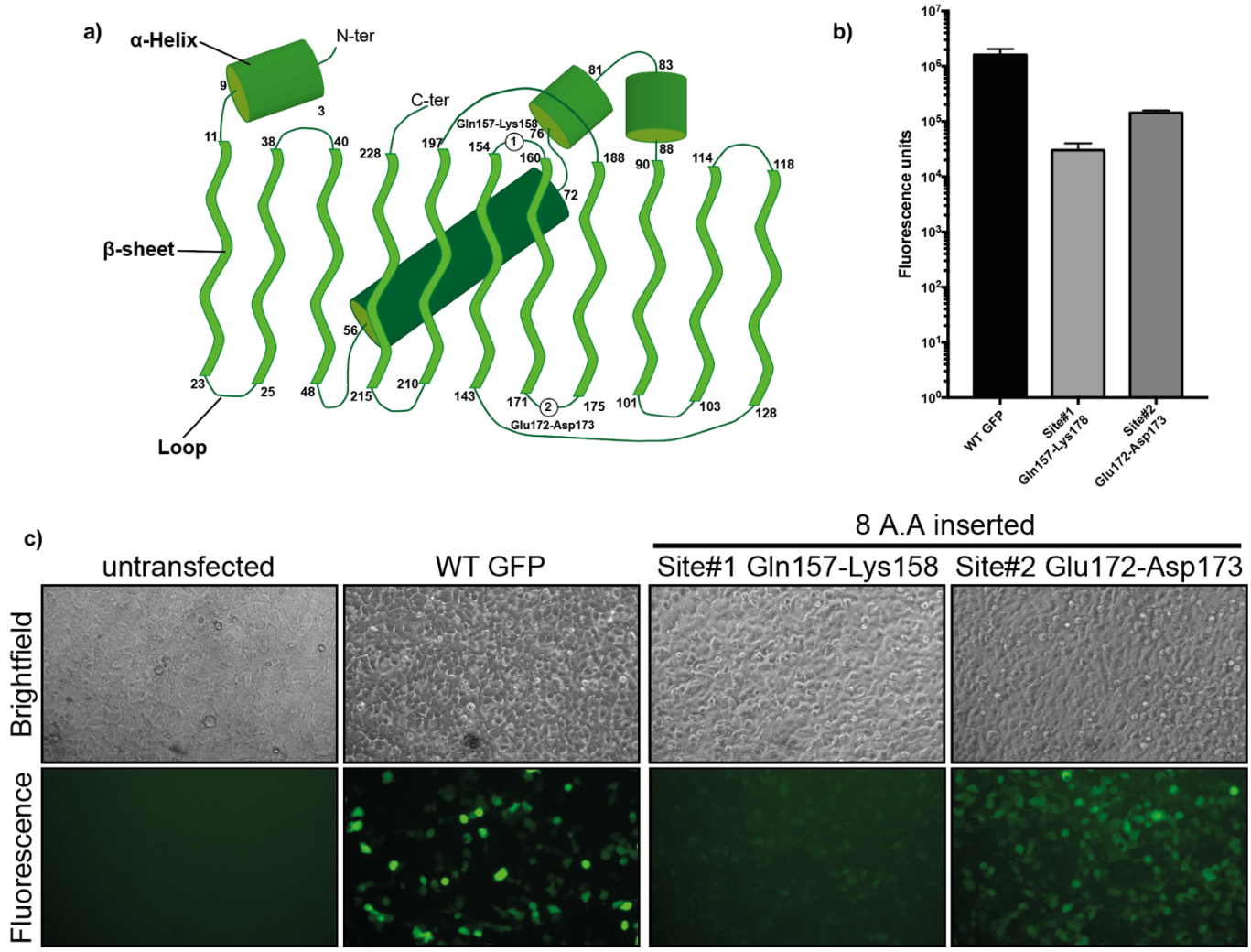
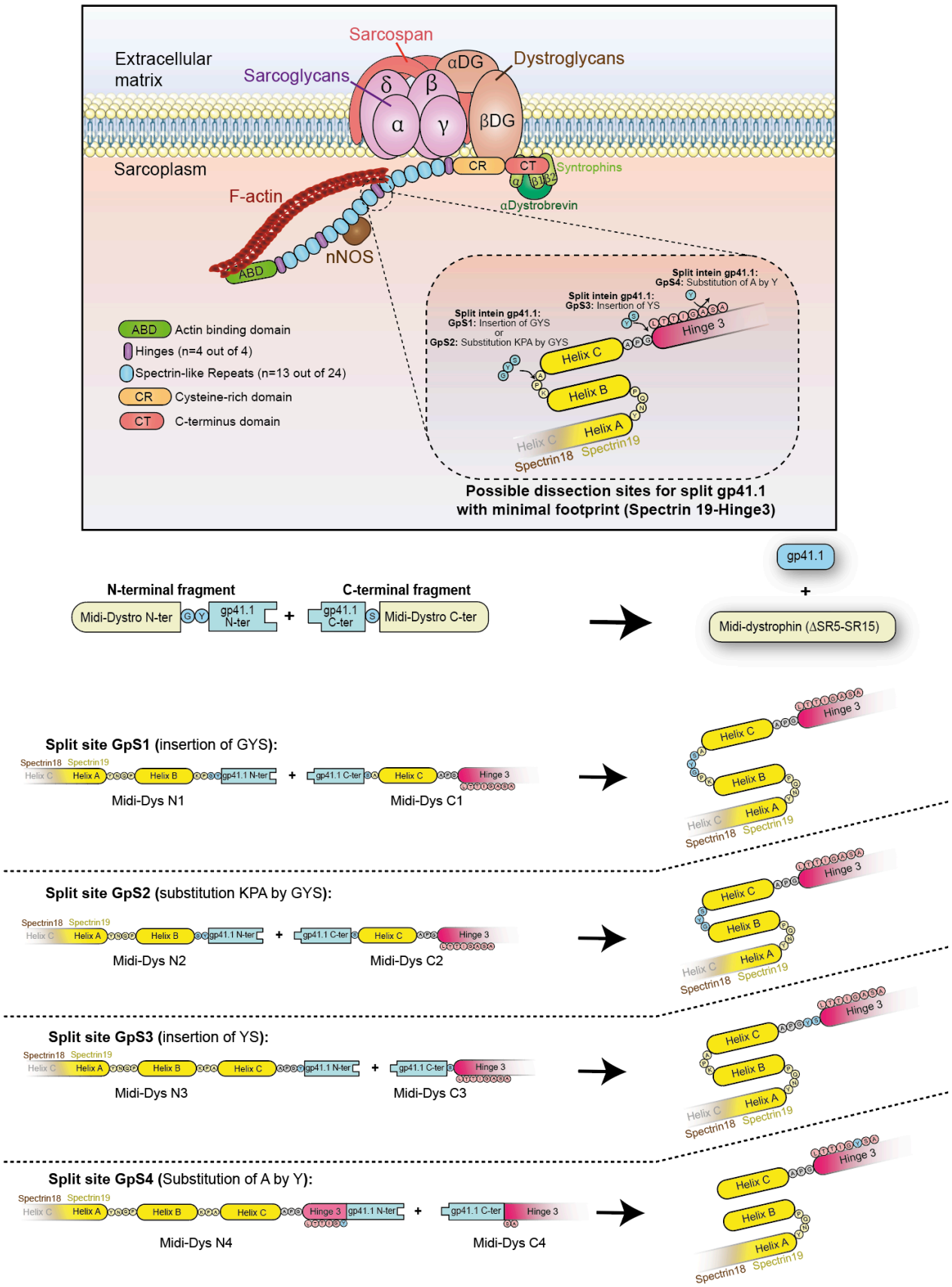
