

ENGINEERED AAV CAPSID PROTEINS

Field of the Invention

The present disclosure relates to an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof. The present disclosure also relates to an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a knob domain of an ultralong CDR-H3, or a portion thereof. The disclosure further relates to a capsid comprising an engineered capsid protein, and to recombinant AAVs comprising said engineered capsid protein or capsid, and their use in therapy. The present disclosure also extends to methods of preparing said engineered capsid proteins, capsids and AAVs.

Background

Viral vector-mediated gene delivery has shown great promise for gene therapy applications, including approved treatments for inherited blindness, spinal muscular atrophy, and rare diseases. Among the viral vectors that are suitable for use in gene therapy, adeno-associated virus (AAV) has a number of advantageous features for clinical applications as compared to other gene therapy vectors such as adenovirus or lentivirus. AAV is non-pathogenic and is smaller in size than lentivirus, resulting in higher diffusion rates in tissues and greater potential for transducing large areas of tissue upon local injection. In addition, the AAV genome has minimal requirements for replication and packaging, enabling much of AAV to be replaced by a gene of interest and its regulatory elements. Genes delivered by AAV do not integrate into the host chromosome, which mitigates risk of undesired effects, and the genes nonetheless remain stable over long periods in host cells (Li & Samulski, *Nature Reviews Genetics* 21 (4): 255-272 (2020)). Finally, many naturally-occurring AAVs and AAV serotypes have been identified in nature, which may differ in their tissue tropism and in their transduction efficiency (Srivastava, *Current Opinion in Virology* 21: 75-80 (2016)). This variety is an advantage in gene therapy, as certain AAVs may be preferred to transduce specific cells or tissues.

The advantage of diversity has been further exploited by protein engineering to create artificial AAVs. The identity of the AAV is determined by the sequences of proteins that make up the protein shell, or capsid, of AAV. Among different variants, most sequence diversity is found in variable or hypervariable regions (VRs or HVRs) of the capsid proteins. By modifying amino

acids in the variable regions or in other parts of the capsid proteins, it has been shown that it was possible to create AAVs with new or improved properties, such as improved transduction efficiency or altered tissue tropism. One approach to capsid engineering is a random mutagenesis, in which entirely new capsids with random mutations are created and screened
5 for desired properties. In another approach, targeted changes are made to a capsid by substituting amino acid residues from the capsid of a first AAV serotype with corresponding residues from the capsid of a second AAV serotype and/or modifying specific amino acid residues in areas of the capsid that interact with binding partners on target cells.

Alternative approaches also include engineering a capsid by inserting heterologous binding
10 moieties, such as peptides, for example to try to enhance recognition of a binding partner on a target cell, thereby improving the ability of AAV to bind and transduce said target cell. For example, in one study, a 14-mer peptide (L14) derived from the laminin fragment P1 was inserted into the common VP3 region of AAV2, as L14 is a target of several cellular integrin receptors and can serve as a viral receptor (Girod et al., Nature Medicine 5: 1052-1056 (1999)).
15 In other studies, insertions of peptides into capsid proteins increased infectivity of rAAVs for retinal cells (e.g. US Patent Application No. US2020121746), cardiac tissues (e.g. PCT Publication No. WO20205889 A1), hepatic tissues (e.g. PCT Publication No. WO2020193799), muscle cells (e.g. PCT Publication No. WO19207132 A1), and GPI-anchored blood-brain barrier ligands (e.g. US Patent Application No. US2020325456).
20 Peptides have also been inserted into capsid proteins to facilitate evasion of neutralizing antibodies (e.g. US Patent No. US10745447).

In another example, a designed ankyrin repeat protein (DARPin), specific for the target HER2/*neu*, a receptor tyrosine kinase overexpressed on human cancer cells, was inserted into the N-terminus of VP2 capsid proteins in AAV2, and the resulting rAAV showed increased
25 transduction efficiency of cells expressing Her2, as well as targeting of Her2-expressing tumors (Münch et al., Mol Therapy 21 (1): 109-118 (2013)). In another example, VHH antibodies selective for cell surface markers (CD38, ARTC2.2, or CD38) were inserted into a VP1 capsid protein to confer binding and transduce cells expressing the cell-specific markers with high selectivity (e.g. US Patent Application No. US20210139563).

30 Nevertheless, while linear peptides are small, the extent of their binding and interaction with targets may be limited by their structure and properties. DARPins are small and can be engineered to bind targets with high affinity, but a disadvantage of DARPins is their concave

binding surface, rigidity, and incomplete randomization of amino acid residues in variable sites. These features limit the range of possible binding targets and require the use of additional engineering to compensate (Shilova & Deyev, *Acta Naturae* 11 (4): 42-53 (2019)).

Also, antibodies and antibody fragments may be relatively large (e.g., a Fab fragment may be around 50 kDa, while a VHH antibody may be around 15 kDa), and their structures and functions may be constrained by requirements for folding and post-translational modifications. The distance between the N- and C- termini in antibody molecules indicates that these molecules would be severely constrained if assembled within an AAV capsid. As an example, the distance between the N- and C- termini can be greater than 30 Å in some antibody fragments (e.g., about 37 Å in a scFv fragment and about 47 Å in a VHH fragment), while the AAV capsid is only about 250 Å in diameter. Finally, an early study indicated that manufacturing of AAV was adversely affected when a single chain antibody was inserted into the N- terminus of VP2 (Yang et al., *Human Gene Therapy* 9 (13) (2008)).

Therefore, there remains a need for new AAV capsid proteins, notably useful in therapy, having improved properties, for example, improved binding, cell and/or tissue specificity, transduction, and/or improved capsid assembly or manufacturability, while retaining a low immunogenicity.

Summary of Invention

The present inventors show for the first time that a AAV capsid protein may be engineered successfully that comprise a bovine ultralong CDR-H3 or a portion thereof, or a knob domain or portion thereof, conferring new binding properties to a capsid comprising said engineered capsid protein.

Surprisingly, the inventors demonstrate that an antibody fragment that comprises a bovine ultralong CDR-H3 or a portion thereof, or a knob domain of a bovine ultralong CDR-H3 or portion thereof may be inserted within a capsid protein, without disrupting either the folding of the antibody fragment, or the capsid assembly and virus production. This was particularly surprising as the knob domain is a disulphide-rich domain, with complex folding properties in particular in the reducing environment inside a cell.

In addition, the antibody fragment retains binding and specificity to its target, conferring new binding properties to the AAV capsids.

Advantageously, knob domains are small, highly specific and may be raised against a high diversity of antigens, and their ability to be incorporated into an AAV capsid and confer new binding properties to capsids provide engineered capsids and AAVs that may be useful in multiple applications, notably in therapeutic applications.

- 5 Thus, in one aspect, there is provided an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof. There is also provided an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a knob domain of an ultralong CDR-H3, or a portion thereof. In some embodiments, the antibody
10 fragment does not comprise a stalk of an ultralong CDR-H3. In some embodiments, the antibody fragment binds an antigen.

In some embodiments, the antibody fragment is 5 amino acids in length or more, 10 amino acids in length or more, 15 amino acids in length or more, 20 amino acids in length or more, 25 amino acids in length or more, 30 amino acids in length or more, 35 amino acids in length
15 or more, 40 amino acids in length or more, 45 amino acids in length or more, 50 amino acids in length or more, 55 amino acids in length or more, or 60 amino acids in length or more. In some embodiments, the antibody fragment is up to 69 amino acids in length. In some embodiments the antibody fragment is between 5 and 55, or between 15 and 50, or between 20 and 45 or between 25 and 40 amino acids in length.

- 20 In some embodiments, the antibody fragment comprises a (Z₁) X₁ C X₂ motif at its N-terminal extremity, wherein:

Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids;

- X₁ is any amino acid residue, preferably selected from the list consisting of Serine, Threonine,
25 Asparagine, Alanine, Glycine, Proline, Histidine, Lysine, Valine, Arginine, Isoleucine, Leucine, Phenylalanine and Aspartic acid; and,

C is cysteine; and,

X₂ is an amino acid selected from the list consisting of Proline, Arginine, Histidine, Lysine, Glycine and Serine.

In some embodiments, the antibody fragment comprises a sequence which is a variant of a naturally occurring sequence. In some embodiments, the antibody fragment further comprises at least one bridging moiety between two amino acids, optionally wherein the bridging moiety is a disulphide bond. In some embodiments, the antibody fragment is fully bovine. In
5 alternative embodiments, the antibody fragment is chimeric.

