau, D. P. *et al.* Age-Related Seroprevalence of Antibodies Against AAV-LK03 in a UK Population Cohort. *Hum. Gene Ther.* **30**, 79–87 (2019).
57. Manno, C. S. *et al.* Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat. Med.* **12**, 342–347 (2006).
58. Scallan, C. D. *et al.* Human immunoglobulin inhibits liver transduction by AAV vectors at low AAV2 neutralizing titers in SCID mice. *Blood* **107**, 1810–1817 (2006).
59. George, L. A. *et al.* Long-Term Follow-Up of the First in Human Intravascular Delivery of AAV for Gene Transfer: AAV2-hFIX16 for Severe Hemophilia B. *Mol. Ther.* **28**, 2073–2082 (2020).
60. Gross, D.-A., Tedesco, N., Leborgne, C. & Ronzitti, G. Overcoming the Challenges Imposed by Humoral Immunity to AAV Vectors to Achieve Safe and Efficient Gene Transfer in Seropositive Patients. *Front. Immunol.* **13**, 857276 (2022).
61. György, B., Fitzpatrick, Z., Crommentuijn, M. H. W., Mu, D. & Maguire, C. A. Naturally enveloped AAV vectors for shielding neutralizing antibodies and robust gene delivery in vivo. *Biomaterials* **35**, 7598–7609 (2014).
62. Mingozzi, F. *et al.* Overcoming Preexisting Humoral Immunity to AAV Using Capsid Decoys. *Sci. Transl. Med.* **5**, (2013).
63. Chicoine, L. *et al.* Plasmapheresis Eliminates the Negative Impact of AAV Antibodies on Microdystrophin Gene Expression Following Vascular Delivery. *Mol. Ther.* **22**, 338–347 (2014).

64. Leborgne, C. *et al.* IgG-cleaving endopeptidase enables in vivo gene therapy in the presence of anti-AAV neutralizing antibodies. *Nat. Med.* **26**, 1096–1101 (2020).
65. Ilyinskii, P. O. *et al.* Enhancement of liver-directed transgene expression at initial and repeat doses of AAV vectors admixed with ImmTOR nanoparticles. *Sci. Adv.* **7**, eabd0321 (2021).
66. Mingozi, F. *et al.* Pharmacological Modulation of Humoral Immunity in a Nonhuman Primate Model of AAV Gene Transfer for Hemophilia B. *Mol. Ther.* **20**, 1410–1416 (2012).
67. Wu, Z., Yang, H. & Colosi, P. Effect of Genome Size on AAV Vector Packaging. *Mol. Ther.* **18**, 80–86 (2010).
68. Trapani, I., Tornabene, P. & Auricchio, A. Large gene delivery to the retina with AAV vectors: are we there yet? *Gene Ther.* **28**, 220–222 (2021).
69. Vidal, P. *et al.* Rescue of GSDIII Phenotype with Gene Transfer Requires Liver- and Muscle-Targeted GDE Expression. *Mol. Ther.* **26**, 890–901 (2018).
70. Kodippili, K. *et al.* Dual AAV Gene Therapy for Duchenne Muscular Dystrophy with a 7-kb *Mini-Dystrophin* Gene in the Canine Model. *Hum. Gene Ther.* **29**, 299–311 (2018).
71. Potter, R. A. *et al.* Systemic Delivery of Dysferlin Overlap Vectors Provides Long-Term Gene Expression and Functional Improvement for Dysferlinopathy. *Hum. Gene Ther.* **29**, 749–762 (2018).
72. Padula, A. *et al.* Full-length ATP7B reconstituted through protein trans-splicing corrects Wilson disease in mice. *Mol. Ther. - Methods Clin. Dev.* **26**, 495–504 (2022).
73. Tornabene, P. *et al.* Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina. *Sci. Transl. Med.* **11**, eaav4523 (2019).
74. Willett, K. & Bennett, J. Immunology of AAV-Mediated Gene Transfer in the Eye. *Front. Immunol.* **4**, (2013).
75. Rangarajan, S. *et al.* AAV5–Factor VIII Gene Transfer in Severe Hemophilia A. *N. Engl. J. Med.* **377**, 2519–2530 (2017).
76. Manini, A., Abati, E., Nuredini, A., Corti, S. & Comi, G. P. Adeno-Associated Virus (AAV)-Mediated Gene Therapy for Duchenne Muscular Dystrophy: The Issue of Transgene Persistence. *Front. Neurol.* **12**, 814174 (2022).
77. Murillo, O. *et al.* Liver Expression of a MiniATP7B Gene Results in Long-Term Restoration of Copper Homeostasis in a Wilson Disease Model in Mice. *Hepatology* (2019).
78. Mingozi, F. & High, K. A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **122**, 23–36 (2013).
79. Finn, J. D. *et al.* Proteasome Inhibitors Decrease AAV2 Capsid derived Peptide Epitope Presentation on MHC Class I Following Transduction. *Mol. Ther.* **18**, 135–142 (2010).
80. Nathwani, A. C. *et al.* Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N. Engl. J. Med.* **371**, 1994–2004 (2014).
81. Nathwani, A. C. Gene therapy for hemophilia. *Hematology* **2019**, 1–8 (2019).
82. Nathwani, A. C. *et al.* Adenovirus-Associated Virus Vector–Mediated Gene Transfer in Hemophilia B. *N. Engl. J. Med.* **365**, 2357–2365 (2011).
83. Lek, A., Atas, E., Hesterlee, S. E., Byrne, B. J. & Bönnemann, C. G. Meeting Report: 2022 Muscular Dystrophy Association Summit on ‘Safety and Challenges in Gene Transfer Therapy’. *J. Neuromuscul. Dis.* **10**, 327–336 (2023).
84. Bönnemann, C. G. *et al.* Dystrophin Immunity after Gene Therapy for Duchenne’s Muscular Dystrophy. *N. Engl. J. Med.* **388**, 2294–2296 (2023).
85. Collaud, F. *et al.* Preclinical Development of an AAV8-hUGT1A1 Vector for the Treatment of Crigler-Najjar Syndrome. *Mol. Ther. - Methods Clin. Dev.* **12**, 157–174 (2019).

86. Chand, D. *et al.* Hepatotoxicity following administration of onasemnogene abeparvovec (AVXS-101) for the treatment of spinal muscular atrophy. *J. Hepatol.* **74**, 560–566 (2021).
87. Wilson, J. M. & Flotte, T. R. Moving Forward After Two Deaths in a Gene Therapy Trial of Myotubular Myopathy. *Hum. Gene Ther.* **31**, 695–696 (2020).
88. D’Amico, A. *et al.* Hepatobiliary disease in XLMTM: a common comorbidity with potential impact on treatment strategies. *Orphanet J. Rare Dis.* **16**, 425 (2021).
89. Ertl, H. C. J. Immunogenicity and toxicity of AAV gene therapy. *Front. Immunol.* **13**, 975803 (2022).
90. Palazzi, X. *et al.* Biodistribution and Tolerability of AAV-PHP.B-CBh- *SMN1* in Wistar Han Rats and Cynomolgus Macaques Reveal Different Toxicologic Profiles. *Hum. Gene Ther.* **33**, 175–187 (2022).
91. Guillou, J. *et al.* Fatal thrombotic microangiopathy case following adeno-associated viral *SMN* gene therapy. *Blood Adv.* **6**, 4266–4270 (2022).
92. George, J. N. & Nester, C. M. Syndromes of Thrombotic Microangiopathy. *N. Engl. J. Med.* **371**, 654–666 (2014).
93. Konkle, B. A. *et al.* BAX 335 hemophilia B gene therapy clinical trial results - potential impact of CpG sequences on gene expression. *Blood* (2020).
94. Pipe, S. *et al.* ‘101HEMB01 Is a Phase 1/2 Open-Label, Single Ascending Dose-Finding Trial of DTX101 (AAVrh10FIX) in Patients with Moderate/Severe Hemophilia B That Demonstrated Meaningful but Transient Expression of Human Factor IX (hFIX)’ (Poster). *Blood 130 (Supplement 1): 3331* (2017).
95. Faust, S. M. *et al.* CpG-depleted adeno-associated virus vectors evade immune detection. *J. Clin. Invest.* **123**, 2994–3001 (2013).
96. Nault, J.-C. *et al.* Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat. Genet.* **47**, 1187–1193 (2015).
97. Logan, G. J. *et al.* Identification of liver-specific enhancer–promoter activity in the 3′ untranslated region of the wild-type AAV2 genome. *Nat. Genet.* **49**, 1267–1273 (2017).
98. Dalwadi, D. A. *et al.* AAV integration in human hepatocytes. *Mol. Ther.* (2021).
99. Nguyen, G. N. *et al.* A long-term study of AAV gene therapy in dogs with hemophilia A identifies clonal expansions of transduced liver cells. *Nat. Biotechnol.* **39**, 47–55 (2021).
100. Bell, P. *et al.* No Evidence for Tumorigenesis of AAV Vectors in a Large-Scale Study in Mice. *Mol. Ther.* **12**, 299–306 (2005).
101. Li, H. *et al.* Assessing the potential for AAV vector genotoxicity in a murine model. *Blood* **117**, 3311–3319 (2011).
102. Donsante, A. *et al.* AAV Vector Integration Sites in Mouse Hepatocellular Carcinoma. *Science* **317**, 477–477 (2007).
103. Chandler, R. J. *et al.* Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy. *J. Clin. Invest.* **125**, 870–880 (2015).
104. Wang, P.-R. *et al.* Induction of hepatocellular carcinoma by in vivo gene targeting. *Proc. Natl. Acad. Sci.* **109**, 11264–11269 (2012).
105. Dill, T. & Naya, F. A Hearty Dose of Noncoding RNAs: The Imprinted DLK1-DIO3 Locus in Cardiac Development and Disease. *J. Cardiovasc. Dev. Dis.* **5**, 37 (2018).
106. Ferla, R. *et al.* Low incidence of hepatocellular carcinoma in mice and cats treated with systemic adeno-associated viral vectors. *Mol. Ther. - Methods Clin. Dev.* **20**, 247–257 (2021).
107. Kaiser, J. Experimental gene therapy for hemophilia probably did not cause patient’s liver tumor. *Science* (2021).
108. Eipel, C. Regulation of hepatic blood flow: The hepatic arterial buffer response revisited. *World*