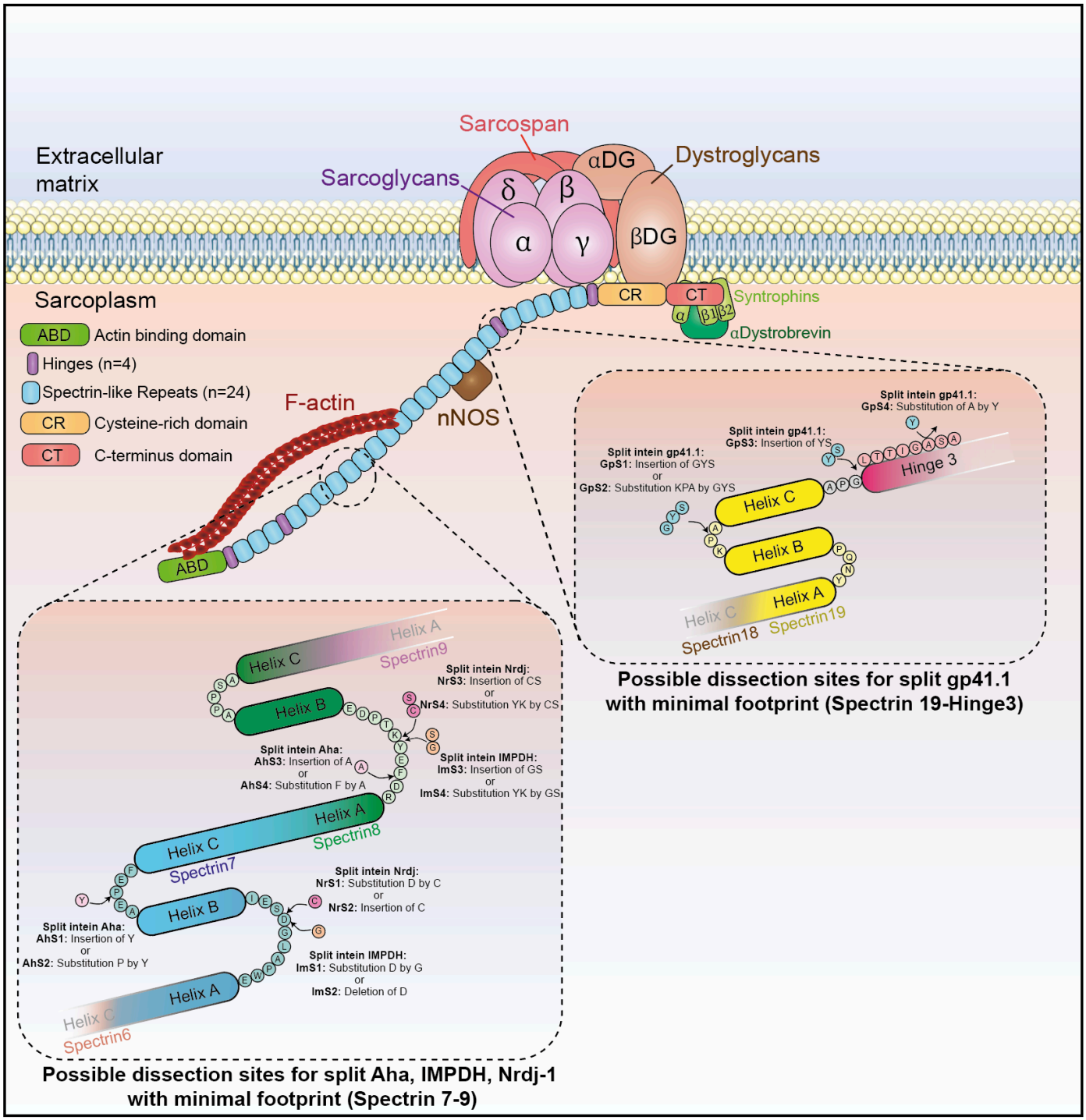