In some embodiments, the AAV capsid protein comprises a naturally occurring, or a variant or an artificial AAV sequence or a combination thereof. In some embodiments, the AAV capsid protein comprises a sequence of an AAV selected from the group consisting of AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10,
10 AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof.

In some embodiments, the antibody fragment is inserted within the AAV capsid protein, optionally via a linker. In some embodiments, the antibody fragment is inserted within the AAV capsid protein via one linker, wherein optionally the linker is genetically fused to the antibody fragment, optionally at its C-terminal end. In alternative embodiments, the antibody
15 fragment is inserted within the AAV capsid protein, via at least two linkers. In some embodiments, at least one linker is fused, optionally genetically, to the N-terminal of the antibody fragment, and at least one linker is fused, optionally genetically, to the C-terminal end of the antibody fragment. In some embodiments, the AAV capsid protein is a VP1, a VP2, or VP3.

20 In some embodiments, the antibody fragment is inserted within the common VP3 region of the AAV capsid protein, optionally, within the GH loop of the common VP3 region. In some embodiments, the antibody fragment is inserted within the variable region VR-IV of the AAV capsid protein. In some embodiments, the antibody fragment is inserted within the AAV capsid protein after amino acid residue Gly455 of AAV9 with reference to SEQ ID NO: 40 (or a
25 corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the variable region VR-V of the AAV capsid protein. In some embodiments, the antibody fragment is inserted within the AAV capsid protein after amino acid residue Asn498 of AAV9 with reference to SEQ ID NO:40, or a corresponding amino acid residue of another AAV. In some embodiments, the antibody fragment is inserted within the
30 variable region VR-VIII of the AAV capsid protein. In some embodiments, the antibody fragment is inserted within the AAV capsid protein after amino acid residue Gln588 of AAV9 with reference to SEQ ID NO:40, or a corresponding amino acid residue of another AAV. In

some embodiments, the AAV capsid protein is VP1 or VP2 and the antibody fragment is inserted within the VP1/VP2 common region of the capsid protein. In some embodiments, the antibody fragment is inserted within the N-terminus of the VP2 capsid protein. In some embodiments, the antibody fragment is inserted within the N-terminus of the VP2 capsid protein of AAV9 before amino acid residue Thr1 with reference to SEQ ID NO:41, or a corresponding amino acid in another AAV.

In another aspect, there is provided an AAV capsid comprising an engineered capsid protein according to the invention. In some embodiments, there is provided an AAV capsid comprising two engineered capsid proteins according to the invention, wherein the first and second engineered capsid proteins are respectively VP1 and VP2, or VP1 and VP3, or VP2 and VP3. In some embodiments, there is provided an AAV capsid comprising three engineered capsid proteins according to the invention, wherein the first, second and third engineered capsid proteins are respectively VP1, VP2 and VP3.

In another aspect, there is provided a nucleic acid encoding an engineered capsid protein or a capsid according to the invention. In another aspect, there is provided a vector comprising said nucleic acid.

In another aspect, there is provided a recombinant adeno-associated virus (rAAV) particle comprising the engineered capsid protein or capsid or nucleic acid or vector according to the invention, and a transgene. In some embodiments, the transgene encodes a peptide, a polypeptide or a nucleic acid molecule, optionally wherein the nucleic acid molecule is a small interfering RNA (siRNA), small or short hairpin RNA (shRNA), microRNA (miRNA).

In another aspect, there is provided a host cell comprising said nucleic acid or said vector and/or which produces a rAAV particle comprising the engineered capsid protein or capsid according to the invention, and a transgene.

In another aspect, there is provided a method for producing an AAV particle comprising an engineered capsid protein according to the invention, and a transgene, said method comprising: a) providing a first vector comprising a first nucleotide sequence encoding an AAV capsid protein, and a second nucleotide sequence encoding the antibody fragment; wherein the first nucleotide sequence and the second nucleotide sequence are genetically fused optionally via a nucleotide sequence coding for a linker; b) providing a second vector comprising the transgene, c) providing a third, Helper vector, d) transfecting a host cell with the first, second and third vector; and e) recovering the AAV from the host cell.

In another aspect, there is provided a vector or an AAV particle according to the invention, for use as a medicament.

Detailed Description of the Invention

As used in this disclosure, the singular forms “a”, “an”, and “the” include plural referents unless
5 the content clearly dictates otherwise. Thus, for example, reference to “a nucleic acid” includes “nucleic acids”, and the like.

The term “comprises” (comprise, comprising) should be understood to have its normal meaning in the art, i.e. that the stated feature or group of features is included, but that the term does not exclude any other stated feature or group of features from also being present. The term “consists of”
10 should also be understood to have its normal meaning in the art, i.e. that the stated feature or group of features is included, to the exclusion of further features. For every embodiment in which “comprises” or “comprising” is used, we anticipate a further embodiment in which “consists of” or “consisting of” is used. Thus, every disclosure of “comprises” should be considered to be a disclosure of “consists of”.

15 *Antibody fragments*

In one aspect, the present disclosure provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof. In some embodiments, the present disclosure provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment,
20 wherein the antibody fragment comprises a knob domain of a bovine ultralong CDR-H3, or a portion thereof.

Antibody fragments for use in the context of the present disclosure encompass whole bovine ultralong CDR-H3 and any portion thereof, preferably any functionally active portion thereof
25 (e.g. which binds to a target or an antigen of interest). Therefore, in some embodiments, the present disclosure provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof which bind an antigen. In some embodiments, the present disclosure provides an engineered capsid protein comprising an AAV capsid protein and an antibody
30 fragment, wherein the antibody fragment comprises a knob domain of a bovine ultralong CDR-H3, or a portion thereof which bind an antigen.

Whole antibodies also known as “immunoglobulins (Ig)” generally relate to intact or full-length antibodies i.e. comprising the elements of two heavy chains and two light chains, interconnected by disulphide bonds, which assemble to define a characteristic Y-shaped three-dimensional structure. Classical natural whole antibodies are monospecific in that they bind one antigen type, and bivalent in that they have two independent antigen binding domains. The terms “intact antibody”, “full-length antibody” and “whole antibody” are used interchangeably to refer to a monospecific bivalent antibody having a structure similar to a native antibody structure, including an Fc region as defined herein.

Each light chain is comprised of a light chain variable region (abbreviated herein as V_L) and a light chain constant region (C_L). Each heavy chain is comprised of a heavy variable region (abbreviated herein as V_H) and a heavy chain constant region (C_H) constituted of three constant domains C_{H1} , C_{H2} and C_{H3} , or four constant domains C_{H1} , C_{H2} , C_{H3} and C_{H4} , depending on the Ig class. The “class” of an Ig or antibody refers to the type of constant region and includes IgA, IgD, IgE, IgG and IgM and several of them can be further divided into subclasses, e.g. IgG1, IgG2, IgG3, IgG4. The constant regions of the antibodies may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (e.g., effector cells) and the first component (C1q) of the classical complement system.

The term “constant domain(s)”, “constant region”, as used herein are used interchangeably to refer to the domain(s) of an antibody which is outside the variable regions. The constant domains are identical in all antibodies of the same isotype but are different from one isotype to another. Typically, the constant region of a heavy chain is formed, from N to C terminal, by $CH1$ -hinge- $CH2$ - $CH3$ -optionnaly $CH4$, comprising three or four constant domains.

“Fc”, “Fc fragment”, “Fc region” are used interchangeably to refer to the C-terminal region of an antibody comprising the constant region of an antibody excluding the first constant region domain. Thus, Fc refers to the last two constant domains, C_{H2} and C_{H3} , of IgA, IgD, and IgG, or the last three constant domains of IgE and IgM, and the flexible hinge N-terminal to these domains.

The V_H and V_L regions of a whole antibody can be further subdivided into regions of hypervariability (or “hypervariable regions”) determining the recognition of the antigen, termed complementarity determining regions (CDR), interspersed with regions that are more

structurally conserved, termed framework regions (FR). Each V_H and V_L is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The CDRs and the FR together form a variable region. By convention, the CDRs in the heavy chain variable region of an antibody or antigen-binding fragment thereof are referred as CDR-H1, CDR-H2 and CDR-H3 and in the light chain variable region as CDR-L1, CDR-L2 and CDR-L3. They are numbered sequentially in the direction from the N-terminus to the C-terminus of each chain.