*J. Gastroenterol.* **16**, 6046 (2010).

109. Shetty, S., Lalor, P. F. & Adams, D. H. Liver sinusoidal endothelial cells — gatekeepers of hepatic immunity. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 555–567 (2018).
110. Ehrlich, A., Duche, D., Ouedraogo, G. & Nahmias, Y. Challenges and Opportunities in the Design of Liver-on-Chip Microdevices. *Annu. Rev. Biomed. Eng.* **21**, 219–239 (2019).
111. Jackson, C. L. & McNiven, M. A. The Hepatocellular Secretory Pathway. in *The Liver* (eds. Arias, I. M. et al.) 75–85 (Wiley, 2020).
112. Costa-Verdera, H. *et al.* Hepatic expression of GAA results in enhanced enzyme bioavailability in mice and non-human primates. *Nat. Commun.* **12**, 6393 (2021).
113. Zhukouskaya, V. V. *et al.* A novel therapeutic strategy for skeletal disorders: Proof of concept of gene therapy for X-linked hypophosphatemia. *Sci. Adv.* **7**, eabj5018 (2021).
114. Zheng, M. & Tian, Z. Liver-Mediated Adaptive Immune Tolerance. *Front. Immunol.* **10**, 2525 (2019).
115. Bénéchet, A. P. *et al.* Dynamics and genomic landscape of CD8<sup>+</sup> T cells undergoing hepatic priming. *Nature* **574**, 200–205 (2019).
116. Moini, M. Review on immunosuppression in liver transplantation. *World J. Hepatol.* **7**, 1355 (2015).
117. Thomson, A. W., Vionnet, J. & Sanchez-Fueyo, A. Understanding, predicting and achieving liver transplant tolerance: from bench to bedside. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 719–739 (2020).
118. Bartolo, L. *et al.* Dual muscle-liver transduction imposes immune tolerance for muscle transgene engraftment despite preexisting immunity. *JCI Insight* **4**, e127008 (2019).
119. Poupiot, J. *et al.* Role of Regulatory T Cell and Effector T Cell Exhaustion in Liver-Mediated Transgene Tolerance in Muscle. *Mol. Ther. - Methods Clin. Dev.* **15**, 83–100 (2019).
120. Thomsen, G. *et al.* Biodistribution of onasemnogene abeparvovec DNA, mRNA and SMN protein in human tissue. *Nat. Med.* **27**, 1701–1711 (2021).
121. Lachmann, R. H. Enzyme replacement therapy for lysosomal storage diseases. *Curr. Opin. Pediatr.* **23**, 588–593 (2011).
122. Chowdary, P. *et al.* Phase 1–2 Trial of AAVS3 Gene Therapy in Patients with Hemophilia B. *N. Engl. J. Med.* **387**, 237–247 (2022).
123. Ljung, R. *et al.* Inhibitors in haemophilia A and B: Management of bleeds, inhibitor eradication and strategies for difficult-to-treat patients. *Eur. J. Haematol.* **102**, 111–122 (2019).
124. de Vries, J. M. *et al.* Pompe disease in adulthood: effects of antibody formation on enzyme replacement therapy. *Genet. Med.* **19**, 90–97 (2017).
125. George, L. A. *et al.* Hemophilia B Gene Therapy with a High-Specific-Activity Factor IX Variant. *N. Engl. J. Med.* **377**, 2215–2227 (2017).
126. Pipe, S. W. *et al.* Gene Therapy with Etranacogene Dezaparvovec for Hemophilia B. *N. Engl. J. Med.* **388**, 706–718 (2023).
127. Ozelo, M. C. *et al.* Valoctocogene Roxaparvovec Gene Therapy for Hemophilia A. *N. Engl. J. Med.* **386**, 1013–1025 (2022).
128. Brunetti-Pierri, N. *et al.* Liver-Directed Adeno-Associated Virus–Mediated Gene Therapy for Mucopolysaccharidosis Type VI. *NEJM Evid.* **1**, (2022).
129. Puzzo, F. *et al.* Rescue of Pompe disease in mice by AAV-mediated liver delivery of secretable acid  $\alpha$ -glucosidase. *Sci. Transl. Med.* **9**, eaam6375 (2017).
130. Strauss, K. A. *et al.* Crigler-Najjar Syndrome Type 1: Pathophysiology, Natural History, and Therapeutic Frontier. *Hepatology* **71**, 1923–1939 (2020).

131. Scharre, S. *et al.* Predicting the disease severity in male individuals with ornithine transcarbamylase deficiency. *Ann. Clin. Transl. Neurol.* **9**, 1715–1726 (2022).
132. Ronzitti, G. *et al.* A translationally optimized AAV-UGT1A1 vector drives safe and long-lasting correction of Crigler-Najjar syndrome. *Mol. Ther. - Methods Clin. Dev.* **3**, 16049 (2016).
133. Lee, Y. M. *et al.* Long-term safety and efficacy of AAV gene therapy in the canine model of glycogen storage disease type Ia. *J. Inherit. Metab. Dis.* **41**, 977–984 (2018).
134. Strnad, P. *et al.* Fazirsiran for Liver Disease Associated with Alpha<sub>1</sub>-Antitrypsin Deficiency. *N. Engl. J. Med.* **387**, 514–524 (2022).
135. Zabaleta, N. *et al.* CRISPR/Cas9-mediated glycolate oxidase disruption is an efficacious and safe treatment for primary hyperoxaluria type I. *Nat. Commun.* **9**, 5454 (2018).
136. Pankowicz, F. P. *et al.* Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia. *Nat. Commun.* **7**, 12642 (2016).
137. Barzi, M. *et al.* Rescue of glutaric aciduria type I in mice by liver-directed therapies. *Sci. Transl. Med.* **15**, eadf4086 (2023).
138. Cunningham, S. C., Dane, A. P., Spinoulas, A. & Alexander, I. E. Gene Delivery to the Juvenile Mouse Liver Using AAV2/8 Vectors. *Mol. Ther.* **16**, 1081–1088 (2008).
139. Coppoletta, J. & Wolbach, S. Body Length and Organ weights of infants and children : A study of the body length and normal weights of the more vital organs of the body between birth and twelve years of age. *Am J Pathol* **9**, 55–70 (1932).
140. Raynaud, P., Carpentier, R., Antoniou, A. & Lemaigre, F. P. Biliary differentiation and bile duct morphogenesis in development and disease. *Int. J. Biochem. Cell Biol.* **43**, 245–256 (2011).
141. Gordillo, M., Evans, T. & Gouon-Evans, V. Orchestrating liver development. *Development* **142**, 2094–2108 (2015).
142. Bosma, P. J., Wits, M. & Oude-Elferink, R. P. J. Gene Therapy for Progressive Familial Intrahepatic Cholestasis: Current Progress and Future Prospects. *Int. J. Mol. Sci.* **22**, 273 (2020).
143. Jauze, L., Monteillet, L., Mithieux, G., Rajas, F. & Ronzitti, G. Challenges of Gene Therapy for the Treatment of Glycogen Storage Diseases Type I and Type III. *Hum. Gene Ther.* **30**, 1263–1273 (2019).
144. Murillo, O. *et al.* Long-term metabolic correction of Wilson’s disease in a murine model by gene therapy. *J. Hepatol.* **64**, 419–426 (2016).
145. Aronson, S. J. *et al.* Liver-directed gene therapy results in long-term correction of progressive familial intrahepatic cholestasis type 3 in mice. *J. Hepatol.* **71**, 153–162 (2019).
146. Weber, N. D. *et al.* Gene therapy for progressive familial intrahepatic cholestasis type 3 in a clinically relevant mouse model. *Nat. Commun.* **10**, 5694 (2019).
147. Siew, S. M. *et al.* Prevention of Cholestatic Liver Disease and Reduced Tumorigenicity in a Murine Model of PFIC Type 3 Using Hybrid AAV-piggyBac Gene Therapy. *Hepatology* **70**, 2047–2061 (2019).
148. Weber, N. D. *et al.* Rescue of infant progressive familial intrahepatic cholestasis type 3 mice by repeated dosing of AAV gene therapy. *JHEP Rep.* 100713 (2023).
149. López-Terrada, D. & Finegold, M. J. Tumors of the liver. in *Liver Disease in Children* (eds. Suchy, F. J., Sokol, R. J. & Balistreri, W. F.) 728–759 (Cambridge University Press, 2014).
150. Chandler, R. J. *et al.* Promoterless, Nuclease-Free Genome Editing Confers a Growth Advantage for Corrected Hepatocytes in Mice With Methylmalonic Acidemia. *Hepatology* **73**, 2223–2237 (2021).
151. Cunningham, S. C. *et al.* Modeling correction of severe urea cycle defects in the growing murine liver using a hybrid recombinant adeno-associated virus/piggyBac transposase gene delivery system. *Hepatology* **62**, 417–428 (2015).