Extended data Fig. 1: Validation of GFP as a platform for split intein screening.
a) Diagram illustrating the topology of the GFP folding pattern with the chromophore, alpha helices, beta sheets, and the connecting loops with the two tested splitting sites (site#1: Glutamine 157-Lysine158, site#2: Glutamic acid 172-Aspartic acid 173). The numbers represent the delimiting residues at the beginning and the end of the secondary structures.
b) Mean fluorescence intensities of transfected HEK293 cells with either WT GFP or two mutated GFP harboring 8-amino acid insertions within the tested sites. **c**) Brightfield and fluorescent microscopy pictures of living HEK293 cells transfected with the WT or mutated GFP. Site#2 was more permissive to the insertion of an octapeptide and, thus, was selected as a splitting site where different split inteins were inserted for initial screening.



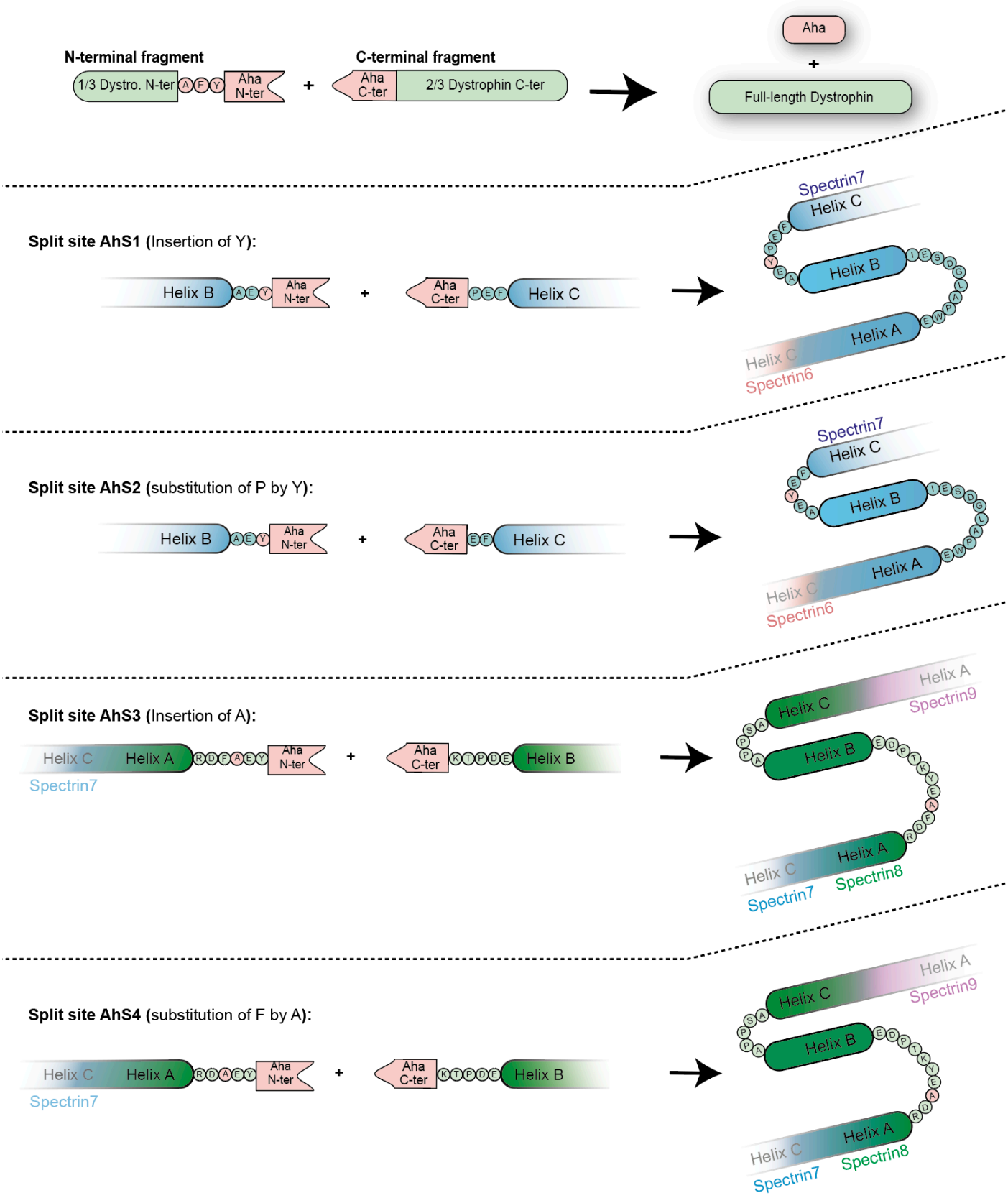
Extended data Fig. 2: Illustration of splitting sites tested to generate the midi-Dys (ΔSR5-15) using split gp41.1.

Split gp41.1 with an optimized footprint (**G**ly-T**y**r/**S**er) was tested in four locations. The resulting midi-Dys (SR5-15) harbor a footprint of between one to three residues. These additional residues were inserted in loops between Helix B and C of Spectrin-like repeat 19, the linker between Helix C and Hinge 3, or within Hinge 3 to minimize their impact on protein folding.

