

A novel synthetic biology approach for MPRA-optimized protein cloning to stop the progression of Tay-Sachs Disease*

Abhinav Dubba, Aditya Mukker, Miya Nithin, Nila Prabhu, and Abhinav Yamujala ▫ Lambert High School, Suwanee, GA

Reviewed on 3 May 2025; Accepted on 9 June 2025; Published on 27 October 2025

*Tay-Sachs disease is a rare, fatal, neurodegenerative disorder caused by mutations in the HEXA gene, which results in the absence or dysfunction of the alpha subunit of the lysosomal enzyme β -hexosaminidase A (β -HexA). It affects 1 in 320,000 babies in the United States and those with it have an average lifespan of two to four years. The enzyme is vital to an individual's health as it prevents GM2 ganglioside buildup in the brain and spinal cord. When the enzyme is not produced, neural activity rapidly deteriorates, leading to worsening effects that result in early childhood mortality. To combat this problem, we propose a synthetic biology approach that involves engineering a plasmid to overproduce the alpha subunit of β -HexA and optimizing it using Massive Parallel Reporter Assays (MPRA) arrays to increase the expression specificity of the plasmid and reduce off-target effects. The plasmid will be cloned in *E. Coli* bacteria for protein production. SDS PAGE and western blotting will be used to qualitatively evaluate protein production. The MPRA arrays will be used for transcriptional efficiency of the plasmid as well as validation of plasmid efficiency. This will be done by transfecting the cells in a petri dish and measuring the amount of RNA produced. The MPRA arrays will allow for minimized off-target toxicities and optimization of the plasmid's effectiveness and accuracy.*

Keywords: *Tay-Sachs Disease, Synthetic Biology, MPRA Arrays, Protein Cloning, β -hexosaminidase A*



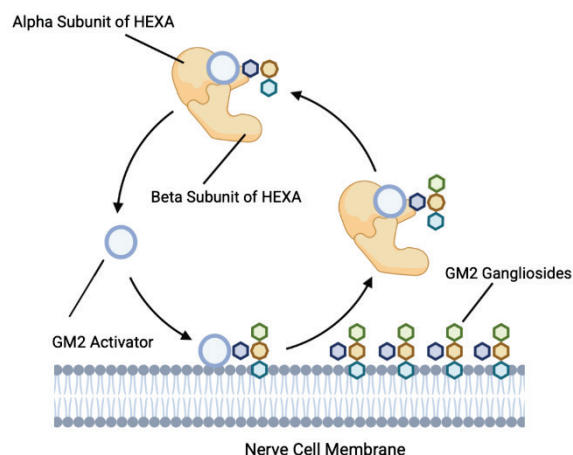
Tay-Sachs disease is a genetic disorder that typically manifests in infancy and early childhood development. It is characterized by vision and hearing loss, paralysis, involuntary muscle twitches, dysphagia, increasingly severe seizures, and other major symptoms. Its cause is the buildup of GM2 gangliosides, which are fatty acids that indirectly impair the central nervous system.

Specifically, various mutations in the

HEXA gene lead to underproduction of the alpha subunit of the β -hexosaminidase-A protein (see Figures 1 and 2).

Over 210 different mutations can cause this disease, so it would be difficult to target each mutation (Leal, 2020). Current treatments include various forms of therapy and medications. However, there is no clear cure and these treatments are simply mitigative remedies. Our solution targets the outcome of the issue by overproducing the

* The authors were mentored by Catherine Sharer from Lambert High School and Alyssa Morrow from Ginkgo Bioworks. Please direct correspondence to: morrowalyssa@gmail.com. This is an Open Access article, which was copyrighted by the authors and published by BioTreks in 2025. It is distributed under the terms of the Creative Commons Attribution License, which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



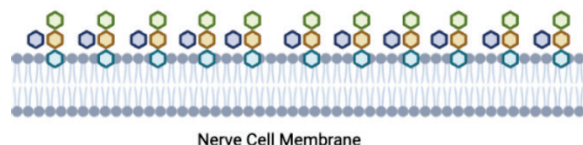
Individual without Tay-Sachs Disease

Figure 1. Production of GM2 Gangliosides via the Alpha Subunit of the HEXA Gene for someone without Tay-Sachs.

alpha subunit of the lysosome enzyme β -hexosaminidase-A.