152. Qiu, M. *et al.* Lipid nanoparticle-mediated codelivery of Cas9 mRNA and single-guide RNA achieves liver-specific in vivo genome editing of *Angptl3*. *Proc. Natl. Acad. Sci.* **118**, e2020401118 (2021).
153. Musunuru, K. *et al.* In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434 (2021).
154. Rothgangl, T. *et al.* In vivo adenine base editing of PCSK9 in macaques reduces LDL cholesterol levels. *Nat. Biotechnol.* **39**, 949–957 (2021).
155. Li, N. *et al.* CRISPR/Cas9-Mediated Gene Correction in Newborn Rabbits with Hereditary Tyrosinemia Type I. *Mol. Ther.* **29**, 1001–1015 (2021).
156. Song, C.-Q. *et al.* Adenine base editing in an adult mouse model of tyrosinaemia. *Nat. Biomed. Eng.* **4**, 125–130 (2019).
157. Böck, D. *et al.* In vivo prime editing of a metabolic liver disease in mice. *Sci. Transl. Med.* **14**, eabl9238 (2022).
158. Darras, B. T., Menache-Starobinski, C. C., Hinton, V. & Kunkel, L. M. Dystrophinopathies. in *Neuromuscular Disorders of Infancy, Childhood, and Adolescence* 551–592 (Elsevier, 2015).
159. Li, H. *et al.* Systemic AAV9.BVES delivery ameliorates muscular dystrophy in a mouse model of LGMDR25. *Mol. Ther.* S1525001622006700 (2022).
160. Elangkovan, N. & Dickson, G. Gene Therapy for Duchenne Muscular Dystrophy. *J. Neuromuscul. Dis.* **8**, S303–S316 (2021).
161. Lek, A. *et al.* *Unexpected Death of a Duchenne Muscular Dystrophy Patient in an N-of-1 Trial of rAAV9-delivered CRISPR-transactivator.* (2023) doi:10.1101/2023.05.16.23289881.
162. Hakim, C. H. *et al.* A Five-Repeat Micro-Dystrophin Gene Ameliorated Dystrophic Phenotype in the Severe DBA/2J-mdx Model of Duchenne Muscular Dystrophy. *Mol. Ther. - Methods Clin. Dev.* **6**, 216–230 (2017).
163. Bourdon, A. *et al.* Evaluation of the dystrophin carboxy-terminal domain for micro-dystrophin gene therapy in cardiac and skeletal muscles in the DMDmdx rat model. *Gene Ther.* **29**, 520–535 (2022).
164. Le Guiner, C. *et al.* Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. *Nat. Commun.* **8**, 16105 (2017).
165. Xu, L. *et al.* Adeno-associated Virus 9 Mediated FKRP Gene Therapy Restores Functional Glycosylation of  $\alpha$ -dystroglycan and Improves Muscle Functions. *Mol. Ther.* **21**, 1832–1840 (2013).
166. Sahenk, Z. *et al.* Systemic delivery of AAVrh74.tMCK.hCAPN3 rescues the phenotype in a mouse model for LGMD2A/R1. *Mol. Ther. - Methods Clin. Dev.* **22**, 401–414 (2021).
167. McKee, K. K. & Yurchenco, P. D. Amelioration of muscle and nerve pathology of Lama2-related dystrophy by AAV9-laminin- $\alpha$ LN linker protein. *JCI Insight* **7**, e158397 (2022).
168. Akman, H. O., Oldfors, A. & DiMauro, S. Glycogen Storage Diseases of Muscle. in *Neuromuscular Disorders of Infancy, Childhood, and Adolescence* 735–760 (Elsevier, 2015).
169. Laforêt, P. *et al.* Deep morphological analysis of muscle biopsies from type III glycogenesis (GSDIII), debranching enzyme deficiency, revealed stereotyped vacuolar myopathy and autophagy impairment. *Acta Neuropathol. Commun.* **7**, 167 (2019).
170. Cagin, U. *et al.* Rescue of Advanced Pompe Disease in Mice with Hepatic Expression of Secretable Acid  $\alpha$ -Glucosidase. *Mol. Ther.* **28**, 2056–2072 (2020).
171. Childers, M. K. *et al.* Gene Therapy Prolongs Survival and Restores Function in Murine and Canine Models of Myotubular Myopathy. *Sci. Transl. Med.* **6**, (2014).
172. Shieh, P. *et al.* 'ASPIRO gene therapy trial in X-Linked Myotubular Myopathy (XLMTM): Update on preliminary efficacy and safety findings' [Abstract OP018]. *Genet. Med.* **24**, S350 (2022).

