Modeling the AAV Rep Protein Towards Restricted Replication for Safer Gene Therapy

Leah Whitfield[1,3,4], Roger B. Sutton[2], Matthew L. Hirsch[1,3,4]
University of North Carolina at Chapel Hill[1]
Texas Tech University Health Sciences Center[2], UNC Gene Therapy Center[3], Carolina Eye Research Institute[4]

Introduction: Adeno-Associated Virus (AAV) is a reportedly helper-dependent and non-pathogenic virus. A capsid surrounds the 4.7 kb AAV genome flanked by hairpin-loop inverted terminal repeats (ITRs).[2] Rep proteins bind the Rep-Binding Element (RBE) on the ITRs and nick at the terminal resolution site (trs) through a catalytic tyrosine residue to replicate DNA.[3] Rep proteins have three regions involved in this event, the DNA-Binding Loop (L_{DB}), α-D (αD), and α-E (αE) domains.[4] Promiscuity in Rep binding/nicking of the ITRs among AAV serotypes and recombinant AAV genomes result in non-specific replication and safety concerns for AAV gene therapy.[5]

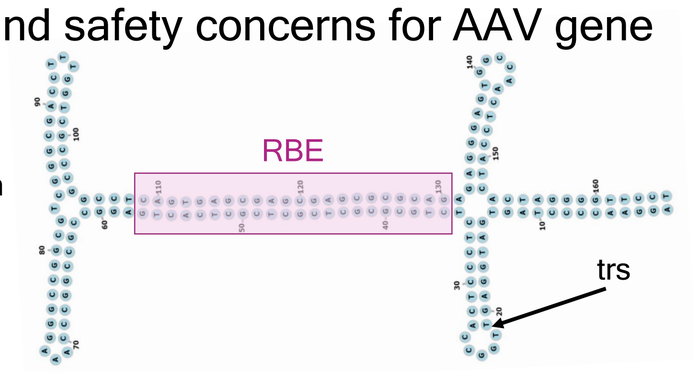

Fig. 1. Secondary Structure of ITR2. ITR2 secondary structure labeled with letter of nucleotide and ITR sequence positions.[1,5] The RBE is boxed in pink, and the nicking site is labeled with an arrow.

Objective: The objective of this work is to rationally design a unique Rep/ITR functional origin of replication that cannot be cross-replicated by AAV Rep proteins found in nature.

Methods: As a first step towards the objective, protein modeling was used to guide rational mutagenesis. The Rep DNA binding/nicking region of AAV serotype 2 was modeled using RoseTTAFold and analyzed in PyMOL.[4,6,7,8] The prediction demonstrates a DNA binding domain like helix-loop-helix conformations frequently observed in binding domains (Fig. 2).

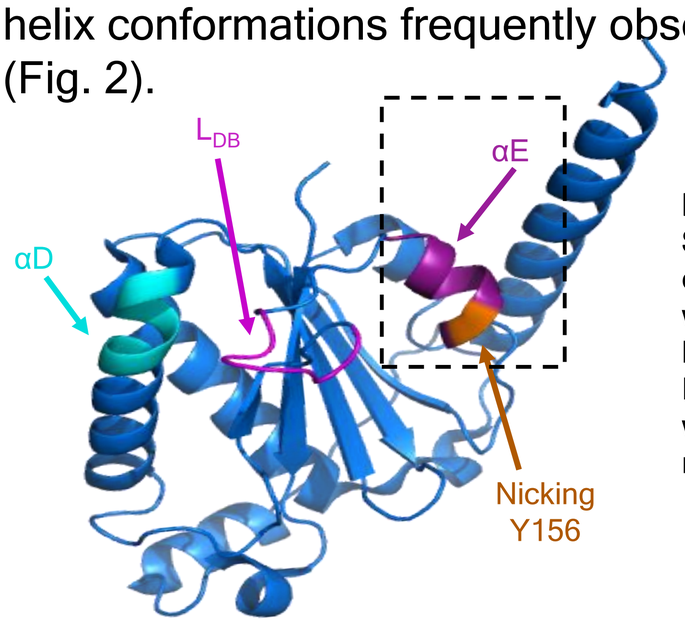

Fig. 2. Labelled Structure of Rep2 Showing Binding Domains. Structure of wild-type Rep2 (wtRep2) in PyMOL with labelled domains.[4,8] The cyan helix is the αD, the pink loop is the L_{DB}, and the purple helix is αE domain which contains the nicking Tyrosine marked in orange.[3]

Since nicking ability necessary for replication is thought to create specificity, altering the structural position of the nicking Y156 while maintaining the helical structure of the αE domain was investigated to restrict cross-replication of natural ITR structures.[5,3] To do this, we first proposed mutation of the catalytic Y156 to alanine to eliminate the ITR nicking activity (Y156A, Fig. 3). Then, an α-helical turn with tyrosine was inserted on the C-terminus of the αE domain to create a triple helix (thA). Finally, helix-breaking proline residues that bracket the turns of the α-helix were mutated to cysteine to maintain the structural position of the nicking tyrosine (thB, Fig. 3).