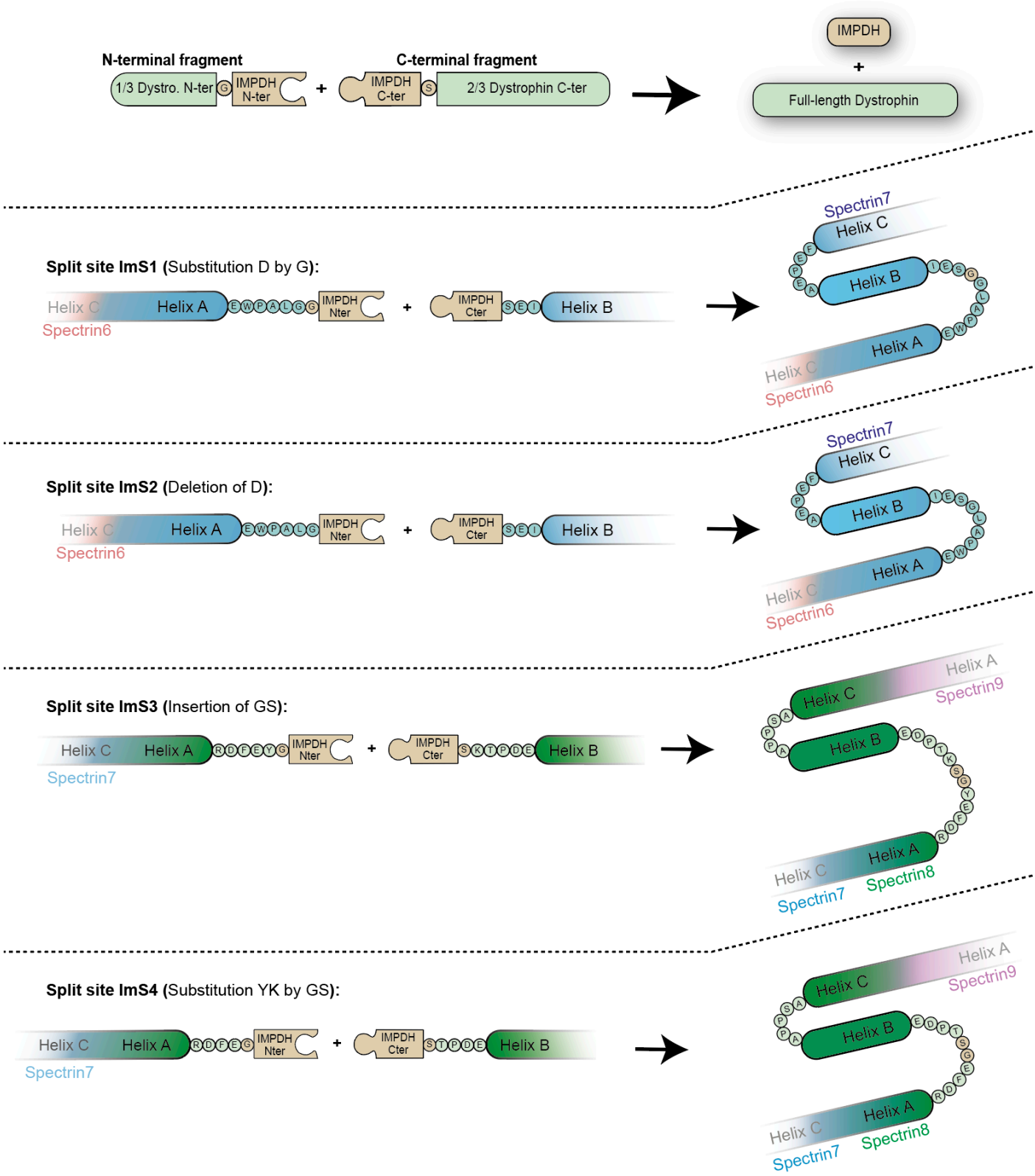


Extended data Fig. 3: Illustration splitting sites tested with triple AAV vector strategy to express the full-length dystrophin.

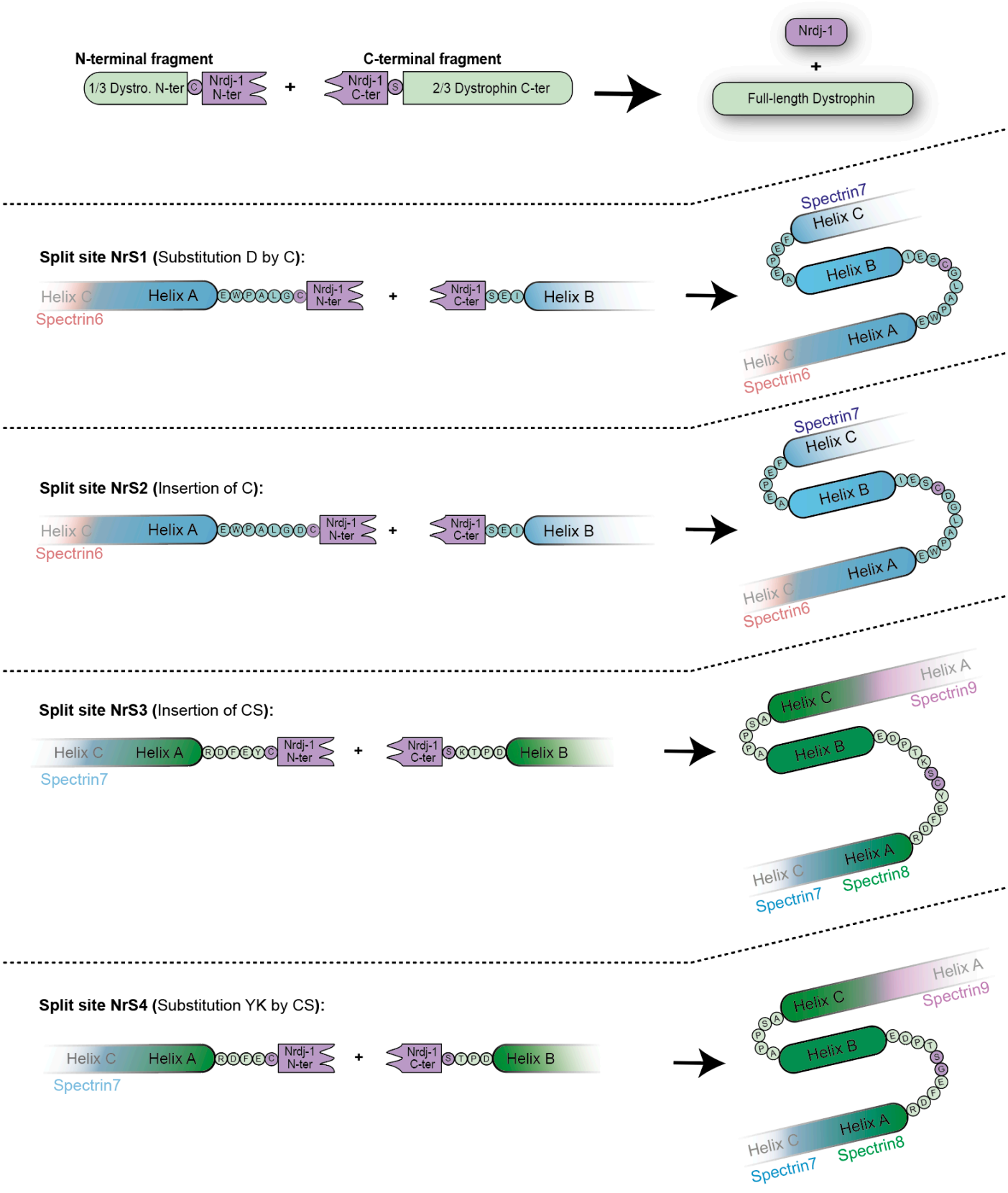
Optimized split Aha, IMPDH, and Nrdj1 with minimal footprints were inserted in loops located between Helix A & B or Helix B & C of Spectrin-like repeat 7 as well as Helix A & B of Spectrin-like repeat 8. These combinations join the N-terminal (from ABD to Spectrin-like repeat 7/8) to the middle fragment (from Spectrin-like repeat 7/8 to repeat 19/Hinge3). Split gp41.1 was evaluated to fuse the middle fragment to the C-terminal fragment at the same four splitting sites validated above with the dual strategy.