CDRs are conventionally numbered according to a system devised by Kabat et al. This system is set forth in Kabat et al., 1991, in *Sequences of Proteins of Immunological Interest*, US Department of Health and Human Services, NIH, USA (hereafter “Kabat et al. (supra)”). This numbering system is used in the present specification except where otherwise indicated. The Kabat residue designations do not always correspond directly with the linear numbering of the amino acid residues. The actual linear amino acid sequence may contain fewer or additional amino acids than in the strict Kabat numbering corresponding to a shortening of, or insertion into, a structural component, whether framework or complementarity determining region (CDR), of the basic variable domain structure. The correct Kabat numbering of residues may be determined for a given antibody by alignment of residues of homology in the sequence of the antibody with a “standard” Kabat numbered sequence.

The CDRs of the heavy chain variable domain are located at residues 31-35 (CDR-H1), residues 50-65 (CDR-H2) and residues 93-102 (CDR-H3) according to the Kabat numbering system. However, according to Chothia (Chothia, C. and Lesk, A.M. *J. Mol. Biol.*, 196, 901-917 (1987)), the loop equivalent to CDR-H1 extends from residue 26 to residue 32. Thus, unless indicated otherwise ‘CDR-H1’ as employed herein is intended to refer to residues 26 to 35, as described by a combination of the Kabat numbering system and Chothia’s topological loop definition. The CDRs of the light chain variable domain are located at residues 24-34 (CDR-L1), residues 50-56 (CDR-L2) and residues 89-97 (CDR-L3) according to the Kabat numbering system. Based on the alignment of sequences of different members of the immunoglobulin family, numbering schemes have been proposed and are for example described in Kabat et al., 1991, and Dondelinger et al., *Frontiers in Immunology*, Vol 9, article 2278 (2018).

Different species exhibit a diversity of CDR-H3 lengths. Some bovine antibodies have been characterized by unusually long CDR-H3 (so called “bovine ultralong CDR-H3”) with lengths

of up to 69 residues, representing 1-15 % of the bovine repertoire, whereas more conventional bovine antibodies have CDR-H3 of around 23 residues. Camelid single chain antibodies have up to 24 residues and shark IgNAR antibodies have up to 27 residues. The CDR-H3 are too long to be accommodated by any of these numbering schemes, but alternative systems have
5 been used, as the one discussed in Stanfield et al. (Stanfield, R. L., Wilson, I. A. & Smider, V. V. Conservation and diversity in the ultralong third heavy-chain complementarity-determining region of bovine antibodies. *Sci Immunol* 1, (2016) (hereinafter “Stanfield et al. (supra)”).

“Bovine CDR-H3” as used herein encompasses all CDR-H3 found in bovines, including bovine regular CDR-H3 and bovine ultralong CDR-H3.

10 The term “Bovine ultralong CDR-H3” refers to the subset of CDR-H3 which has the features of characterized ultralong CDR-H3 as defined hereinafter, notably comprising a duplication of the IGHVI-7 gene segment. The ultralong CDR-H3 has been found in bovine IgG of all classes.

Bovine ultralong CDR-H3 have been characterized by a very unusual tridimensional structure comprising a “stalk domain” and a “knob domain”. The stalk domain is composed of two
15 antiparallel β strands (each strand generally corresponding to about 12 residues). The knob domain is a disulfide rich domain which comprises a loop motif and sits atop of the stalk, which serves as a bridge to link the knob domain with the main bovine antibody scaffold. The CDR-H3 is derived from DNA rearrangement of variable (V), diversity (D), and joining (J) gene segments. The ultralong CDR-H3 are encoded by the VH_{BUL} (Bovine Ultra Long), D_{H2} , and
20 J_{H1} gene segments, and their length is due to an unusually long D_{H2} segment. Ultralong CDR-H3 have been characterized by a duplication of the IGHVI-7 gene segment.

The “stalk domain” of bovine ultralong CDR-H3 has been characterised by its structure notably. The skilled person will appreciate that the definition of a “stalk domain” may rely on
25 crystal structure analysis and/or sequencing information, notably as he will understand that the stalk domain position and structure may vary slightly from one ultralong CDR-H3 to another, e.g. in terms of size. The term “stalk domain” will be generally appreciated by the skilled person to correspond to the antiparallel β strands that bridge the knob domain with the main bovine antibody scaffold. The length of the stalk β strands can differ, notably from long β
30 strands (12 or more residues) to shorter β strands.

The skilled person will appreciate that the definition of a knob domain may rely on crystal structure analysis and/or sequencing information, notably as he will understand that the knob

domain position and structure may vary slightly from one ultralong CDR-H3 to another, e.g. in terms of size, cysteine content, disulphide bond content. In particular, the sequence of ultralong CDR-H3 can be determined by well-known sequencing methods, and the skilled person will be able to identify the minimal sequence which define a knob domain, based for example on a comparative analysis, with well characterised ultralong CDR-H3 as well as stalk and knob domains thereof, e.g. by alignment with well-known and/or standard nucleic and/or amino acid sequences, and/or based on crystal structure analysis.

As mentioned above, the ultralong CDR-H3 are too long to be accommodated by any of existing numbering scheme, but alternative systems have been used, as the one discussed in Stanfield et al. (supra). Structural analysis has also been provided for example by Wang et al. (Wang, F. *et al.* Reshaping antibody diversity. *Cell* 153, 1379-1393 (2013)).

The conserved Cysteine at position 92(Kabat) and the conserved Tryptophan at position 103(Kabat) respectively defines the start and the end of the CDR-H3. The germline encoded VH_{BUL} D_{H2} J_{H1} has the following sequence:

CTTVHQSCPDGYSYGYGCGYGYGCSGYDCYGYGGYGGYGGYGYSSYSYSYTYEYYVDA
WGQGLLVTVSS

(VH_{BUL}; followed by D_{H2} gene region in bold; followed by J_{H1} gene region underlined; The sequence coding the CDR-H3 is in italic, between positions 92 and 103 according to Kabat)

Kabat numbering system may be used for heavy-chain residues 1 to 100 and 101 to 228 but residues between 100 and 101 (corresponding to residues encoded by D_{H2} and J_{H1} genes) do not accommodate to the Kabat numbering system and may be numbered differently, for example sequentially with a D identifier, as described in Stanfield et al. (supra), with the conserved Cysteine residue at the start of D_{H2} being “D2”, followed by D3, D4 etc...).

Following Cys H92, the common motif TTVHQ (positions 93-97 in the germline VH_{BUL}, according to Kabat) starts the ascending strand of the β -stalk region of the CDR-H3. The length between the end of the VH_{BUL} and the “CPD” conserved motif in D_{H2} is variable due to differences in junctional diversity formed through V-D recombination. In Stanfield, those junctional residues are referred as “a,b,c” following H100 residue, depending on the length.

The D_{H2} region has been characterised to encode the knob domain and part of the descending strand of the stalk region. D_{H2} begins with a conserved Cysteine which is part of a conserved “CPD” motif in the germline sequence, which characterises the beginning of the knob domain.

The knob domain terminates at the beginning of the descending strand of the β -stalk region. The descending strand of the β -stalk region has been characterised by alternating aromatic-aliphatic residues in some ultralong CDR-H3. The descending strand of the β -stalk region ends with the residues encoded by the genetic J region, followed by residue H101, H102 according to Kabat.

In the context of the present disclosure, the minimal sequence that may define a knob domain corresponds to the portion of the ultralong CDR-H3 encapsulated by disulphide bonds, more particularly the minimal knob domain sequence starts from the first cysteine residue of an ultralong CDR-H3 and ends with the last cysteine residue of the ultralong CDR-H3. Therefore, a minimal knob domain typically comprises at least two cysteines. In one embodiment, the knob domain sequence starts from one residue preceding the first cysteine residue of an ultralong CDR-H3 and ends after the residue subsequent to the last cysteine residue of the ultralong CDR-H3. Additional amino acids may be present in the N-terminal extremity and/or in the C-terminal extremity of the knob domain sequence, preferably up to 5 additional amino acids may be present in the N-terminal and/or in the C-terminal extremity.

In one aspect, there is provided an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a knob domain of an ultralong CDR-H3, or a portion thereof. In one embodiment, the antibody fragment consists of the knob domain of a bovine ultralong CDR-H3, i.e. is a full-length knob domain, notably comprised between the ascending stalk and the descending stalk of the ultralong CDR-H3. In one embodiment, the antibody fragment comprises or consists of a portion of the knob domain of a bovine ultralong CDR-H3. In some embodiments, the antibody fragment binds to an antigen.