173. Sellier, P. *et al.* Muscle-specific, liver-detargeted ADENO-ASSOCIATED VIRUS gene therapy rescues Pompe phenotype in adult and neonate *GAA*<sup>-/-</sup> mice. *J. Inherit. Metab. Dis.* jimd.12625 (2023).
174. Huichalaf, C. *et al.* In vivo overexpression of frataxin causes toxicity mediated by iron-sulfur cluster deficiency. *Mol. Ther. - Methods Clin. Dev.* **24**, 367–378 (2022).
175. Belbellaa, B., Reutenauer, L., Messaddeq, N., Monassier, L. & Puccio, H. High Levels of Frataxin Overexpression Lead to Mitochondrial and Cardiac Toxicity in Mouse Models. *Mol. Ther. - Methods Clin. Dev.* **19**, 120–138 (2020).
176. *Inborn metabolic diseases: diagnosis and treatment.* (Springer, 2012).
177. Sentner, C. P. *et al.* Glycogen storage disease type III: diagnosis, genotype, management, clinical course and outcome. *J. Inherit. Metab. Dis.* **39**, 697–704 (2016).
178. Adeva-Andany, M. M., González-Lucán, M., Donapetry-García, C., Fernández-Fernández, C. & Ameneiros-Rodríguez, E. Glycogen metabolism in humans. *BBA Clin.* **5**, 85–100 (2016).
179. Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D. & Tagliabracci, V. S. Glycogen and its metabolism: some new developments and old themes. *Biochem. J.* **441**, 763–787 (2012).
180. Berg, J. M., Tymoczko, J. L., Stryer, L. & Stryer, L. *Glycogen Metabolism.* In: *Biochemistry.* (W.H. Freeman, 2007).
181. Bhagavan, N. V. Carbohydrate Metabolism II: Gluconeogenesis, Glycogen Synthesis and Breakdown, and Alternative Pathways. in *Medical Biochemistry* 275–305 (Elsevier, 2002). doi:10.1016/B978-012095440-7/50017-2.
182. Soty, M., Gautier-Stein, A., Rajas, F. & Mithieux, G. Gut-Brain Glucose Signaling in Energy Homeostasis. *Cell Metab.* **25**, 1231–1242 (2017).
183. Kassiotis, C., Rajabi, M. & Taegtmeier, H. Metabolic Reserve of the Heart: The Forgotten Link Between Contraction and Coronary Flow. *Prog. Cardiovasc. Dis.* **51**, 74–88 (2008).
184. Gentry, M. S., Guinovart, J. J., Minassian, B. A., Roach, P. J. & Serratos, J. M. Lafora disease offers a unique window into neuronal glycogen metabolism. *J. Biol. Chem.* **293**, 7117–7125 (2018).
185. Sun, R. C. *et al.* Brain glycogen serves as a critical glucosamine cache required for protein glycosylation. *Cell Metab.* **33**, 1404-1417.e9 (2021).
186. Onkar, A., Sheshadri, D. & Ganesh, S. Glycogen: the missing link in neuronal autophagy? *Autophagy* **16**, 2102–2104 (2020).
187. Rajas, F., Gautier-Stein, A. & Mithieux, G. Glucose-6 Phosphate, a Central Hub for Liver Carbohydrate Metabolism. *Metabolites* **9**, 282 (2019).
188. Seranova, E. *et al.* Dysregulation of autophagy as a common mechanism in lysosomal storage diseases. *Essays Biochem.* **61**, 733–749 (2017).
189. Halaby, C. A. *et al.* Liver fibrosis during clinical ascertainment of glycogen storage disease type III: a need for improved and systematic monitoring. *Genet. Med.* **21**, 2686–2694 (2019).
190. Hijazi, G. *et al.* A retrospective longitudinal study and comprehensive review of adult patients with glycogen storage disease type III. *Mol. Genet. Metab. Rep.* **29**, 100821 (2021).
191. Shen, J. *et al.* Mutations in exon 3 of the glycogen debranching enzyme gene are associated with glycogen storage disease type III that is differentially expressed in liver and muscle. *J. Clin. Invest.* **98**, 352–357 (1996).
192. Michon, C.-C. *et al.* Cognitive profile of patients with glycogen storage disease type III: a clinical description of seven cases. *J. Inherit. Metab. Dis.* **38**, 573–580 (2015).
193. Kishnani, P. S. *et al.* Glycogen Storage Disease Type III diagnosis and management guidelines. *Genet. Med.* **12**, 446–463 (2010).
194. Valayannopoulos, V. *et al.* Successful Treatment of Severe Cardiomyopathy in Glycogen

- Storage Disease Type III With D,L-3-Hydroxybutyrate, Ketogenic and High-Protein Diet. *Pediatr. Res.* **70**, 638–641 (2011).
195. Zmasek, C. M. & Godzik, A. Phylogenomic analysis of glycogen branching and debranching enzymatic duo. *BMC Evol. Biol.* **14**, 183 (2014).
196. Zhai, L., Feng, L., Xia, L., Yin, H. & Xiang, S. Crystal structure of glycogen debranching enzyme and insights into its catalysis and disease-causing mutations. *Nat. Commun.* **7**, 11229 (2016).
197. Shen, M., Gong, X. & Xiang, S. Crystal structures of glycogen-debranching enzyme mutants in complex with oligosaccharides. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **77**, 420–426 (2021).
198. Cheng, A., Zhang, M., Okubo, M., Omichi, K. & Saltiel, A. R. Distinct mutations in the glycogen debranching enzyme found in glycogen storage disease type III lead to impairment in diverse cellular functions. *Hum. Mol. Genet.* **18**, 2045–2052 (2009).
199. Liu, K.-M., Wu, J.-Y. & Chen, Y.-T. Mouse model of glycogen storage disease type III. *Mol. Genet. Metab.* **111**, 467–476 (2014).
200. Pagliarani, S. *et al.* Glycogen storage disease type III: A novel Agl knockout mouse model. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1842**, 2318–2328 (2014).
201. Yi, H. *et al.* Characterization of a canine model of glycogen storage disease type IIIa. *Dis. Model. Mech.* **5**, 804–811 (2012).
202. Pursell, N. *et al.* Inhibition of Glycogen Synthase II with RNAi Prevents Liver Injury in Mouse Models of Glycogen Storage Diseases. *Mol. Ther.* **26**, 1771–1782 (2018).
203. Mederacke, I. *et al.* Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. *Nat. Commun.* **4**, 2823 (2013).
204. Ma, S. *et al.* Characterisation of Lamp2-deficient rats for potential new animal model of Danon disease. *Sci. Rep.* **8**, 6932 (2018).
205. Nemazanyy, I. *et al.* Defects of Vps15 in skeletal muscles lead to autophagic vacuolar myopathy and lysosomal disease. *EMBO Mol. Med.* **5**, 870–890 (2013).
206. Yi, H. *et al.* Correction of glycogen storage disease type III with rapamycin in a canine model. *J. Mol. Med.* **92**, 641–650 (2014).
207. Sun, B. *et al.* Alglucosidase alfa enzyme replacement therapy as a therapeutic approach for glycogen storage disease type III. *Mol. Genet. Metab.* **108**, 145–147 (2013).
208. Löser, P., Jennings, G. S., Strauss, M. & Sandig, V. Reactivation of the Previously Silenced Cytomegalovirus Major Immediate-Early Promoter in the Mouse Liver: Involvement of NFκB. *J. Virol.* **72**, 180–190 (1998).
209. Lim, J.-A., Choi, S. J., Gao, F., Kishnani, P. S. & Sun, B. A Novel Gene Therapy Approach for GSD III Using an AAV Vector Encoding a Bacterial Glycogen Debranching Enzyme. *Mol. Ther. - Methods Clin. Dev.* **18**, 240–249 (2020).
210. Berling, É. *et al.* Narrative review of glycogen storage disorder type III with a focus on neuromuscular, cardiac and therapeutic aspects. *J. Inherit. Metab. Dis.* 521–533 (2021).
211. Mendell, J. R. *et al.* Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017).
212. Von Drygalski, A. *et al.* Etranacogene dezaparvovec (AMT-061 phase 2b): normal/near normal FIX activity and bleed cessation in hemophilia B. *Blood Adv.* **3**, 3241–3247 (2019).
213. Russell, S. *et al.* Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *The Lancet* **390**, 849–860 (2017).
214. Jauze, L. *et al.* Dual AAV gene therapy in combination with rapamycin rescues muscular

lysosomal pathway impairment and reverse hepatic inflammation response in a mouse model of GSDIII [Manuscript in preparation].

215. Mingozi, F., Ronzitti, G. & Vidal, P. Treatment of glycogen storage disease type III (Patent WO2018162748 A1).
216. Potter, R. A. *et al.* Dose-Escalation Study of Systemically Delivered rAAVrh74.MHCK7.microdystrophin in the *mdx* Mouse Model of Duchenne Muscular Dystrophy. *Hum. Gene Ther.* **32**, 375–389 (2021).
217. Zhang, P. *et al.* Immunodominant Liver-Specific Expression Suppresses Transgene-Directed Immune Responses in Murine Pompe Disease. *Hum. Gene Ther.* **23**, 460–472 (2012).
218. Rossiaud, L. *et al.* Pathological modeling of glycogen storage disease type III with CRISPR/Cas9 edited human pluripotent stem cells. *Front. Cell Dev. Biol.* **11**, 1163427 (2023).
219. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
220. David, A., Islam, S., Tankhilevich, E. & Sternberg, M. J. E. The AlphaFold Database of Protein Structures: A Biologist's Guide. *J. Mol. Biol.* **434**, 167336 (2022).
221. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
222. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
223. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
224. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).
225. Wang, B. *et al.* Construction and analysis of compact muscle-specific promoters for AAV vectors. *Gene Ther.* **15**, 1489–1499 (2008).
226. Lu, C. *et al.* Spectrum of AGL mutations in Chinese patients with glycogen storage disease type III: identification of 31 novel mutations. *J. Hum. Genet.* **61**, 641–645 (2016).
227. Decostre, V. *et al.* Long term longitudinal study of muscle function in patients with glycogen storage disease type IIIa. *Mol. Genet. Metab.* **122**, 108–116 (2017).
228. Perveen, S. *et al.* Spectrum of amyloglucosidase mutations in Asian Indian patients with Glycogen storage disease type III. *Am. J. Med. Genet. A.* **182**, 1190–1200 (2020).
229. Shaiu, W.-L., Kishnani, P. S., Shen, J., Liu, H.-M. & Chen, Y.-T. Genotype-Phenotype Correlation in Two Frequent Mutations and Mutation Update in Type III Glycogen Storage Disease. *Mol. Genet. Metab.* **69**, 16–23 (2000).
230. Liang, Y. *et al.* Genotypic and clinical analysis of 49 Chinese children with hepatic glycogen storage diseases. *Mol. Genet. Genomic Med.* **8**, (2020).
231. Ko, J. S. *et al.* A mutation analysis of the AGL gene in Korean patients with glycogen storage disease type III. *J. Hum. Genet.* **59**, 42–45 (2014).
232. Nazari, F. *et al.* Distinct Clinical and Genetic Findings in Iranian Patients With Glycogen Storage Disease Type 3. *J. Clin. Neuromuscul. Dis.* **19**, 203–210 (2018).
233. Kumar, T. V. *et al.* Molecular and clinical profiling in a large cohort of Asian Indians with glycogen storage disorders. *PLOS ONE* **17**, e0270373 (2022).
234. Rouillon, J. *et al.* Serum proteomic profiling reveals fragments of MYOM3 as potential biomarkers for monitoring the outcome of therapeutic interventions in muscular dystrophies. *Hum. Mol. Genet.* **24**, 4916–4932 (2015).