Systems level

Our solution is a biological system to externally overproduce functional β -hexosaminidase A (HexA) protein, which is then given to Tay-Sachs patients. The system begins with engineered cells transformed with a novel plasmid designed to overproduce the HEXA α -subunit, which is then purified to form the active HexA enzyme. This recombinant HexA is then transported to neuronal lysosomes, where it breaks down accumulated GM2 gangliosides. This treatment uses the body's natural lysosomal uptake mechanisms, bypassing the genetic defect. By maintaining HexA levels by using periodic doses of the protein, we create a sustained corrective effect that halts neurodegeneration.



Individual with Tay-Sachs Disease

Figure 2. Production of GM2 Gangliosides via the Alpha Subunit of the HEXA Gene for someone with Tay-Sachs.

This protein replenishment approach interacts with human physiology at three key levels:

(1) Cellular uptake into the central nervous system via receptor-mediated endocytosis (Jeong, 2022)

(2) Lysosomal uptake via the Mannose-6-Phosphate (M6P) pathway (Coutinho, 2012)

(3) Enzymatic breakdown of the GM2 ganglioside inside of the central nervous system (Huff and Daly, 2019)

This protein-replacement strategy guarantees immediate, controlled, and scalable intervention without permanently altering or harming the genes of the patient. By optimizing the interactions between the proteins and their biological environments, this treatment has the potential to play a key role in extending and improving the lives of Tay-Sachs patients, representing an important step to a definitive cure for Tay-Sachs.

Device level

To insert the plasmid into neuronal cells and encourage lysosomal uptake, we plan to utilize receptor-mediated endocytosis and employ the Mannose-6-Phosphate (M6P) pathway. Endocytosis is the quickest and simplest method to ensure successful delivery of the treatment into the cell. Receptor-mediated phagocytosis – a form of endocytosis used for large molecules – would allow the cellular membrane to fold inward and form a vesicle around the plasmid; however, this method does not guarantee lysosomal uptake. Therefore, we plan on coupling this technique with the M6P pathway, which involves a specific residue recognized by M6P – allowing entry into the lysosomal system (Coutinho 2021). By guaranteeing that this residue is present in the plasmid, we can ensure lysosomal uptake and the eventual breakdown of GM2 gangliosides.

Parts level

For experimentation, we plan on inserting the HEXA gene into the pUC19 plasmid in place of the LacZ gene. pUC19 was chosen due to

its high copy number and its cloning ease. To ensure compatibility with the chosen *E. coli* plasmid, we will remove the introns to bacterially translate the human HEXA gene. Although *E. coli* cannot perform post-translational modification, this limitation is acceptable at this stage of the project as the current focus is on the transcriptional activity rather than the protein activity. Once the introns are spliced out, the new exon chain can be inserted into pUC19, where it will then rapidly replicate under optimal conditions (40°C in LB medium with antibiotics) (J.F. & Russell, 2001). We will then test the plasmid's effectiveness using an SDS-PAGE gel, the most optimal form of gel electrophoresis for protein separation and identification, as well as a western blot to ensure the presence of HEXA. Once the presence of HEXA is confirmed, the protein will be extracted and placed in a glycerol-phenol solution to preserve it and to stabilize it against any temperature strain.

Additionally, during experimentation, MPRA arrays, a machine learning method, will be used to validate the effectiveness of the plasmid. The amount of RNA that is produced in the cell after transfection will be compared to the amount of DNA produced. This data will then be inputted into the model to approximate the ratio between RNA and DNA. We also plan to compare a neuron-specific cell line (high HEXA RNA expression) to Tay-Sachs disease and compare it to a liver cell line (low HEXA RNA expression) to measure the RNA levels of the plasmids, allowing us to test for any off-target toxicity and/or negative effects and make our plasmid more specific and efficient.

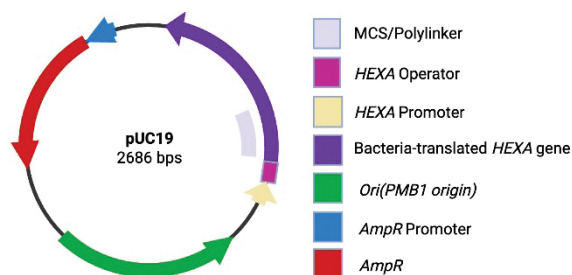


Figure 3. Designed plasmid to overproduce the alpha subunit of beta hexosaminidase-A.

Next steps

After initial testing with the mentioned plasmid, we plan to also test our idea using mammalian cells, as well as testing for successful protein extraction after the cell is cloned. Once the protein has been properly extracted, we plan on adding it to a glycerol solution to store it and ensure shelf life, as well as testing two different treatment administration methods: a nasal spray or an epidural injection. The first method, a nasal spray, would utilize the olfactory and trigeminal nerves found in the nasal cavity to get the protein to the brain through a spray solution containing the beta hexosaminidase-A. The second method, an epidural injection, would utilize the same solution containing the protein but would instead be injected directly into the spinal cord to be spread throughout the central nervous system.