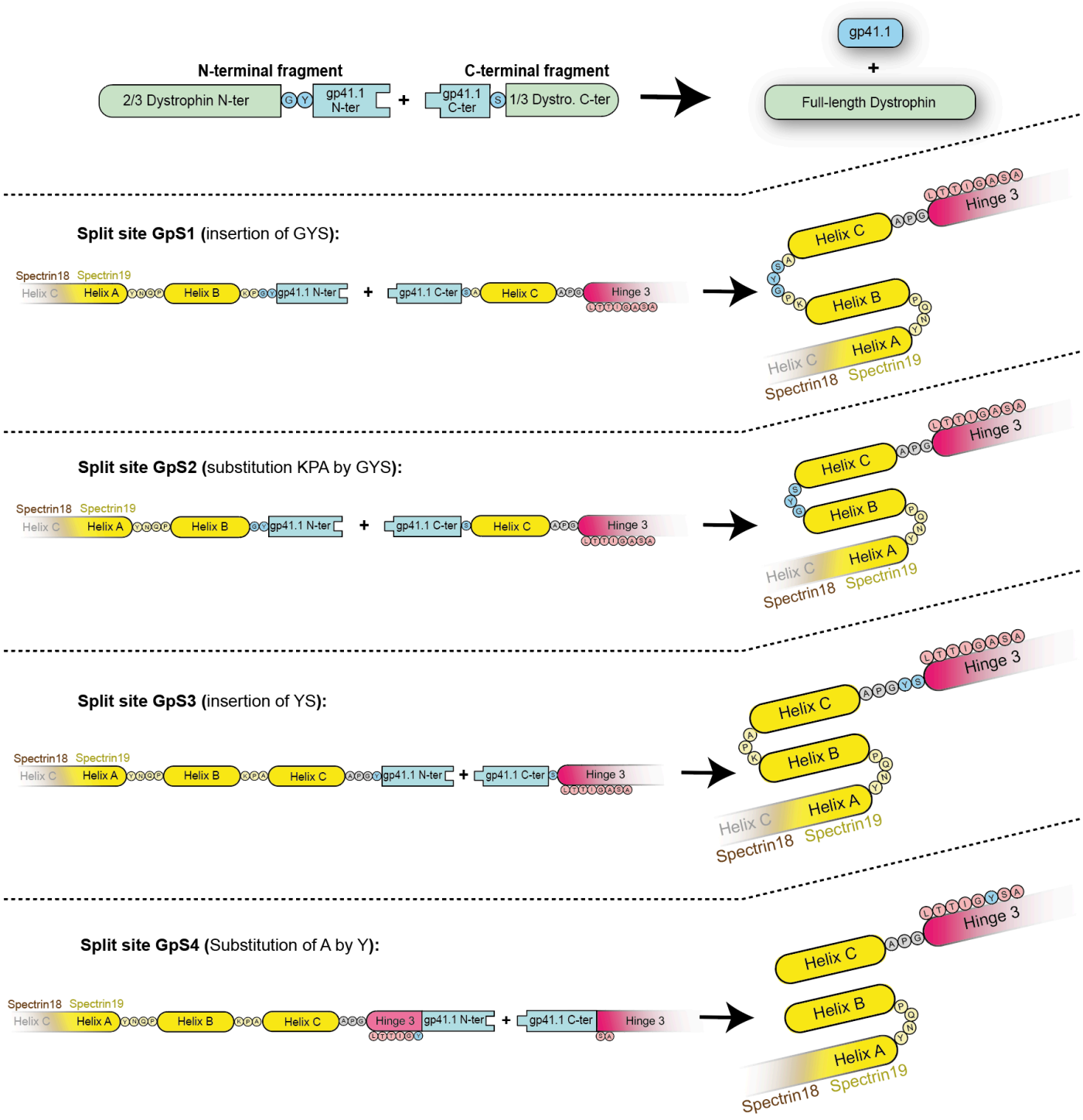
Extended data Fig. 4: Splitting sites tested with Aha.
Optimized split Aha with a minimal footprint (**A**la-Glu-T**y**r) was tested in four sites located in loops between Helix B & C of Spectrin-like repeat 7, or Helix A & B of Spectrin repeat 8. By using the native amino acids of dystrophin as part of the catalyzing reaction, the PTS mediated by the optimized Aha results in a minimal footprint with the insertion of only one residue (T**y**r in AhS1 or AhS2, or **A**la in AhS3 or AhS4).