235. Shi, Y., Inoue, H., Wu, J. C. & Yamanaka, S. Induced pluripotent stem cell technology: a decade of progress. *Nat. Rev. Drug Discov.* **16**, 115–130 (2017).
236. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
237. Yiu, W. H. *et al.* Complete Normalization of Hepatic G6PC Deficiency in Murine Glycogen Storage Disease Type Ia Using Gene Therapy. *Mol. Ther.* **18**, 1076–1084 (2010).
238. Mutel, E. *et al.* Targeted deletion of liver glucose-6 phosphatase mimics glycogen storage disease type 1a including development of multiple adenomas. *J. Hepatol.* **54**, 529–537 (2011).
239. Butterfield, J. S. S. *et al.* IL-15 blockade and rapamycin rescue multifactorial loss of factor VIII from AAV-transduced hepatocytes in hemophilia A mice. *Mol. Ther.* **30**, 3552–3569 (2022).
240. Brooks, E. D. *et al.* Long-term complications of glycogen storage disease type Ia in the canine model treated with gene replacement therapy. *J. Inherit. Metab. Dis.* **41**, 965–976 (2018).

# SCIENTIFIC COMMUNICATIONS

## **SCIENTIFIC COMMUNICATIONS**

### **1. PUBLICATIONS (RELATED TO THE PHD):**

- **Gardin, A.**, Rouillon, J., Montalvo-Romeral, V., Rossiaud, L., Vidal, P., Launay, R., Vie, M., Krimi-Benchekroun, Y., Bertin, B., Dubreuil, G., Nozi, J., Jauze, L., Fragnoud, R., Daniele, N., Van Wittenberghe, L., Esque, J., André, I., Nissan, X., Hoch, L. & Ronzitti, G. Single rAAV vector gene therapy with a mini-GDE transgene corrects heart and muscle impairment in mouse, rat, and human models of glycogen storage disease type III. **Submitted to *The Journal of Clinical Investigation*, 2023.**
- Rossiaud, L., Fragner, P., Barbon, E., **Gardin, A.**, Benabides, M., Pellier, E., Cosette, J., El-Kassar, L., Giraud-Triboult, K., Nissan, X., Ronzitti, G. & Hoch, L. Pathological modeling of Glycogen Storage Disease type III with CRISPR/Cas9 edited human pluripotent stem cells. *Front. Cell Dev. Biol.* 11, 1163427 (2023).
- **Gardin, A.**, Remih, K., Gonzales, E., Andersson, E. R. & Strnad, P. Modern therapeutic approaches to liver-related disorders. *J. Hepatol.* 76, 1392–1409 (2022).

### **2. ORAL PRESENTATIONS :**

- **Journée recherche de la filière FILFOIE:** “Liver-directed gene therapy for inborn errors of metabolism”, 06/06/2023, Paris.
- **American Society of Gene and Cell Therapy (ASGCT):** “Single AAV vector gene therapy with a mini-GDE transgene corrects muscle impairment in mouse, rat and human cellular models of glycogen storage disease type III”, 18/05/2023, Los Angeles.
- **French-American Innovation Days (FAID): clinical trials in gene therapy:** “Liver-directed gene therapy for inherited metabolic disorders”, 17/11/2022, San Diego.

### **3. PATENT:**

Rouillon, J., **Gardin, A.**, Ronzitti, G. N-terminal truncated GDE for the treatment of glycogen storage disease III. Application EP22305841, **under evaluation.**

## RESUME EN FRANÇAIS

## RESUME EN FRANÇAIS

### « OPTIMISATION D'UNE APPROCHE DE THERAPIE GENIQUE POUR LA GLYCOGENOSE DE TYPE III EN UTILISANT DES VECTEURS VIRAUX RECOMBINANTS DE TYPE ADENO-ASSOCIES »

#### 1. INTRODUCTION

L'objectif de cette thèse est l'optimisation d'une approche de thérapie génique utilisant des vecteurs viraux de type adéno-associés recombinants (rAAV) afin de corriger la glycoséose de type III (GSDIII), une maladie héréditaire du métabolisme du glycogène. Dans les sections suivantes, nous décrirons brièvement la maladie, le principe de la thérapie génique utilisant des vecteurs rAAV et les principales difficultés rencontrées lors de sa mise en œuvre, et enfin les avantages et inconvénients de son application à la GSDIII.

##### a. GLYCOGENOSE DE TYPE III

La GSDIII est une maladie génétique rare affectant le métabolisme du glycogène. Elle est liée à des variants dans le gène *AGL* conduisant à l'absence d'enzyme débranchante du glycogène (GDE) fonctionnelle. Il en résulte une accumulation de glycogène dans le foie, le cœur et les muscles squelettiques des patients. La maladie se manifeste dès la première année de vie par une augmentation de la taille du foie (hépatomégalie), des hypoglycémies de jeûne court, une élévation des transaminases plasmatiques (ASAT et ALAT) et un retard de croissance. Une fibrose hépatique est présente tôt dans la vie et peut conduire à une cirrhose à l'âge adulte. Les patients adultes peuvent présenter des tumeurs hépatiques (carcinomes hépatocellulaires) et nécessiter une transplantation hépatique. Dans l'enfance et surtout à partir de l'adolescence, la majorité des patients (85%, forme GSDIIIa) développent une atteinte musculaire progressive pouvant conduire à une perte de la marche à l'âge adulte chez un tiers des patients. Ils présentent aussi une atteinte cardiaque associant une hypertrophie ventriculaire gauche, un risque de mort subite à l'âge adulte et parfois une insuffisance cardiaque pouvant nécessiter une transplantation cardiaque. Il n'existe pas de traitement curatif pour la GSDIII. Les patients doivent prendre des repas fréquents riches en carbohydrates complexes et éventuellement suivre un régime hyperprotidique afin de contrôler les hypoglycémies. Bien que la prise en charge nutritionnelle améliore le phénotype métabolique, elle n'a que peu d'effet sur la progression des atteintes musculaire et cardiaque, qui restent les plus invalidantes à l'âge adulte. Il est donc important de pouvoir développer de nouveaux traitements pour cette pathologie, comme par exemple une approche de thérapie génique.

#### b. THÉRAPIE GÉNIQUE UTILISANT DES VECTEURS ADENO-ASSOCIÉS RECOMBINANTS

La thérapie génique repose sur l'introduction de matériel génétique (*gene addition*), la suppression de l'expression d'un gène (*gene silencing*) ou la modification du génome du patient (édition génique ou *gene editing*) dans le but de corriger une maladie, génétique ou parfois acquise. Les séquences d'ADN ou la machinerie d'édition génique sont apportées aux cellules cibles par des vecteurs, qui peuvent être de nature virale ou non-virale. Parmi les vecteurs viraux, les vecteurs de type rAAV sont les plus utilisés pour réaliser une approche de thérapie génique dite *in vivo*, dans laquelle le vecteur est injecté au patient par voie intraveineuse ou par voie locale (intraoculaire, intracérébrale, intramusculaire). Le vecteur rAAV atteint ses organes cibles, délivre son cargo et permet la modification souhaitée.