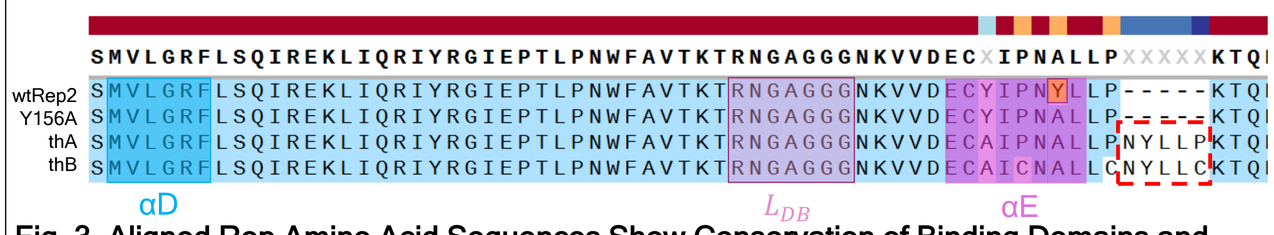

Fig. 3. Aligned Rep Amino Acid Sequences Show Conservation of Binding Domains and Extension of the αE Domain to Alter the Position of the Nicking Tyrosine. Amino acids of wild-type AAV Rep serotype 2 (wtRep2) with labeled αD, L_{DB}, and αE domains.[4] The catalytic Tyrosine is marked in orange.[3] Rep mutants also investigated include Rep Y156A, extension of the αE domain by one helical turn (thA, red box), and thA with three proline residues substituted with cysteines (thB).

Results: Previous reports have demonstrated that mutation of Y156 eliminates Rep-mediated nicking of the ITR.[3] To determine the structural impact of this on the αE domain, modeling was performed using RoseTTAFold by substituting Y156 with A.[6,7,8] The results demonstrate conservation of the αE helix (Fig. 4).

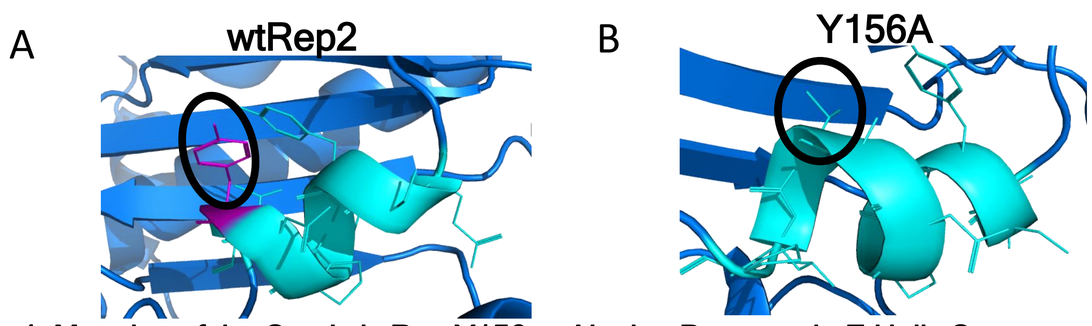

Fig. 4. Mutation of the Catalytic Rep Y156 to Alanine Preserved αE Helix Structure. RoseTTAFold was used for protein modeling.[6,7] (A) The wild-type Rep2 (wtRep2) αE helix is shown in cyan with the catalytic Tyrosine (Y) in magenta.[3] (B) wtRep2 Y156A was made by mutating Y156 to A. A is smaller and neutral nonpolar amino acid making it less reactive and less sterically hindering than Y, a neutral polar molecule.[9]

It was **hypothesized** that duplicating the second turn of the αE catalytic domain (NYLLP) to create a third turn would create a unique nicking interface for later functional selection of a mutant ITR while maintaining the αE helical structure. Modeling of this mutant Rep protein (termed Rep Triple Helix or thA, Fig. 3) revealed that the αE helical structure was not maintained presumably due to proline helix-breakers that create steric hindrance.[9] To test this, the proline residues located on each helical turn were changed to cysteine (termed thB, Fig. 3) and then modeled using RoseTTAFold.[6,7,8,9] The results demonstrate that thB maintains αE helical structure with a single turn extension including a potentially catalytic tyrosine.