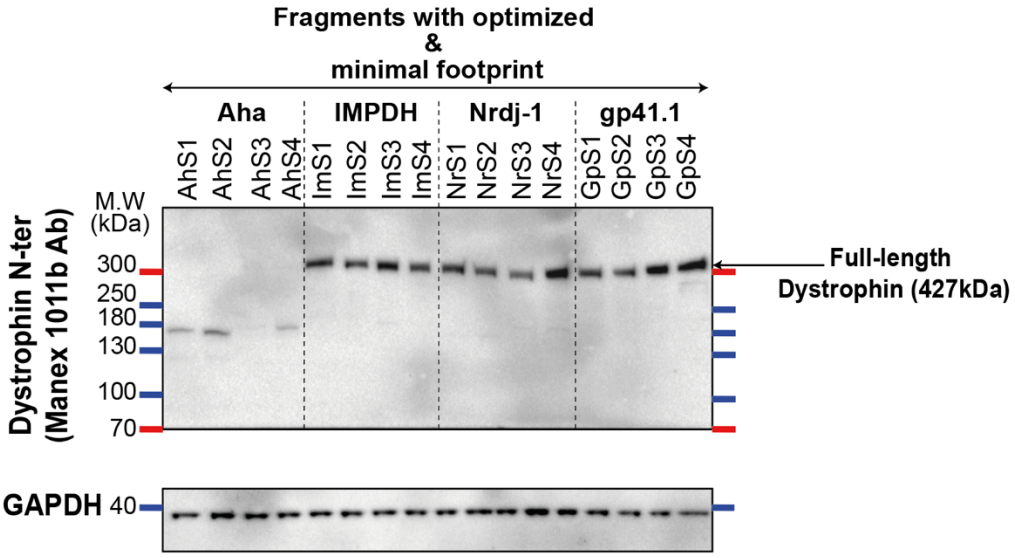
Extended data Fig. 5: Splitting sites tested with IMPDH.
Optimized split IMPDH with a minimal footprint (**G**ly/**S**er) was inserted in four sites located in loops between Helix A & B of Spectrin-like repeat 7 or Spectrin-like repeat 8. The PTS mediated by the optimized IMPDH results in a minimal footprint with the insertion of one (**G**ly in ImS1) or two residues (**G**ly-**S**er in Ims3 & ImS4), while in ImS2 native **G**ly and **S**er were used as catalyzing amino acids.



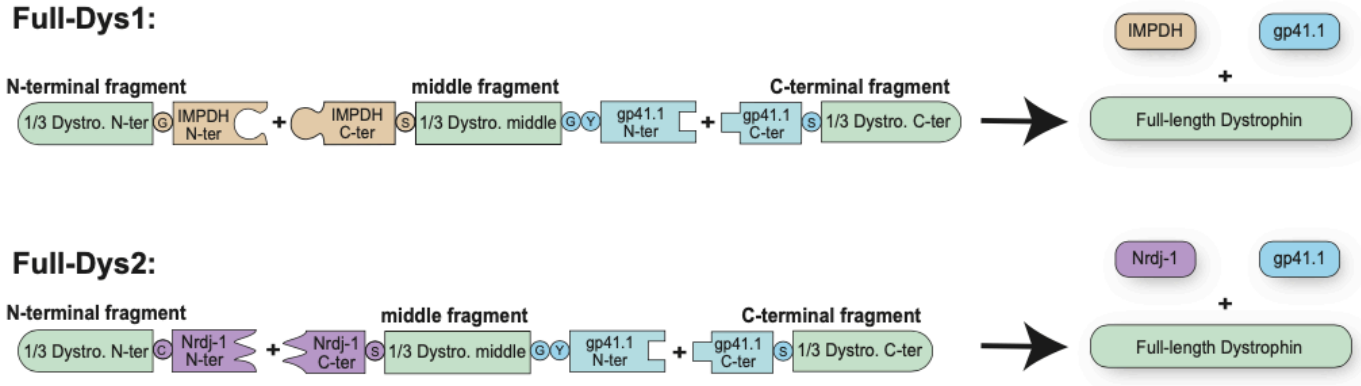
Extended data Fig. 6: Splitting sites tested with Nrdj-1.
The same splitting sites tested with IMPDH were evaluated with Nrdj-1. The PTS mediated by the optimized Nrdj-1 results in a minimal footprint with either substitution of the native Asp (**D**) by **C**ys (in NrS1), or insertion of a new residue (**C**ys in NrS2). Similarly, two residues (**C**ys-**S**er) were inserted in NrS3 in the loop between Helix A and B of Spectrin repeat 8, while in NrS4 substitution of native T**y**r-lys (**K**) by the footprint (**C**ys-**S**er).



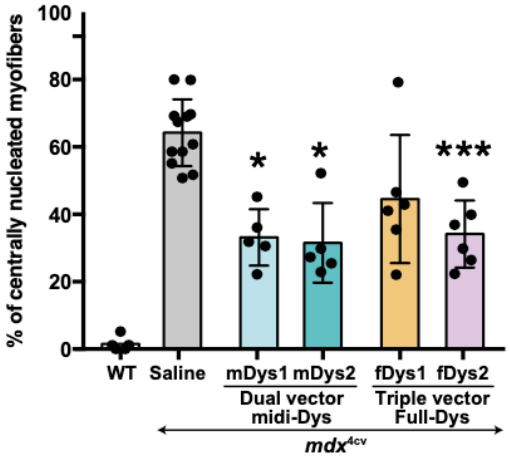
Extended data Fig. 7: Splitting sites tested with gp41.1 in triple vector strategy. Optimized split gp41.1 was tested to join an N-terminal dystrophin fragment carrying two-thirds of the protein (from ABD to Spectrin-like repeat 19/Hinge 3) to the C-terminal one-third of dystrophin (from Spectrin-like repeat 19/Hinge 3 to C-terminus) using the same split sites used to reconstitute the midi-Dys (SR5-15).