Les vecteurs rAAV sont basés sur le virus adéno-associé (AAV), un virus à ADN simple brin non-enveloppé, non pathogène et déficient possédant un génome de 4,7 kb. Ces virus existent à l'état naturel et sont capables d'infecter les cellules humaines. Dans un rAAV, le génome viral a été remplacé par une cassette d'expression comportant le gène d'intérêt, un promoteur permettant son expression dans le tissu souhaité et des séquences répétées terminales (ITRs) permettant l'encapsidation du génome. Les protéines de capsides déterminent le tropisme tissulaire du rAAV et de très nombreux sérotypes ont été découverts ou développés par ingénierie (capsides chimériques). Les rAAV sont des vecteurs viraux en majorité non-intégratifs, qui persistent sous forme d'épisome dans le noyau des cellules infectées. En cas de prolifération cellulaire, l'épisome n'est pas répliqué et sera donc progressivement perdu par « dilution » au cours des mitoses successives. Il ne peut donc être utilisé que pour cibler des cellules quiescentes ou peu proliférantes, comme le foie adulte, les muscles et le système nerveux central. Il s'agit d'une approche prometteuse pour le traitement des maladies monogéniques, qui a montré son efficacité dans des pathologies affectant le foie, les muscles, l'œil ou le système nerveux central. De très nombreux essais cliniques sont actuellement en cours, avec des résultats prometteurs. Plusieurs produits de thérapie génique basés sur des rAAV ont été mis sur le marché ces dernières années, notamment pour traiter l'hémophilie A et B (ciblant le foie), la rétinite pigmentaire (ciblant la rétine) et l'amyotrophie spinale infantile (ciblant le système nerveux central). En dépit de ces bons résultats, plusieurs complications, parfois non prédites par les études précliniques réalisées dans des modèles animaux, ont été mises en évidence lors des essais cliniques chez l'homme. Ces complications comprennent la génération d'anticorps neutralisants anti-rAAV à très fort titres, empêchant toute réadministration, le développement d'une réponse immunitaire dirigée contre le vecteur et la mise en évidence de toxicités notamment hépatiques après injection de fortes doses de vecteurs rAAV.

Tout d'abord, les AAV sont des virus existants à l'état naturel capables d'infecter l'homme, principalement pendant l'enfance. Il en résulte la présence d'anticorps neutralisants « naturels » chez une fraction importante de la population, variable en fonction du sérotype et pouvant dépasser 50% des patients pour certains sérotypes comme l'AAV2. De plus, une réactivité croisée entre les différents sérotypes est relativement fréquente. La présence de ces anticorps neutralisants préexistants empêche l'utilisation des sérotypes en question et restreint le nombre de patients pouvant bénéficier de ce type d'approche de thérapie génique. De plus, l'injection du vecteur rAAV chez des patients séro-négatifs conduit inévitablement au développement de taux très élevés d'anticorps neutralisants, empêchant toute réadministration du sérotype en question. Ainsi, la perte des cellules corrigées (secondaire à une réponse immune par exemple) ou la présence d'une prolifération tissulaire (comme chez l'enfant en croissance par exemple) induit une perte de correction sans possibilité de réinjection du vecteur du fait des anticorps neutralisants. Plusieurs approches sont en cours d'évaluation dans des modèles animaux afin d'échapper aux anticorps neutralisants, de les éliminer transitoirement ou d'éviter leur développement après une première injection de vecteur rAAV. Aucune de ces approches n'a pour l'instant été évaluée chez l'homme.

La seconde difficulté associée à l'utilisation de ce type d'approche est le développement d'une réponse immunitaire chez les patients recevant le vecteur rAAV. Les premiers essais de thérapie génique réalisés chez des patients atteints d'hémophilie B ont révélé le développement d'une réponse immunitaire cytotoxique médiée par les lymphocytes T et dirigée contre les antigènes de capsid du virus. Cette réponse a conduit à la destruction des hépatocytes transduits et à une perte de correction. Il s'agit d'une réponse immune dose-dépendante partiellement contrôlable par l'utilisation transitoire d'immunosuppresseurs et notamment de corticostéroïdes. A l'inverse, peu de réponses immunes ont été démontrées contre le transgène, possiblement en lien avec une tolérance liée à l'expression du transgène dans le foie, un organe connu pour ces propriétés tolérogènes. Les essais suivants ont aussi révélé la présence d'une toxicité plus précoce, liée à l'activation de la réponse immunitaire innée et notamment celle du complément, parfois sévère et pouvant conduire au décès.

Enfin, le ciblage du muscle et du système nerveux central avec des vecteurs rAAV pose des questions spécifiques. En effet, les vecteurs rAAV sont très bien adaptés au ciblage du foie, qui est probablement une cible naturelle des AAV, alors que le ciblage des autres organes nécessite de fortes doses de rAAV. Les essais cliniques réalisés pour traiter des maladies neuromusculaires comme l'amyotrophie spinale infantile ou la myopathie à myotubularine (MTM) ont révélé la présence d'une toxicité hépatique, liée à une surcharge des hépatocytes aux fortes doses utilisées pour cibler le muscle et la moelle épinière. Plusieurs patients traités atteints de MTM sont décédés d'une insuffisance

hépatique secondaire à cette toxicité, probablement aggravée par une atteinte hépatique sous-jacente spécifique de cette pathologie. Ainsi, il est impératif de mettre au point des vecteurs rAAV capables de cibler ces deux organes sans cibler le foie, ce qui est d'autant plus important pour des patients ayant aussi une atteinte hépatique sous-jacente, comme les patients atteints de MTM mais aussi les patients atteints de GSDIII. Plusieurs approches ont été développées, incluant la génération de vecteurs rAAV possédant une insertion peptidique au niveau de la capsid, permettant un ciblage plus spécifique du muscle tout en réduisant le ciblage du foie.

### C. THÉRAPIE GÉNÉRIQUE BASÉE SUR LES RAAV POUR LA GSDIII

Plusieurs éléments font de la GSDIII une cible idéale pour une approche de thérapie génique basée sur les rAAV : la présence de variants dans un gène unique (*AGL*), la physiopathologie bien caractérisée de la maladie et la possibilité d'améliorer le phénotype en restaurant uniquement une fraction même faible de l'activité enzymatique normale de GDE. En effet, un grand nombre de maladies métaboliques voient leurs symptômes et leur pronostic très largement améliorés en présence d'une activité enzymatique résiduelle, même très faible (5% par exemple), ce qui suggère qu'il n'est pas forcément nécessaire d'obtenir une correction complète pour observer un bénéfice significatif.

A l'inverse, plusieurs difficultés plus spécifiques à la GSDIII existent. Premièrement, la taille de la séquence codante de GDE est de 4,6 kb ce qui dépasse les capacités d'encapsulation du vecteur rAAV, après ajout des séquences régulatrices et ITRs indispensables. Deuxièmement, il est nécessaire de corriger deux organes très différents, le foie et les muscles (incluant le cœur), nécessitant un développement technologique afin d'obtenir une capsid et des séquences régulatrices capables de cibler et d'exprimer le transgène dans ces deux tissus. Troisièmement, la GSDIII s'accompagne d'une atteinte hépatique chronique avec prolifération et fibrose, ce qui peut impacter à la fois la transduction du foie par les vecteurs rAAV mais aussi la persistance du vecteur à long terme avec un risque de perte de correction secondaire à la dilution du vecteur du fait de la prolifération. De plus, un risque théorique de décompensation de l'atteinte hépatique existe en cas d'utilisation de vecteurs rAAV à fortes doses, surtout du fait de l'atteinte hépatique inhérente à la GSDIII, à l'image de celle observée dans les essais cliniques ciblant la MTM.

Plusieurs approches ont été développées pour surmonter ces difficultés, incluant l'utilisation d'une stratégie à deux vecteurs et l'utilisation d'une enzyme bactérienne orthologue de GDE. Dans la première stratégie, développée dans notre équipe, deux vecteurs sont utilisés, chacun comportant la moitié de la cassette d'expression. Les deux vecteurs sont ensuite injectés par voie intraveineuse dans

un modèle de souris de la GSDIII (souris *Agf<sup>-/-</sup>*). Après avoir atteints leurs cellules cibles, les deux vecteurs recombinent et produisent un transgène complet, permettant ainsi l'expression de GDE. Cette approche a démontré son efficacité pour la traitement de l'atteinte musculaire des souris malades, mais n'a pas permis de corriger la totalité du phénotype hépatique. Les inconvénients de cette stratégie sont la nécessité d'utiliser deux vecteurs (et donc une dose double), le caractère peu efficace de la recombinaison entre les deux vecteurs et les risques de toxicité, principalement liés à l'expression potentielle d'une protéine tronquée non fonctionnelle issue du premier vecteur. La seconde stratégie utilise la pullulanase, une enzyme bactérienne homologue de GDE, dont la séquence codante fait 2,2 kb et peut donc être facilement encapsidée dans un vecteur rAAV. Un vecteur rAAV exprimant la pullulanase permet la correction du phénotype musculaire et hépatique 12 semaines après injection intraveineuse du vecteur, mais on observe une perte de correction complète dans les semaines qui suivent du fait d'une réponse immunitaire cellulaire intense liée au caractère immunogène de la pullulanase, d'origine bactérienne. Ainsi, les deux approches publiées présentent chacune des limitations spécifiques et une optimisation apparaît indispensable avant de pouvoir appliquer ce type d'approche aux patients atteints de GSDIII.