In some embodiments, the antibody fragment does not comprise a stalk of an ultralong CDR-H3, notably of a bovine ultralong CDR-H3. In some embodiments, the engineered capsid protein comprising an AAV capsid protein and an antibody fragment does not comprise a stalk of an ultralong CDR-H3 notably of a bovine ultralong CDR-H3.

In one embodiment, the antibody fragment comprises at least two, or at least four, or at least six, or at least eight, or at least ten cysteine residues. In one embodiment, the antibody fragment comprises two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen or fourteen cysteine residues. In one embodiment, the antibody fragment comprises between two cysteine

residues and ten cysteine residues. In one embodiment, the antibody fragment comprises between four cysteine residues and eight cysteine residues.

Two cysteine residues may bridge together to form a disulphide bond within the antibody fragment, for example within the knob domain.

- 5 In one embodiment, the antibody fragment comprises at least one, or at least two, or at least three, or at least four, or at least five disulphide bonds. In one embodiment, the antibody fragment comprises one, two, three, four, five, six, or seven disulphide bonds. In one embodiment, the antibody fragment comprises between one disulphide bond and five disulphide bonds. In one embodiment, the antibody fragment comprises between two
10 disulphide bonds and four disulphide bonds.

It will be appreciated that an increased content in cysteine residues will increase the possibility to form disulphide bonds within the antibody fragment. Such disulphide bonds contribute to form a loop motif within the antibody fragment, which may be advantageous to increase the stability, and/or rigidity and/or binding specificity and/or binding affinity of the antibody
15 fragment for a target or antigen of interest.

In one embodiment, the antibody fragment is 5 amino acids in length or more, 10 amino acids in length or more, 15 amino acids in length or more, 20 amino acids in length or more, 25 amino acids in length or more, 30 amino acids in length or more, 35 amino acids in length or more, 40 amino acids in length or more, 45 amino acids in length or more. In one embodiment,
20 the antibody fragment is up to 50 amino acids in length or up to 55 amino acids in length. In one embodiment, the antibody fragment is 5 amino acids in length or more, 10 amino acids in length or more, 15 amino acids in length or more, 20 amino acids in length or more, 25 amino acids in length or more, 30 amino acids in length or more, 35 amino acids in length or more, 40 amino acids in length or more, 45 amino acids in length or more, and is up to 55 amino acids
25 in length. In one embodiment, the antibody fragment is a portion of a bovine ultralong CDR-H3 which is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, or 55 amino acids in length. In one embodiment, the antibody fragment is a portion of a knob domain of a bovine ultralong CDR-H3 which is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,
30 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, or 55 amino acids in length. In one embodiment, the antibody fragment is between 5 and 55, or between 15 and 50, or between 20

and 45, or between 25 and 40 amino acids in length. In one embodiment, the antibody fragment is a knob domain of a bovine ultralong CDR-H3 which is between 5 and 55, or between 15 and 50, or between 20 and 45, or between 25 and 40 amino acids in length.

In one embodiment, the antibody fragment comprises a (Z₁) X₁ C X₂ motif at its N-terminal
5 extremity, wherein:

- a. Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids; and,
- b. X₁ is any amino acid residue; and,
- c. C is cysteine; and,
- 10 d. X₂ is an amino acid selected from the list consisting of Proline, Arginine, Histidine, Lysine, Glycine and Serine.

In one embodiment, the antibody fragment comprises a knob domain or any portion thereof which comprises a (Z₁) X₁ C X₂ motif at its N-terminal extremity, as defined above. Z₁ as
15 defined in the present invention represents any amino acid or any sequence of 2, 3, 4, or 5 independently selected amino acids that may be the same or different. In one embodiment, Z₁ is 1 amino acid. In another embodiment, Z₁ is 2 amino acids, which may be the same or different. In another embodiment, Z₁ is 3 amino acids, which may be the same or different. In another embodiment, Z₁ is 4 amino acids, which may be the same or different. In another
20 embodiment, Z₁ is 5 amino acids, which may be the same or different.

In one embodiment, X₁ is selected from the list consisting of Serine, Threonine, Asparagine, Alanine, Glycine, Proline, Histidine, Lysine, Valine, Arginine, Isoleucine, Leucine, Phenylalanine and Aspartic acid. Thus, in one aspect, the invention provides an antibody fragment, which comprises a (Z₁) X₁ C X₂ motif at its N-terminal extremity, wherein:

- 25 a. Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids; and,
- b. X₁ is any amino acid residue, preferably selected from the list consisting of Serine, Threonine, Asparagine, Alanine, Glycine, Proline, Histidine, Lysine, Valine, Arginine, Isoleucine, Leucine, Phenylalanine and Aspartic acid; and,
- 30 c. C is cysteine; and,
- d. X₂ is an amino acid selected from the list consisting of Proline, Arginine, Histidine, Lysine, Glycine and Serine.

In one embodiment, the antibody fragment comprises a knob domain or any portion thereof which comprises a (Z₁) X₁ C X₂ motif at its N-terminal extremity, as defined above. In one embodiment, the antibody fragment comprises a (Z₁)X₁ C X₂ motif at its N-terminal extremity, wherein C is cysteine; and X₁ is selected in the list consisting of Serine (S), Threonine (T), Asparagine (N), Alanine (A), Glycine (G), Proline (P), Histidine (H), Lysine (K), Valine (V), Arginine (R), Isoleucine (I), Leucine (L), Phenylalanine (F) and Aspartic acid (D), and X₂ is selected from the list consisting of Proline (P), Arginine (R), Histidine (H), Lysine (K), Glycine (G) and Serine (S), and wherein Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids.

In one embodiment, the N-terminal extremity of the antibody fragment comprises a motif which comprises 3 amino acid residues, corresponding to a X₁C X₂ motif, selected in the list consisting of SCP, TCP, NCP, ACP, GCP, PCR, HCP, SCR, KCP, VCP, TCH, RCP, ICP, ICR, HCR, LCR, SCK, SCG, NCP, TCS, DCP and FCR.

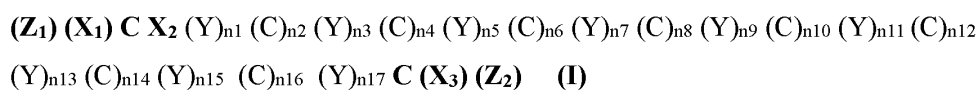
In one embodiment, the N-terminal extremity of the antibody fragment is initiated by a motif selected in the list consisting of (Z₁)SCP, (Z₁)TCP, (Z₁)NCP, (Z₁)ACP, (Z₁)GCP, (Z₁)HCP, (Z₁)KCP, (Z₁)VCP, (Z₁)RCP, (Z₁)ICP, (Z₁)DCP, wherein Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids.

In one embodiment, the antibody fragment comprises a (AB)_n and/or (BA)_n motif, wherein A is any amino acid residue, B is an aromatic amino acid selected from the group consisting of: tyrosine (Y), phenylalanine (F), tryptophan (W), and histidine (H), and wherein n is 1, 2, 3 or 4.

In one embodiment, A is an aliphatic amino acid residue. An aliphatic amino acid is an amino acid containing an aliphatic side chain functional group. Aliphatic amino acid residues include Alanine, isoleucine, leucine, proline, and valine.

In one embodiment, the antibody fragment comprises a motif of 2-8 amino acids which is rich in aromatic and/or aliphatic amino acids. In one embodiment, the antibody fragment comprises a motif of 2-8 amino acids which comprises at least 2, or at least 3 or at least 4, or at least 5 amino acids selected from the group consisting of: tyrosine (Y), phenylalanine (F), tryptophan (W), and histidine (H).

In some embodiments, the antibody fragment comprises or consists of the sequence of formula (I):



5 wherein:

C represents one cysteine residue; and,

Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids; and,

10 X₁ is present or absent, and when X₁ is present, X₁ is any amino acid residue, preferably selected from the list consisting of Serine, Threonine, Asparagine, Alanine, Glycine, Proline, Histidine, Lysine, Valine, Arginine, Isoleucine, Leucine, Phenylalanine and Aspartic acid; and,

X₂ is selected from the list consisting of Proline, Arginine, Histidine, Lysine, Glycine and Serine; and,

15 Z₂ is present or absent, and when Z₂ is present, Z₂ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids; and,

n₂, n₄, n₆, n₈, n₁₀, n₁₂, n₁₄ and n₁₆ are independently 0 or 1; and,

Y represents any amino acid or any sequence of amino acids that may be the same or different; and,

20 n₁, n₃, n₅, n₇, n₉, n₁₁, n₁₃, n₁₅ and n₁₇ represent the number of amino acids in Y, and are independently selected from 0 to 22, preferably from 1 to 15; and,

at least one of n₁, n₃, n₅, n₇, n₉, n₁₁, n₁₃, n₁₅ and n₁₇ is not equal to 0; and,

25 X₃ is present or absent, and when X₃ is present, X₃ represents any amino acid, preferably selected from the list consisting of Leucine, Serine, Glycine, Threonine, Tryptophan, Asparagine, Tyrosine, Arginine, Isoleucine, aspartic acid, Histidine, Glutamic acid, Valine, Lysine, Proline; and,

wherein the peptide is up to 55 amino acids in length.