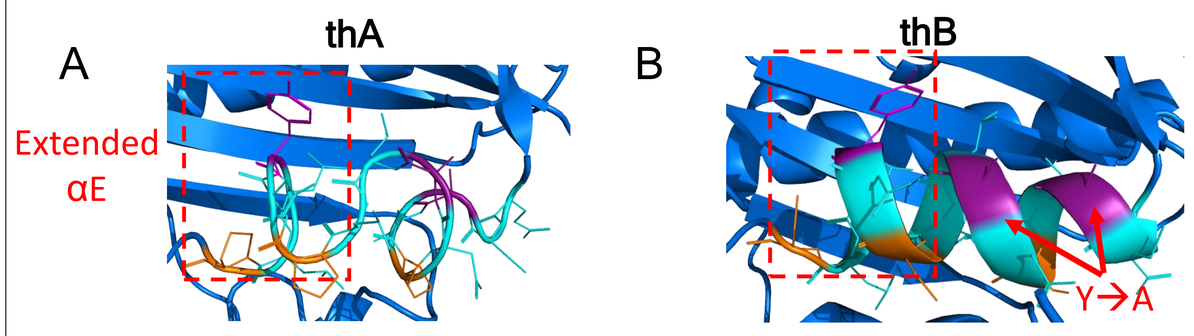

Fig. 5. Extension of the Rep αE Domain by a Single Turn while Preserving the Helix. (A) The extended αE domain (red box) on Y156A (Figs. 3, 4B) reveals helix disruption. (B) The sequence described in (A) with three proline to cysteine substitutions (Fig. 3) demonstrated conservation of αE helix in the presence of the additional inserted turn.

Conclusions:

- Modeling of the Rep DNA binding and ITR nicking domain suggests a binding domain like helix-loop-helix
- Ablating the Rep tyrosine (Y156A) to eliminate ITR nicking did not alter αE helix conformation.
- Extending the αE domain by duplication of the second turn resulted in loss of helical structure
- Proline to cysteine substitutions on the extended turn of the αE maintained helical structure

Future Directions:

- Rep Mutant Production/Characterization. The mutant Rep proteins in Fig. 3 will be generated via site-directed mutagenesis and analyzed by Western analysis and AAV vector production (qPCR, reporter transduction).
- Rep thB Selection of a Mutant ITR. Rep thB will be used replicate and package transgenic genomes flanked by mutant ITRs in a CMV-GFP library. Capsid packaged ITRs will be sequenced and subjected to successive production rounds using Rep thB toward selection of a single ITR sequence.

Acknowledgements: Hirsch and Sutton Labs, Jackie Bower

(1) *Kerpedjiev P, Hammer S, Hofacker IL (2015).* Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. Bioinformatics 31(20):3377-9.
(3) Yoon M, Smith DH, Ward P, Medrano FJ, Aggarwal AK, Linden RM. 2001. Amino-terminal domain exchange redirects origin-specific interactions of adeno-associated virus rep78 in vitro. Journal of Virology. 75(7):3230-3239. doi:10.1128/JVI.75.7.3230-3239.2001. [accessed 2023 Apr 11]. https://pubmed.ncbi.nlm.nih.gov/11238849/.
(4) Santosh V, Musayev FN, Jaiswal R, Zárate-Pérez F, Vandewinkel B, Dierckx C, Endicott M, Sharifi K, Dryden K, Henckaerts E, et al. 2020. The Cryo-EM structure of AAV2 Rep68 in complex with ssDNA reveals a malleable AAA+ machine that can switch between oligomeric states. Nucleic Acids Research. 48(22):12983-12999. doi:10.1093/nar/gkaa1133. https://pubmed.ncbi.nlm.nih.gov/33270897/.
(5) Hewitt FC, Samulski RJ. 2010. Creating a Novel Origin of Replication through Modulating DNA-Protein Interfaces. Lesniak M, editor. PLoS ONE. 5(1):e8850. doi:10.1371/journal.pone.0008850.
(6) M. Baek, et al., Accurate prediction of protein structures and interactions using a three-track neural network, Science (2021). https://www.science.org/doi/10.1126/science.abj8754-
(7) I.R. Humphreys, J. Pei, M. Baek, A. Krishnakumar, et al, Computed structures of core eukaryotic protein complexes, Science (2021). https://www.science.org/doi/10.1126/science.abm4805-
(8) DeLano, W.L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
(9) Why Proline Breaks Alpha Helix and Beta Sheet Structures, J. Kahn UMCP. wwwbiochemumdedu. https://www.biochem.umd.edu/kahn/teach_res/whynopro/.