Dans la première partie de ce travail, nous avons développé une approche de thérapie génique utilisant une forme tronquée fonctionnelle de l'enzyme GDE, capable d'être encapsidée dans un unique vecteur rAAV et permettant la correction de l'atteinte musculaire et cardiaque de trois modèles de GSDIII : le modèle de souris *Agf<sup>-/-</sup>*, un modèle de rat récemment généré par notre équipe et un modèle cellulaire humain issu de cellules pluripotentes induites (hiPSC) invalidées pour le gène *AGL*, développé par l'équipe de Xavier Nissan au sein de l'institut I-STEM. Cette partie a été soumise pour publication. Dans une seconde partie, nous avons évalué les capacités d'un vecteur rAAV unique à corriger l'atteinte hépatique des souris GSDIII.

## **2. CORRECTION DE L'ATTEINTE MUSCULAIRE ET CARDIAQUE DE LA GSDIII**

Dans cette première partie, nous avons évalué la possibilité de corriger l'atteinte musculaire et cardiaque de plusieurs modèles de GSDIII en utilisant une stratégie avec un vecteur rAAV unique, car ces deux atteintes sont les plus invalidantes à l'âge adulte chez les patients atteints de GSDIII.

### **a. DEVELOPPEMENT D'UN VECTEUR UNIQUE EXPRIMANT UNE FORME TRONQUEE DE GDE**

Tout d'abord, nous avons optimisé les différents éléments de la cassette d'expression en dehors du transgène (promoteur, intron, signal de poly-adénylation/poly-A), afin de sélectionner les éléments les plus petits conservant une efficacité satisfaisante. En utilisant une forme tronquée du promoteur CMV (miniCMV), une poly-A de petite taille (pA58) et en retirant l'intron, nous avons pu

généraliser un vecteur unique exprimant la séquence codante codon-optimisée de la GDE complète d'origine humaine. Ce vecteur rAAV possède une cassette d'expression autour de 5,3 kb, dépassant les capacités d'encapsidation d'un vecteur rAAV (vecteur dit *oversized*). Cependant, il peut néanmoins être produit au prix d'une productivité médiocre et de l'encapsidation de génomes tronqués. Nous avons choisi une capsid virale modifiée, développée dans notre équipe, et possédant une insertion peptidique permettant un ciblage efficace du muscle et du cœur tout en réduisant très fortement le ciblage du foie. Après injection de ce vecteur par voie intraveineuse dans des souris *Agf<sup>-/-</sup>*, nous avons observé une expression de GDE dans les muscles et le cœur et une réduction partielle du glycogène mesuré par mesure enzymatique dans les extraits de tissus et par coloration par acide périodique – Schiff (PAS) sur des coupes de tissus. Du fait de l'utilisation d'un promoteur et d'une capsid ciblant spécifiquement le muscle, aucune correction du foie n'était attendue. Ce vecteur démontre la possibilité de traiter le phénotype musculaire avec un vecteur unique, mais la correction reste partielle et un tel vecteur ne sera dans tous les cas pas utilisable en clinique humaine du fait de sa mauvaise productivité et de l'encapsidation de nombreux génomes tronqués.

La génération d'un vecteur rAAV unique non-*oversized* nécessite donc de réduire la taille du transgène, dernier élément de la cassette n'ayant pas été modifié. La troncature d'une protéine, *a fortiori* d'une enzyme, expose au risque d'obtenir une protéine non fonctionnelle ou instable. En utilisant la prédiction de structure tridimensionnelle de la GDE humaine obtenue par le logiciel Alpha-Fold2, la position des sites catalytiques et en tenant compte des positions des différents variants faux-sens pathogènes décrits chez les patients GSDIII, nous avons identifié les deux régions les plus à même d'être retirées sans affecter l'activité de la protéine : le domaine N-terminal (domaine M1) et le domaine central (domaine C). Nous avons, dans un premier temps, généré plusieurs protéines tronquées dans le domaine C (protéines  $\Delta C$ ). L'injection d'un vecteur rAAV exprimant ces mutants par voie intraveineuse a permis une réduction modérée de l'accumulation de glycogène dans les muscles squelettiques (triceps, quadriceps), mais sans aucun effet observable au niveau du cœur. De plus, ces vecteurs n'ont pas été à même de restaurer la force musculaire des souris *Agf<sup>-/-</sup>*. Ainsi, les mutants  $\Delta C$ , bien que capable de s'exprimer correctement, ne conservent qu'une activité enzymatique modeste qui, de plus, semble tissu spécifique.

Nous avons ensuite généré plusieurs protéines tronquées dans le domaine N-terminal (protéines  $\Delta N$ ter), en utilisant la prédiction de structure obtenue par le logiciel Alpha-Fold2, afin de réaliser les troncatures entre les différentes structures secondaires prédites du domaine N-terminal. L'injection d'un vecteur rAAV exprimant ces protéines tronquées par voie intramusculaire dans le muscle tibial antérieur des souris *Agf<sup>-/-</sup>* a démontré une réduction significative du glycogène dans ce

muscle un mois après injection, et ce dans tous les groupes de souris traitées. Les souris injectées avec le vecteur rAAV exprimant la protéine tronquée  $\Delta Nter2$  présentent la réduction la plus importante du contenu en glycogène, équivalente à celle obtenue avec le vecteur *oversized* exprimant la GDE de taille complète. Ainsi, nous avons sélectionné le vecteur exprimant la  $\Delta Nter2$  pour les expériences suivantes.

b. CORRECTION COMPLETE DU PHENOTYPE MUSCULAIRE DES SOURIS ET DES RATS GSDIII AVEC UN VECTEUR RAAV EXPRIMANT  $\Delta NTER2$

Compte tenu des résultats obtenus après injection intramusculaire, nous avons injecté le vecteur rAAV exprimant la  $\Delta Nter2$  par voie intraveineuse dans les souris *Agf<sup>-/-</sup>* âgées de 4 mois, qui présentent déjà une atteinte musculaire installée (accumulation du glycogène et perte de force musculaire). Nous avons observé une correction complète de l'accumulation de glycogène dans le cœur et les muscles squelettiques, évaluée à la fois par méthode enzymatique dans des extraits de tissus et par coloration au PAS sur coupes de tissus, 3 mois après injection. De plus, une restauration quasi-complète de la force musculaire des animaux était observée, suggérant la possibilité de traiter des animaux ayant une atteinte musculaire déjà installée. Enfin, nous avons mis en évidence une élévation de la myoméline 3 (*Myom3*), un biomarqueur récemment décrit dans les dystrophies musculaires, dans le sérum des souris *Agf<sup>-/-</sup>* avant injection. Après injection du vecteur rAAV exprimant  $\Delta Nter2$ , nous avons observé une normalisation de ce biomarqueur chez les souris *Agf<sup>-/-</sup>* traitées, témoignant là encore de la capacité du vecteur exprimant  $\Delta Nter2$  à corriger un muscle déjà atteint.