Z₁ represents any amino acid or any sequence of 2, 3, 4, or 5 independently selected amino acids that may be the same or different. In one embodiment, Z₁ is 1 amino acid. In another embodiment, Z₁ is 2 amino acids, which may be the same or different. In another embodiment,

Z₁ is 3 amino acids, which may be the same or different. In another embodiment, Z₁ is 4 amino acids, which may be the same or different. In another embodiment, Z₁ is 5 amino acids, which may be the same or different.

5 Z₂ represents any amino acid or any sequence of 2, 3, 4, or 5 independently selected amino acids that may be the same or different. In one embodiment, Z₂ is 1 amino acid. In another embodiment, Z₂ is 2 amino acids, which may be the same or different. In another embodiment, Z₂ is 3 amino acids, which may be the same or different. In another embodiment, Z₂ is 4 amino acids, which may be the same or different. In another embodiment, Z₂ is 5 amino acids, which may be the same or different.

10

Z₁ and Z₂ may comprise any amino acid as long as the properties of the antibody fragment otherwise defined is retained, e.g. binding capability to an antigen of interest.

Brackets are generally used for optional residues or sequences. For example, (C) generally indicates an optional Cysteine residue, in the context of the present disclosure. In some
15 embodiments, the antibody fragment comprises a knob domain of an ultralong CDR-H3 or any portion thereof and comprises or consists of the sequence of formula (I) as described above.

In some embodiments, the antibody fragment of the present invention specifically binds to an antigen of interest, i.e. comprises a specific binding domain to an antigen of interest.
20 “Specifically,” as employed herein is intended to refer to a binding domain that only recognises the antigen to which it is specific or a binding domain that has significantly higher binding affinity to the antigen to which is specific compared to affinity to antigens to which it is non-specific, for example 5, 6, 7, 8, 9, 10 times higher binding affinity.

Preferably, the antibody fragment of the present invention has a specific binding affinity (as
25 measured by its dissociation constant K_D) for its cognate antigen of 10⁻⁵ M or less, 10⁻⁶ M or less, 10⁻⁷ M or less, 10⁻⁸ M or less, 10⁻⁹ M or less, 10⁻¹⁰ M or less, or 10⁻¹¹ M or less. In one embodiment, the antibody fragment of the present invention has a specific binding affinity (as measured by its dissociation constant K_D) for its cognate antigen between 1. 10⁻⁷ M and 1. 10⁻⁸ M, or between 1. 10⁻⁸ M and 1. 10⁻⁹ M, or between 1. 10⁻⁹ M and 1. 10⁻¹⁰ M.

30 Affinity can be measured by known techniques such as surface plasmon resonance techniques including Biacore™. Affinity may be measured at room temperature, 25°C or 37°C. Affinity

may be measured at physiological pH, i.e. at about pH 7.4. In one embodiment, the affinity values as described above are measured using Biacore, notably Biacore 8K, at pH 7.4.

It will be appreciated that the affinity of antibodies fragments provided by the present invention may be altered using any suitable method known in the art.

5 *Antibody fragment variants*

In some aspects, the antibody fragment comprises a sequence which is a variant of a naturally occurring sequence of a bovine antibody, i.e. of a bovine ultralong CDR-H3 or portion thereof. In other words, the present disclosure provides variants of antibody fragments as described above, which comprise non-naturally occurring sequences, i.e. which have been further
10 engineered, for example to improve at least one pharmacokinetic and/or biological function. In such aspects, the antibody fragment comprising a naturally occurring sequence may be referred as "parent".

The present invention also includes antibody fragments, i.e. bovine ultralong CDR-H3 or portions thereof, which comprise sequences which are at least 90%, 91%, 92%, 93%, 94%,
15 95%, 96%, 97%, 98% or 99% similar or identical to a sequence given herein. "Identity", as used herein, indicates that at any particular position in the aligned sequences, the amino acid residue is identical between the sequences. "Similarity", as used herein, indicates that, at any particular position in the aligned sequences, the amino acid residue is of a similar type between the sequences. For example, leucine may be substituted for isoleucine or valine. Other amino
20 acids which can often be substituted for one another include but are not limited to:

- phenylalanine, tyrosine and tryptophan (amino acids having aromatic side chains);
- lysine, arginine and histidine (amino acids having basic side chains);
- aspartate and glutamate (amino acids having acidic side chains);
- asparagine and glutamine (amino acids having amide side chains); and
25 - cysteine and methionine (amino acids having sulphur-containing side chains).

Degrees of identity and similarity can be readily calculated by methods well known, for example the BLAST™ software available from NCBI.

In one embodiment, antibody fragments of the present disclosure are processed to provide
30 improved affinity for a target or antigen. Such variants can be obtained by a number of affinity maturation protocols including mutating the CDR, chain shuffling, use of mutator strains of *E. coli*, DNA shuffling, phage display and sexual PCR. Vaughan *et al* (Nature Biotechnology, 16,

535-539, 1998) discusses these methods of affinity maturation. Another method useful in the context of the present disclosure to improve binding of the antibody fragment at a binding site on the target or antigen of interest is a method as described in WO2014/198951. Improved affinity as employed herein in this context refers to an improvement over the starting antibody
5 fragment. Affinity can be measured as described above.

In one embodiment, the antibody fragment is a variant of a parent bovine antibody fragment which has an affinity which is at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% higher than the affinity of the parent bovine antibody fragment, as measured for example by
10 Biacore.

"Truncated variants" when referring to antibody fragments are those with one or more amino acids in the native or starting amino acid sequence removed from either terminus of the polypeptide.

In some embodiments, the antibody fragment is a variant which has been engineered to
15 comprise a disulfide bond which is in a non-naturally occurring position. This may be engineered into the molecule by introducing cysteine(s) into the amino acid chain at the position or positions required. This non-natural disulfide bond is in addition to or as an alternative to the natural disulfide bond(s) which may be present in the parent antibody fragment. The cysteine(s) in natural positions can be replaced by an amino acid such as serine
20 which is incapable on forming a disulfide bridge. Introduction of engineered cysteines can be performed using any method known in the art. These methods include, but are not limited to, PCR extension overlap mutagenesis, site-directed mutagenesis or cassette mutagenesis (see, generally, Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, 1989; Ausubel et al., Current Protocols in
25 Molecular Biology, Greene Publishing & Wiley-Interscience, NY, 1993). Site-directed mutagenesis kits are commercially available, e.g. QuikChange® Site-Directed Mutagenesis kit (e.g. Stratagene, La Jolla, CA). Cassette mutagenesis can be performed based on Wells *et al.*, 1985, Gene, 34:315-323. Alternatively, mutants can be made by total gene synthesis by annealing, ligation and PCR amplification and cloning of overlapping oligonucleotides.

30

In one aspect, it may be useful to decrease or remove the cysteine residues and/or disulfide bonds in an antibody fragment of the disclosure, e.g. to lower the risk of immunogenicity, i.e. of side reactions occurring during or after the administration to a patient. In such aspect, one

or all of the cysteine(s) in natural positions can be replaced by an amino acid such as serine which is incapable on forming a disulfide bridge. It will be appreciated that alternative bridging moieties may be used to stabilise and/or form a cyclised antibody fragment in the absence of cysteine residues. In one embodiment, the antibody fragment is a variant which has been
5 engineered to remove the cysteine residues and which comprises at least one bridging moiety as defined in the present disclosure. In one embodiment, the antibody fragment is a variant which has been engineered to contain only one, or only two, or only three, or only four, cysteine residues, and/or to contain only one or only two disulphide bonds and which optionally further comprises at least one bridging moiety as defined in the present disclosure.

10 Additional modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of tyrosinyl, seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (Creighton, T. E., *Proteins: Structure and Molecular Properties*, W.H. Freeman and Co., San Francisco, 1983, pp. 79-86).