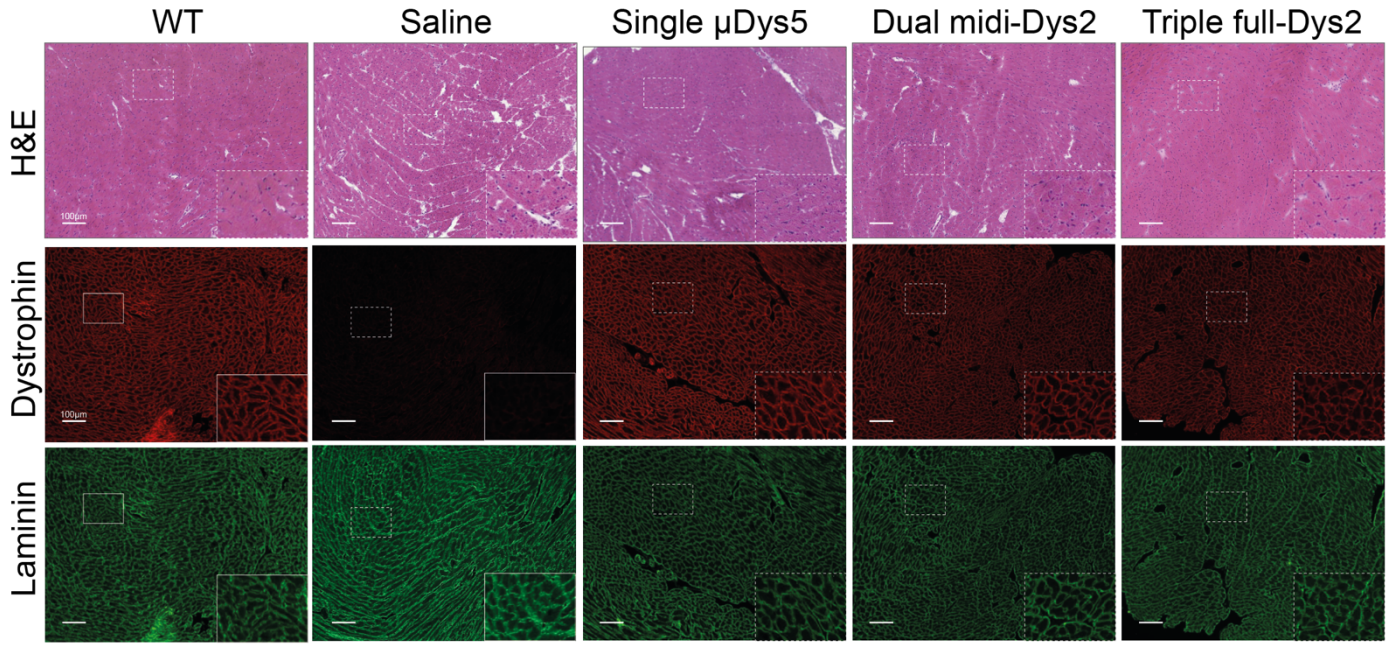
Extended data Fig. 8: Evaluation of PTS to express full-length dystrophin using the optimized split inteins.
The efficacy of each split intein to reconstitute full-length dystrophin from two fragments (one-third + two-thirds of dystrophin) was evaluated *in vitro* at four splitting sites (as depicted in Extended data Fig. 4 to 7)



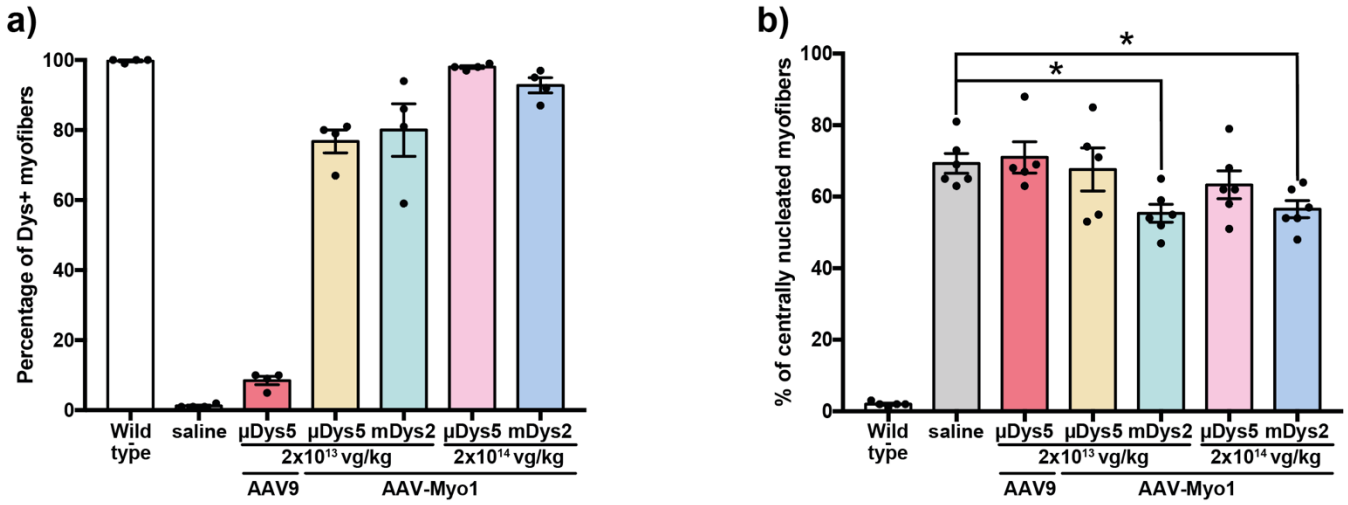
Extended data Fig. 9: Illustration of two combinations tested in the triple vector approach.
Two split inteins, IMPDH (split site: ImS4) and Nrdj-1 (split site: NrS4), were selected from previous screenings and tested for their ability to join the N-terminal to the middle fragment in a triple vector strategy. Split gp41.1 was inserted at the GpS4 split site to join the middle fragment to the C-terminal fragment.