Safety

One possible source of danger for the patients of this treatment is that epidural injections to the spinal cord can cause paralysis (Xia and Gan 2013). An experienced anesthesiologist must administer these injections on a fixed schedule to ensure the patient's safety. This approach can be expensive and risky if the licensed professional makes a mistake. One approach to ensure the safety of the patient throughout the periodic protein administration is a nasal spray. By using a nasal spray, the process of transporting the extracted proteins to the central nervous system (CNS) will be more efficient as it can bypass the blood-brain barrier through the trigeminal and olfactory nerves (Jeong 2022). This would minimize any off-target effects and reduce the risk of potentially harming the patient. Another potential risk for this project is the use of mammalian cells later on in testing, as it may cause increased variability in protein expression. This risk can be mitigated by first testing with *E. coli* cells and checking protein expression and other metrics to ensure safe production.

Discussions

Our system minimizes the biological impacts of Tay-Sachs disease— regardless of the mutations that cause it (over 210 mutations in total). Furthermore, it has the potential to significantly extend an individual's lifespan; however, delivery methods require further research and clinical trials. Due to the blood-brain barrier, it is difficult to transport the treatment into neuronal cells. To combat this, we proposed two solutions: a nasal spray and periodic protein administration.

Our future steps include inserting the modified plasmid into *E.coli*, evaluating protein production, testing the plasmid in mammalian cells – this is not something that we can do through our high school; however, it is something we would like to test in the future – testing insertion methods, and creating a solution that can house the β -HexA to reach the CNS safely through the spinal cord. Our idea for this is a constant dosage or injection plan for patients so that we can keep the GM2 gangliosides cleared from their central nervous system and ensure no further damage is done. This would overproduce the enzyme and clear out the GM2 gangliosides from the brain and spinal cord – potentially saving hundreds of children worldwide.

Currently, our two preferred methods are periodic protein doses over a fixed interval and a nasal spray. These administration methods would achieve two different purposes— reaching either the brain or the spinal cord. The injections would be administered to the spinal cord and may be risky due to its fragile nature, as well as potentially causing nerve damage and possible paralysis(paraplegia) (Petrin 2020). The purpose of the nasal spray would be to reach the brain. When the nasal spray is administered to the patient, it will use either the trigeminal or olfactory nerves to reach the brain, and from there, the earlier implementation of receptor-mediated endocytosis will lead the proteins to the lysosomes to clear the gangliosides. The actual dosage of the prescriptions will be decided using a finger prick system to find the amount of ganglioside buildup in each patient (hundreds of different mutations may cause different levels of protein expression).

Altogether, our solution and future plans will work towards a growing need for a cure to this devastating neurodegenerative disease.

Author contributions

A.D, N.P, A.M, M.N, and A.Y contributed to the abstract and background research. A.D, N.P, and M.N created the video. M.N, A.D and A.Y worked on the device level. A.D, A.Y, and A.M contributed to the systems level. A.Y, N.P, and A.M contributed to the parts level. M.N and N.P created the graphics/visuals. M.N, A.D, N.P and A.Y contributed to the safety level. A.D, N.P, M.N, and A.M contributed to the discussions level.

Acknowledgements

One organization we would like to thank is Ginkgo Bioworks, the leading platform for cell programming, for offering our team guidance and mentoring—primarily through the help of numerous specialists in both plasmid creation and artificial intelligence development. The most influential of these mentors is Dr. Alyssa Morrow. Dr. Morrow has been extremely helpful for our research, especially in artificial intelligence and helped us truly cultivate our ideas for the MPRA Array models and purpose of the AI component of our project. She helped us hone in on our purpose and story and answered our questions with insightful research through either her own knowledge or colleagues' help. Our school advisors, Mrs. Sharer and Mr. Schuyler, were also instrumental in the creation of this project. Together, they constantly checked us on progress, feasibility, and clarity throughout the culmination and result of our idea.

Each author served a specific and meaningful purpose in our overall project and contributed to the overall idea. Our team brainstormed ideas through a number of sessions and online meetings, where each individual brought a different knowledge base to the table – ranging from computation to synthetic biology. M. Nithin, A. Dubba and N. Prabhu were essential in creating multiple visual representations of the plasmid