The antibody fragment of the invention may be cyclised. Cyclisation may be advantageous to
15 confer more resistance to proteolysis, resulting notably in an improved stability. Therefore, in one embodiment, the antibody fragment of the present disclosure further comprises a bridging moiety between two amino acids. Cyclised antibody fragments include any antibody fragments that have as part of their structure one or more cyclic features such as a loop, bridging moiety, and/or an internal linkage. As used herein, the term "bridging moiety" refers to one or more
20 components of a bridge formed between two adjacent or non-adjacent amino acids, unnatural amino acids or non-amino acids in an antibody fragment. Bridging moieties may be of any size or composition.

In one embodiment, a bridging moiety may be between the amino acid residue in N-terminal position and the amino acid residue in C-terminal position such as to create a head-to-tail
25 cyclisation. In one embodiment, a bridging moiety may be between amino acids which are not in terminal position. In one embodiment, the antibody fragment comprises only one bridging moiety between two amino acids. In another embodiment, the antibody fragment comprises more than one bridging moiety between two amino acids, e.g. two, or three, or five bridging moieties, each one being between two amino acids. In one embodiment, the bridging moiety
30 comprises a disulphide bond. In one embodiment, the disulphide bond is formed between two naturally occurring cysteine residues. In another embodiment, the disulphide bond is formed

between cysteine residues, with at least one cysteine residue being engineered, as described above.

In one embodiment, the antibody fragment of the invention is fully bovine. In such embodiment, each and every residue is derived from a bovine germline sequence. In some
5 embodiments, each and every residue is derived from a bovine germline sequence which has undergone affinity maturation for an antigen.

In one embodiment, the antibody fragment of the invention is chimeric. The term "chimeric" refers to an antibody fragment comprising at least two portions, one being derived from a particular source or species, such as bovine, while the other portion is derived from a different
10 source or species, such as human. In one embodiment, the antibody fragment is human/bovine chimeric. In one embodiment, the antibody fragment comprises at least one residue derived from a human sequence. In one embodiment, the antibody fragment comprises one, two, three, four, five, or more residues derived from a human sequence. In one embodiment, the antibody fragment comprises at least two residues derived from a human sequence wherein the at least
15 two residues are contiguous.

In one embodiment, the antibody fragment of the invention is synthetic. The term "synthetic" refers to an antibody fragment that has been produced de novo by synthesis, notably by chemical synthesis. Chemical synthesis approaches have been described, such as solid phase polypeptide synthesis (see e.g., Coin, I et al. (2007); Nature Protocols 2(12):3247-56).

20 In one embodiment, the antibody fragment comprises a sequence selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12, 26, 27, 30, 31, 34 or 35, or any variant thereof as described herein. In one embodiment, the antibody fragment comprises a bovine ultralong CDR-H3 and comprises a sequence selected from the group consisting of SEQ ID NO:4, 8, 12, 27, 31 or 35 or any variant thereof as described herein. In one embodiment, the antibody
25 fragment comprises a knob domain of a bovine ultralong CDR-H3 and comprises a sequence selected from the group consisting of SEQ ID NO: 2, 6, 10, 26, 30, or 34, or any variant thereof as described herein.

AAV capsid proteins

The term "adeno-associated virus or AAV" encompasses any AAV, for example, any serotype
30 or subtype of AAV, any forms of AAV such as a naturally-occurring AAV or any variant or

any derivative thereof, and/or any engineered (i.e. artificial), or recombinant forms of AAV, or any combination of these.

The AAV single-stranded DNA genome comprises two inverted terminal repeats (ITRs) and two open reading frames, containing structural (*cap*) and packaging (*rep*) genes (Hermonat et al., 1984). The AAV genome typically comprises packaging genes, such as *rep* and/or *cap* genes which encode packaging functions for an AAV particle. The *rep* gene encodes one or more of the proteins Rep78, Rep68, Rep52 and Rep40 or variants thereof. The *cap* gene encodes one or more capsid proteins such as VP1, VP2 and VP3 or variants thereof. These proteins make up the capsid of a viral particle such as an AAV particle. In some embodiments, the AAV capsid protein is a virion protein (VP).

VP1, VP2, and VP3 refer to three distinct proteins found in the capsid (or protein shell) of an AAV. The terms “VP” encompasses any AAV VP, for example, from any AAV serotype or subtype of AAV, any forms of AAV such as a naturally-occurring AAV or any variant or any derivative thereof, and/or any engineered (i.e. artificial), or recombinant forms of AAV, or any combination of these.

The VPs are encoded by a single capsid gene *cap* and are produced by alternative mRNA splicing of the transcript and alternative start codon usage. VP1, VP2, and VP3 differ from one another only in their N-terminus.

“VP1” refers to the VP1 capsid protein, which is about 79-82 kDa in size and comprises about 713-738 amino acids. VP1 comprises the entire sequence of VP2 and VP3, and in addition, at its N-terminal, an amino acid sequence which is unique to VP1 and may be referred to as “the VP1 unique region.”

“VP2” refers to the VP2 capsid protein, which is about 64-67 kDa in size and comprises about 580-601 amino acids. VP2 comprises the entire sequence of VP3, and in addition, at its N-terminal, a VP2 amino acid sequence that is not present in VP3 but is shared by VP1 and VP2 and may be referred to as “the VP1/VP2 common region.”

“VP3” refers to the VP3 capsid protein, which is about 59-61 kDa in size and which comprises about 524-544 amino acids. The VP3 amino acid sequence is shared among all the VPs and may be referred to as “the common VP3 region.”

In some embodiments, the AAV capsid protein comprises a naturally occurring, or a variant or derivative, or an artificial AAV sequence or a combination thereof.

As described herein, “naturally-occurring” refers to a form of a biomolecule such as an AAV capsid protein or an AAV capsid gene that may be found in nature and/or may be isolated from a biological source such as a human, a non-human primate, or cells isolated from the biological source. Variants, and/or derivatives of a naturally-occurring AAV may also be found in nature and/or isolated from a biological source. An example of a naturally-occurring AAV protein is an AAV capsid protein of AAV2, which is a serotype of AAV first isolated from human cells. An example of a naturally-occurring capsid gene is an AAV2 *cap* gene that encodes the AAV2 capsid protein. A naturally-occurring biomolecule may also be called a “wild-type” biomolecule. Naturally-occurring AAVs include, but are not limited to AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, or AAV13, or a combination thereof.

These AAVs may be classified according to their similarity, for example, according to genetic sequences which indicate a common ancestor and its lineal descendants. AAV clades comprise clades A, B, C, D, E, and F (Gao et al., J Virology, 78(12): 6381-6388 (2004)).

In contrast to naturally-occurring, “engineered” or “artificial” refers to a form of a molecule such as an AAV capsid protein or an AAV capsid gene that has been generated by an experimental method (for example, produced by de novo synthesis) and/or is not known to occur in nature. As an example, an engineered AAV capsid protein may differ from a naturally-occurring AAV capsid protein (or a variant or derivative) by the addition, substitution, and/or deletion of one or more amino acids as compared to the protein sequence of the naturally-occurring capsid protein. An engineered AAV capsid gene may differ from a naturally-occurring AAV capsid gene by the addition, substitution, and/or deletion of one or more nucleotides as compared to the nucleotide sequence of the naturally-occurring capsid gene. Engineered AAV capsid genes and engineered AAV capsid proteins may be created using molecular techniques such as capsid shuffling, directed evolution, random peptide library insertions, generation of chimeric capsids, site-directed mutagenesis, and more. An example of an engineered AAV capsid protein is an AAV capsid protein that comprises the insertion of a heterologous peptide, as described e.g. in US Patent Application No. US2020121746, or the chimera AAV X-Vivo (AAV-XV) derived from chimeras of AAV12 VP1/2 sequences and the VP3 sequence of AAV6 (Viney et al., Journal of Virology 95 (7) (2021)).

In some embodiments, the AAV capsid protein (e.g., VP1, VP2, or VP3) comprises a sequence of an AAV selected from the group consisting of AAV1, AAV2, AAV true type (AAV-TT),

AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof. In one embodiment, the AAV capsid protein is VP1 of an AAV selected from the group consisting of AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof. In one embodiment, the AAV capsid protein is VP2 of an AAV selected from the group consisting of AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof. In one embodiment, the AAV capsid protein is VP3 of an AAV selected from the group consisting of AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof.