Extended data Fig. 10: Quantification of myofibers with central nuclei following 5 weeks post AAV6 intramuscular injection:
Using TA cross-section stained with H&E, the percentage of myofibers with centrally-located nuclei was determined (*n*=5-8 sample per condition, ~400 myofibers were counted per sample). Data represent mean ± s.e.m. **p*<0.05, ****p*<0.001 for mice treated AAV *versus* saline group (ANOVA test followed by Dunnett's post hoc).

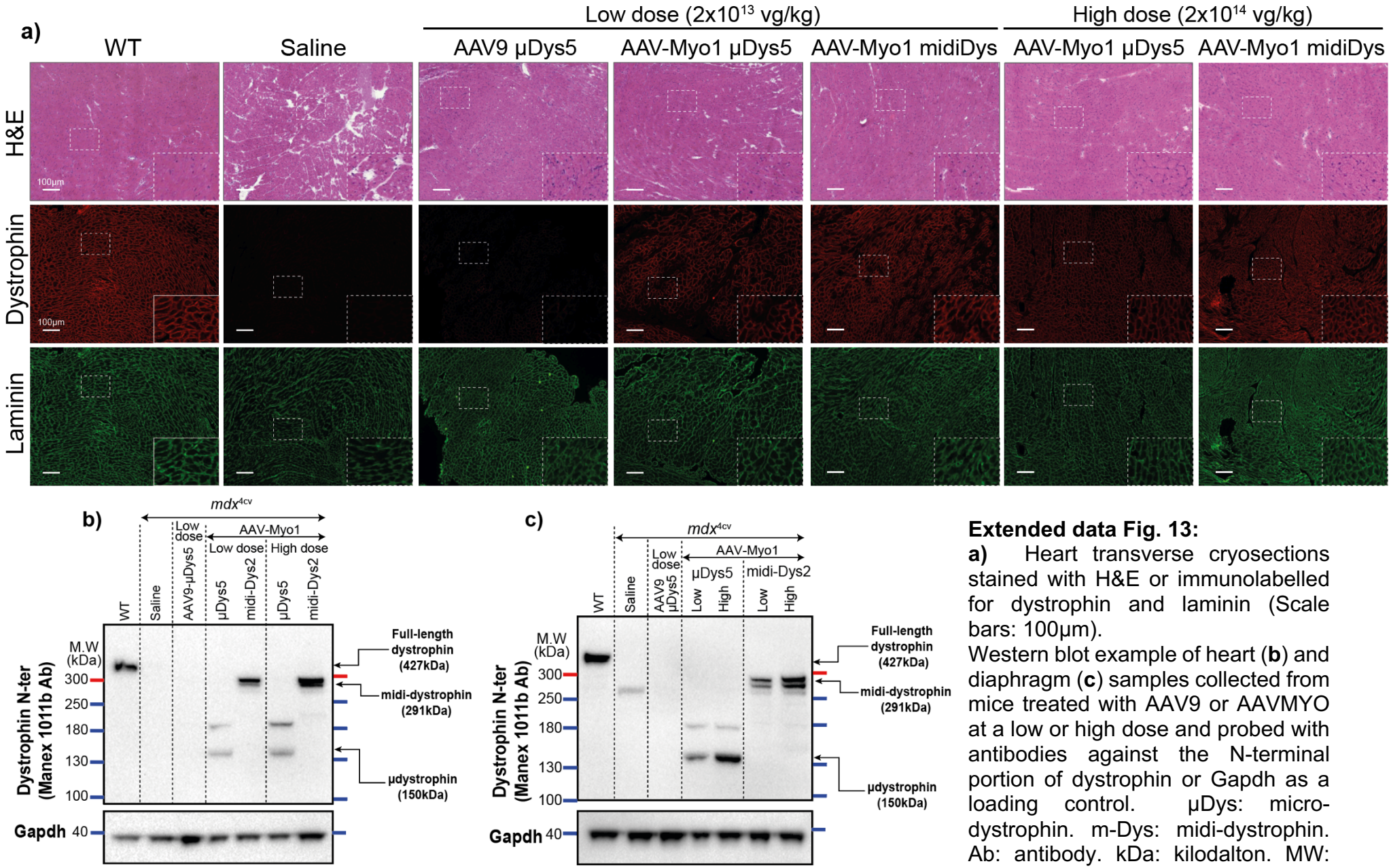


Extended data Fig. 11: General morphology and wide-spread expression of dystrophins in hearts.
Heart sections of control mice or *mdx*^{4cv} males treated with single, dual, or triple AAV6 vectors stained with H&E or double-immunolabeled for dystrophin and laminin (Scale bars: 100µm).



Extended data Fig. 12: Quantification of dystrophin biodistribution and centrally-nucleated myofibers across T.A muscle sections.

a) Percentage of dystrophin-positive myofibers (*n*=4, more than 2000 fiber per each sample). b) Percentage of myofibers with central nuclei measured from cross-sections stained with H&E (*n*=5-6, ~300 fibers per sample). *p<0.05 *versus* *mdx*^{4cv} mice treated with saline (ANOVA test followed by Dunnett's post hoc). µDys: micro-dystrophin. m-Dys: midi-dystrophin. Dys+: Dystrophin positive. vg: viral genome. kg: Kilogram



Extended data Fig. 13:

a) Heart transverse cryosections stained with H&E or immunolabelled for dystrophin and laminin (Scale bars: 100µm).

Western blot example of heart (**b**) and diaphragm (**c**) samples collected from mice treated with AAV9 or AAVMYO at a low or high dose and probed with antibodies against the N-terminal portion of dystrophin or Gapdh as a loading control. µDys: micro-dystrophin. m-Dys: midi-dystrophin. Ab: antibody. kDa: kilodalton. MW: molecular weight.