Ensuite, nous avons évalué ce vecteur rAAV exprimant la  $\Delta Nter2$  dans un modèle de rat de la GSDIII (rat *Agf<sup>-/-</sup>*), récemment développé par notre équipe, et qui récapitule les principales atteintes observées chez les patients. L'injection du vecteur par voie intraveineuse dans des rats âgés de 6 semaines a permis une correction quasi-complète de l'accumulation de glycogène dans les muscles squelettiques et dans le cœur des rats *Agf<sup>-/-</sup>* 3 mois après injection. Nous n'avons pas pu analyser la perte de force musculaire, du fait de l'absence de test validé pour la mettre en évidence dans ce modèle. Comme chez les souris *Agf<sup>-/-</sup>*, nous avons pu mettre en évidence une normalisation du biomarqueur *Myom3* dans le sérum des rats *Agf<sup>-/-</sup>* traités. Ces données confirment l'efficacité du vecteur observée chez les souris *Agf<sup>-/-</sup>*.

c. CORRECTION DE L'ACCUMULATION DE GLYCOGENE DANS UN MODELE CELLULAIRE HUMAIN DE LA MALADIE

Pour valider l'efficacité d'un vecteur rAAV exprimant la  $\Delta Nter2$ , nous avons évalué sa capacité à corriger l'accumulation de glycogène dans un modèle cellulaire humain récapitulant le phénotype de la GSDIII, récemment mis au point par nos collaborateurs de l'institut I-STEM. Ces cellules sont issues de cellules souches pluripotentes induites humaines (hiPSC) invalidées pour le gène *AGL* via la

technologie CRISPR-Cas9 (GSDIII<sup>CRISPR</sup>). Une fois dérivées en myoblastes puis différenciées en myotubes, ces cellules accumulent du glycogène par rapport à un contrôle isogénique après culture dans un milieu sans glucose. Après transduction des cellules GSDIII<sup>CRISPR</sup> et des contrôles isogéniques par un vecteur rAAV exprimant la GFP ou la GDE ΔNter2, nous avons montré la réduction du niveau de glycogène accumulé dans les cellules GSDIII<sup>CRISPR</sup>, par rapport aux cellules transduites avec le vecteur contrôle. L'analyse d'un marquage des cellules par le PAS confirme la réduction du glycogène dans les cellules traitées par le vecteur rAAV exprimant la GDE-ΔNter2. De plus, les cellules GSDIII<sup>CRISPR</sup> ne présentent aucune toxicité spécifique liée à l'expression de la protéine tronquée GDE-ΔNter2, avec une viabilité cellulaire et une expression des marqueurs de différenciation musculaire identiques aux cellules contrôles.

Toutes ces données démontrent la capacité de la protéine tronquée GDE-ΔNter2 à améliorer le phénotype musculaire de plusieurs modèles de GSDIII et suggèrent un intérêt pour le traitement des patients atteints de GSDIII.

### 3. CORRECTION DE L'ATTEINTE HEPATIQUE DE LA GSDIII

Dans cette seconde partie, nous avons évalué la capacité d'un vecteur rAAV unique à corriger l'atteinte hépatique des souris *Agf*<sup>-/-</sup>. Tout d'abord, nous avons caractérisé le phénotype hépatique de ces animaux à différents âges. Ces animaux récapitulent le phénotype observé chez les patients, avec apparition progressive d'une hépatomégalie et accumulation de glycogène dans le foie dès un jeune âge. De plus, les souris *Agf*<sup>-/-</sup> présentent une glycémie basse à tous les âges analysés et une élévation des taux de transaminases plasmatiques. L'analyse histologique révèle la présence d'une surcharge de glycogène dès un jeune âge, et d'une fibrose hépatique progressive, surtout après l'âge de 6 mois. Ainsi, le phénotype est présent tôt dans la vie des animaux, dès l'âge de 6 semaines, en dehors de la fibrose hépatique, présente à partir de 6 mois de vie.

#### a. EVALUATION DU VECTEUR CHEZ DES SOURIS GSDIII PRESENTANT UNE MALADIE HEPATIQUE AVANCEE

L'utilisation du vecteur rAAV utilisé dans la première partie n'est pas possible pour corriger le phénotype hépatique, car la capsidie utilisée ne cible pas le foie et le promoteur utilisé est connu pour présenter une extinction progressive dans le foie des souris. Nous avons donc généré un vecteur rAAV exprimant la séquence codante de la GDE humaine de taille complète sous le contrôle d'un promoteur hépatique et avec une capsidie ciblant le foie. Ce vecteur rAAV a été injecté par voie intraveineuse dans des souris *Agf*<sup>-/-</sup> âgées de 4-5 mois, c'est-à-dire présentant l'ensemble du phénotype hépatique, ce afin de mimer le traitement de patients GSDIII adultes, déjà malades. Après injection du vecteur rAAV, la glycémie et les transaminases sériques sont suivies de façon prospective tous les mois. Les souris *Agf*<sup>-/-</sup> ayant reçu le vecteur rAAV présentent une correction de la glycémie 2 mois après injection, mais

cette correction est progressivement perdue lors du suivi entre 3 et 6 mois après injection. L'expression de GDE dans le foie des animaux 9 mois après injection est uniquement détectable avec le vecteur ciblant le foie et est très faible, insuffisante pour permettre une correction significative du phénotype. En effet, il n'y avait pas de baisse significative du taux de glycogène dans le foie des souris traitées par rapport aux souris injectées avec du PBS, bien que les résultats soient hétérogènes avec des valeurs de glycogène dans les normes pour certaines souris traitées par le vecteur hépatotrope. L'analyse histologique révèle une fibrose hépatique et une vacuolisation des hépatocytes similaires à celles des souris non traitées, bien que des plages d'hépatocytes corrigés (non vacuolisés) soient observables dans le foie des souris traitées. Le niveau de correction apparaît néanmoins insuffisant pour permettre une correction de la glycémie. Les mécanismes qui sous-tendent cette perte d'expression sont en cours de caractérisation. Parmi les hypothèses, on retrouve une correction insuffisante avec une prolifération résiduelle, possiblement du fait d'un traitement trop tardif (fibrose trop avancée), ayant conduit à une dilution progressive du vecteur et/ou une réponse immunitaire dirigée contre le vecteur, ayant conduit à la perte des cellules transduites.

#### b. CORRECTION DE L'ATTEINTE HEPATIQUE DE JEUNES SOURIS ATTEINTES DE GSDIII

Nous avons ensuite utilisé des souris plus jeunes, âgées de 6 semaines lors de l'injection et ne présentant pas de fibrose hépatique significative. Ces animaux ont été injectés par voie intraveineuse avec un vecteur rAAV exprimant la GDE humaine de taille complète sous le contrôle d'un promoteur hépatique plus puissant. Toutes les souris *Agf<sup>-/-</sup>* injectées avec un vecteur rAAV ont présenté une correction rapide de la glycémie dès un mois après l'injection. La correction de la glycémie a été maintenue 6 mois, mais a ensuite été perdue progressivement chez la moitié des animaux. L'analyse du foie 9 mois après injection montre une correction de l'hépatomégalie et de l'accumulation de glycogène chez les souris traitées avec le vecteur rAAV, avec là encore une hétérogénéité, comme pour la mesure de la glycémie, reflétant la co-existence de souris corrigées et de souris ayant perdues progressivement la correction. L'analyse histologique confirme cette variabilité, avec une histologie corrigée (hépatocytes non vacuolés, absence de fibrose) dans le foie des souris ayant une glycémie normale au dernier suivi, alors que l'histologie des souris ayant perdu la correction de la glycémie est similaire à celles des souris *Agf<sup>-/-</sup>* non traitées.

Pour conclure, la correction du phénotype hépatique de la GSDIII en utilisant un vecteur rAAV est possible en traitant les souris à un jeune âge, mais une partie des animaux présentent néanmoins une perte de correction progressive dont le mécanisme reste peu clair. Elle pourrait être liée à une correction insuffisante des hépatocytes, avec persistance d'une prolifération conduisant à une dilution du vecteur, et/ou à une réponse immune dirigée contre le vecteur rAAV. Ce phénomène est

actuellement en cours de caractérisation. Des optimisations supplémentaires sont nécessaires, notamment l'utilisation de la forme tronquée de GDE optimisée dans la première partie de ce travail.

#### **4. CONCLUSION**

Le traitement de la GSDIII en utilisant des vecteurs de type rAAV présente plusieurs défis, à savoir la grande taille de la séquence codante de GDE, la nécessité de corriger plusieurs organes très différents (cœur, muscles, foie) et la présence d'une atteinte hépatique chronique associant fibrose et prolifération des hépatocytes. L'application d'une telle approche à la GSDIII nécessite donc plusieurs optimisations. Dans la première partie de ce travail, nous avons démontré la possibilité de corriger l'atteinte cardiaque et musculaire de plusieurs modèles de GSDIII en utilisant une forme tronquée fonctionnelle de la GDE, capable d'être encapsidée de façon efficace dans un vecteur rAAV unique. Dans la seconde partie de ce travail, nous avons montré la possibilité de corriger l'atteinte hépatique du modèle murin de GSDIII en utilisant un vecteur unique, mais aussi les limites de cette approche, avec une perte de correction progressive chez une partie des animaux traités. Il reste nécessaire d'améliorer le transfert du gène codant GDE dans le foie des souris GSDIII, notamment en utilisant la forme tronquée de GDE générée dans la première partie de ce travail, ainsi que le design de séquences régulatrices permettant l'expression du transgène dans le foie et les muscles, afin de générer un vecteur unique capable de corriger la totalité du phénotype de la maladie.