In one embodiment, the AAV capsid protein (e.g., VP1, VP2, or VP3) comprises a sequence of an AAV-TT wherein the AAV-TT comprises the sequence SEQ ID NO: 69. In one embodiment, the AAV capsid protein is AAV-TT VP1. In one embodiment, the AAV capsid protein is AAV-TT VP2. In one embodiment, the AAV capsid protein is AAV-TT VP3.

In one embodiment, the AAV capsid protein (e.g., VP1, VP2, or VP3) comprises a sequence of an AAV9. In one embodiment, the AAV capsid protein is AAV9 VP1 and comprises the amino acid sequence given in SEQ ID NO:40. In one embodiment, the AAV capsid protein is AAV9 VP2 and comprises the amino acid sequence given in SEQ ID NO:41. In one embodiment, the AAV capsid protein is AAV9 VP3 and comprises the amino acid sequence given in SEQ ID NO:42. The sequences given in SEQ ID NOS: 40, 41, and 42 are aligned in Figure 2 (top, middle and bottom lines respectively).

Engineered capsid proteins

In one aspect, the present disclosure provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof. In another aspect, the present disclosure provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a knob domain of a bovine ultralong CDR-H3, or a portion thereof. In some embodiments, the antibody fragment does not comprise a stalk domain of a bovine ultralong CDR-H3, or any portion thereof, i.e., the engineered capsid protein does not comprise a stalk domain of a bovine ultralong CDR-H3, or any portion thereof. In some embodiments, the antibody fragment binds to an antigen.

In the context of the present disclosure, the terms “Fused to”, “inserted into”, and “conjugated to” may be used interchangeably. Thus, antibody fragment fusion proteins encompass molecules comprising an antibody fragment of the invention fused to, or conjugated to, or inserted into/within an AAV capsid protein.

- 5 In one embodiment, the antibody fragment is inserted within the AAV capsid protein, directly, i.e. without any linker between the capsid protein and the antibody fragment. In such case, it will be appreciated that an amino acid of the antibody fragment forms a bond (e.g., a peptide bond) with an amino acid of the AAV capsid protein. The antibody fragment and the capsid protein therefore form a fusion protein wherein a first nucleotide sequence encoding the capsid
10 protein is directly genetically fused to a second nucleotide sequence encoding the antibody fragment. Both the capsid protein and the antibody fragment may be translated from the same open reading frame (ORF) or from two ORFs that are aligned.

In another embodiment, the antibody fragment of the invention is inserted within the AAV capsid protein, via a linker, i.e. via one or more linkers.

- 15 In one embodiment, the antibody fragment comprises a sequence selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12, 26, 27, 30, 31, 34 or 35, or any variant thereof as described herein.

- In one embodiment, the antibody fragment is inserted within the AAV capsid protein via one linker (i.e. a single linker). In one embodiment, the antibody fragment is inserted within the
20 AAV capsid protein via one linker wherein the linker is genetically fused to the antibody fragment. In one embodiment, the linker is positioned at the N-terminus of the antibody fragment. In one embodiment, the linker is positioned at the C-terminus of the antibody fragment. In one embodiment, the sequence of an antibody fragment and linker positioned at its C-terminus comprises any sequence from SEQ ID NO: 19 to 21, SEQ ID NO: 23 to 25, SEQ
25 ID NO: 28, SEQ ID NO: 32, or SEQ ID NO: 36.

- In another embodiment, the antibody fragment is inserted within the AAV capsid protein via at least two linkers. The at least two linkers may be the same or different. In some embodiments, at least one linker is fused, optionally genetically, to the N-terminal end of the antibody fragment, and at least one linker is fused, optionally genetically, to the C-terminal end of the
30 antibody fragment. In some embodiments, the antibody fragment is inserted within the AAV capsid protein via two linkers, wherein one linker is positioned at the N-terminal end of the

antibody fragment and the second linker is positioned at the C-terminal end of the antibody fragment, as illustrated below:

--AAV capsid protein – Linker – Antibody fragment – Linker- AAV capsid protein - - - -

In some embodiments, the first and second linkers are the same, i.e. comprise the same amino acid sequence. In one embodiment, the sequence of Linker – Antibody fragment – Linker comprises any sequence from SEQ ID NO: 13 to 18, SEQ ID NO:22, SEQ ID NO: 29, SEQ ID NO: 33, SEQ ID NO: 37.

In some embodiments, the first and second linkers are different, i.e. comprise different amino acid sequences.

10

In one embodiment, the linker(s) is/are a peptide linker. The term “peptide linker” as used herein refers to a peptide comprised of amino acids. A range of suitable peptide linkers will be known to the person of skill in the art. In one embodiment, the linker is a flexible linker. In one embodiment, the linker is selected from a sequence comprised in the list consisting of SEQ ID NO: 103 to SEQ ID NO: 145, as represented in Table 13.

15

Table 13. Linkers

SEQ ID NO:	SEQUENCE
103	(S)GGGGTGGGGS
104	SGGGGSGGGGTGGGGS
105	SGGGGSE
106	DKTHTS
107	(S)GGGGS
108	(S)GGGGG
109	(S)GGGGSGGGGS
110	(S)GGGGSGGGGSGGGGS
111	(S)GGGGSGGGGSGGGGSGGGGS
112	(S)GGGGSGGGGSGGGGSGGGGSGGGGS
113	AAAGSG-GASAS
114	AAAGSG-XGGGS-GASAS
115	AAAGSG-XGGGSXGGGS –GASAS
116	AAAGSG- XGGGSXGGGSXGGGS –GASAS

117	AAAGSG- XGGGSXGGGSXGGGSXGGGS-GASAS
118	AAAGSG-XS-GASAS
119	PGGNRGTTTTRRPATTTGSSPGPTQSHY
120	ATTTGSSPGPT
121	ATTTGS
-	GS
122	EPSGPISTINSPPSKESHKSP
123	GTVAAPSVFIFPPSD
124	GGGGIAPSMVGGGGS
125	GGGGKVEGAGGGGGS
126	GGGGSMKSHDGGGGS
127	GGGGNLITIVGGGGS
128	GGGGVVPSLPGGGGS
129	GGEKSIPGGGGS
130	RPLSYRPPFPFGFPSVRP
131	YPRSIYIRRRHPPSLTT
132	TPSHLSHILPSFGLPTFN
133	RPVSPFTFPRLSNSWLPA
134	SPA AHFPR SIPRPGPIRT
135	APGPSAPSHRSLPSRAFG
136	PRNSIHFLHPLL VAPLGA
137	MPSLSGVLQVRYLSPPDL
138	SPQYPSPLTLTLPPHPSL
139	NPSLNPPSYLHRAPSRIS
140	LPWRTSLLPSLPLRRP
141	PPLFAKGPVGLLSRSFPP
142	VPPAPVVSLRSAHARPPY
143	LRPTPPRVRSYTCCPTP-
144	PNVAHVLPLLTVPWDNLR
145	CNPLLPLCARSPAVRTFP

(S) is optional in sequences 103 and 107 to 112.

In some embodiments, the antibody fragment is inserted within the AAV capsid protein via one linker comprising the sequence SEQ ID NO:95 (GGGGG) (i.e. SEQ ID NO: 108 without the optional Serine). In some embodiments, the antibody fragment is inserted within the AAV capsid protein via one linker comprising the sequence GGGGSGGGGS (SEQ ID NO: 97) (i.e. SEQ ID NO: 109 without the optional Serine). In some embodiments, the linker is positioned at the C-terminal end of the antibody fragment. In some embodiments, the linker comprising SEQ ID NO:95 is encoded by a nucleotide acid sequence comprising ggtggaggcgggggt (SEQ ID NO: 96). In some embodiments, the linker comprising SEQ ID NO:97 is encoded by a nucleotide acid sequence comprising ggt gga ggc ggg agt gga ggt ggc ggg agt (SEQ ID NO: 98). In some embodiments, the linker comprising SEQ ID NO:97 is encoded by a nucleotide acid sequence comprising ggc ggg ggt gga agt ggc gga ggt gga agt (SEQ ID NO: 99).

In some embodiments, the antibody fragment is inserted within the AAV capsid protein via two linkers, wherein one linker is positioned at the N-terminal end of the antibody fragment and the second linker is positioned at the C-terminal end of the antibody fragment, and wherein both linkers comprise the sequence SEQ ID NO: 95. In some embodiments, the antibody fragment is inserted within the AAV capsid protein via two linkers, wherein one linker is positioned at the N-terminal end of the antibody fragment and the second linker is positioned at the C-terminal end of the antibody fragment, and wherein both linkers comprise the sequence SEQ ID NO: 97. In some embodiments, one of the two linkers is encoded by sequence SEQ ID NO: 98 and the other linker is encoded by sequence SEQ ID NO: 98 or SEQ ID NO: 99.

Figure 1 shows an illustration of engineered viruses comprising capsid proteins inserted with an antibody fragment, such as a knob domain (Fig. 1A, 1C), or a bovine ultralong CDR-H3 (Fig. 1B, 1D), optionally via a linker (Fig. 1E-1F respectively).

In some embodiments, the antibody fragment may be conjugated to the capsid protein, e.g. by chemical conjugation. The antibody fragment may be attached via bioconjugation, for example, via covalent bonds formed directly with the capsid protein, or formed indirectly via an adaptor or linker. In some embodiments, the antibody fragment conjugated to the capsid protein forms a cyclic peptide.

Methods for chemical conjugation may comprise direct conjugation of the antibody fragment to lysine residues or cysteine residues on the capsid, for example via NHS-ester chemistry (to primary amines in lysine residues or the N-terminus of polypeptide chains) or maleimide chemistry (to thiol groups in cysteine residues). Further techniques for bioconjugation may

include click chemistry, biotin-based attachment, incorporation of unnatural amino acids on the capsid, polymers, atom-transfer radical polymerization, and more (Chen et al., Wiley Interdiscip Rev Nanomed Nanobiotechnol. 11 (3): e1545 (2019)). A further exemplary method for conjugating a ligand to an AAV capsid is described in WO 2017212019, wherein the ligand is covalently linked to a primary amino group of a capsid polypeptide via a -CSNH- bond. Alternatively, bispecific antibodies may be used to bind an AAV capsid and cell-specific receptor (Bartlett et al., Nat. Biotechnol. 17 (2): 181-186 (1999)). In some embodiments, functionalized peptides can be conjugated to the AAV capsid via an isothiocyanate moiety that readily reacts with a primary amine. For example, an antibody fragment may be linked to an isothiocyanate via a PEG linker, and the isothiocyanate may react with a lysine residue on the AAV capsid.

In some embodiments, amino acid residues in a capsid protein may be mutated to lysine or cysteine residues in order to facilitate the chemical conjugation at locations in the capsid protein which do not have lysine or cysteine residues. For example, if specific amino acid residues in loop regions of the capsid protein are exposed and/or available for binding, these residues may be changed to a lysine or cysteine residue.

In some embodiments, the present invention provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment as described herein, wherein the antibody fragment is inserted within a AAV capsid protein VP1 (optionally via a linker). In some embodiments, the present invention provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment as described herein, wherein the antibody fragment is inserted within a AAV capsid protein VP2 (optionally via a linker). In some embodiments, the present invention provides an engineered capsid protein comprising an AAV capsid protein and an antibody fragment as described herein, wherein the antibody fragment is inserted within a AAV capsid protein VP3 (optionally via a linker).

In certain embodiments, the antibody fragment may be comprised within a region of the AAV capsid protein that mediates binding to heparan sulfate proteoglycan (HSPG), a cell surface molecule that acts as a receptor for AAV (Büning & Srivastava, Molecular Therapy - Methods & Clinical Development 12 (15): 248-265 (2019)).

In some embodiments, the antibody fragment is inserted within a capsid protein without deleting any amino acid of the capsid protein, i.e., all amino acids of the AAV capsid protein before insertion are present in the capsid protein when inserted with the antibody fragment.

Accordingly, the antibody fragment may be inserted before, after an amino acid residue of the AAV capsid protein, or in between two amino acid residues of the AAV capsid protein.

The antibody fragment may be comprised within a structural or topological element in the AAV capsid protein. Structurally, AAV capsid proteins have a conserved core structure, consisting of an eight-stranded, β -barrel motif, α A helix, and loops connecting the β strands. The loops may form features in the surface topology of the AAV capsid, for example, in AAV9 the DE loop (connect β D and β E strands) protrusions form the central channel at each five-fold axis and the HI loops (connect β H and β I strands) lie on the floor of the depression around each five-fold axis (DiMattia et al., *Journal of Virology* 86 (12): 6947-6958 (2012)). Accordingly, the antibody fragment may be comprised within a loop in the AAV capsid protein, such as a surface-exposed loop of the AAV capsid. The antibody fragment may be comprised at any site along the loops and in particular may be comprised within the outermost portion of a loop, as it may maximize exposure to a binding target.

The antibody fragment may be comprised within regions of the AAV capsid protein where amino acids residues show the greatest diversity across AAVs, e.g., in a variable region. The term “variable region” (also “VR”) or “hypervariable region” (also “HVR”) refers to regions on an AAV capsid protein (e.g., VP1, VP2, and/or VP3) which show variation when AAV capsid proteins from different AAVs are compared. In general, the variable regions comprise amino acids that are believed to mediate certain functions of the AAV capsid proteins, such as cell and/or tissue tropism, specificity, and/or efficiency of transduction, recognition of AAV by neutralizing antibodies in the host, the innate immune response of the host, and downstream processing or processability of the AAV (Vandenberghe et al., *Gene Therapy* 16: 311-319 (2009)). For example, VP3 contains nine variable regions in VP3 (numbered VR-I to VR-IX) whose sequence diversity across AAV serotypes has been reported to contribute to functional differences observed between the serotypes (Tseng & Agbandje-McKenna, *Frontiers in Immunology* 5 (9) (2014)).

Variable regions may be found in between β strands in AAV capsid proteins. VR-I may be found in between β B and β C strands; VR-II may be found in between the β D and β E strands; VR-III may be found in between the β E and β F strands; VR-IV, VR-V, VR-VI, VR-VII, and VR-VIII may be found in between the β G and β H strands; and VR-IX may be found after the β I strand. The length of the VRs varies depending on the AAV, but ranges from about 4 to 18 amino acid residues (DiMattia et al., *Journal of Virology* 86 (12): 6947-6958 (2012)) and

(Govindasamy et al., Journal of Virology 80 (23) (2020)). VRs may be found in the loops connecting the β strands. For example, VR-IV, VR-V, and VR-VIII are located in loops at the top of the protrusions (Tseng & Agbandje-McKenna, Frontiers in Immunology 5 (9) (2014)). VR-IV, VR-V, VR-VI, VR-VII, and VR-VIII connecting the β G and β H strands are comprised
5 in a loop referred to as “GH loop”.

In some embodiments, the antibody fragment is inserted within the common VP3 region of the AAV capsid protein. The capsid protein may be VP1, VP2, or VP3. In some embodiments, the antibody fragment is inserted within the common VP3 region of VP1. In some embodiments, the antibody fragment is inserted within the common VP3 region of VP2. In some
10 embodiments, the antibody fragment is inserted within the common VP3 region of VP3.

In some embodiments, the antibody fragment is inserted within the GH loop comprised in the common VP3 region. In one embodiment, the antibody fragment is inserted within the GH loop of VP1. In one embodiment, the antibody fragment is inserted within the GH loop of VP2. In one embodiment, the antibody fragment is inserted within the GH loop of VP3.

15 In some embodiments, the antibody fragment is inserted within the variable region VR-IV of the AAV capsid protein (within the GH loop comprised in the common VP 3 region). In some embodiments, the antibody fragment is inserted within the variable region VR-IV of an AAV9 capsid protein. In some embodiments, the antibody fragment is inserted within the AAV capsid protein after amino acid residue Gly455 of AAV9 with reference to SEQ ID NO: 40 (or a
20 corresponding amino acid residue of another AAV).

In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP1 after amino acid residue Gly455 of AAV9 with reference to SEQ ID NO: 40 (or a corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP2 after amino acid residue Gly318 of
25 AAV9 with reference to SEQ ID NO: 41 (or a corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP3 after amino acid residue Gly253 of AAV9 with reference to SEQ ID NO: 42 (or a corresponding amino acid residue of another AAV).

In some embodiments, the antibody fragment is inserted within the variable region VR-V of the AAV capsid protein. In some embodiments, the antibody fragment is inserted within the
30 variable region VR-V of an AAV9 capsid protein. In some embodiments, the antibody

fragment is inserted within the AAV capsid protein after amino acid residue Asn498 of AAV9 with reference to SEQ ID NO:40, or a corresponding amino acid residue of another AAV.

In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP1 after amino acid residue Asn498 of AAV9 with reference to SEQ ID NO: 40 (or a
5 corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP2 after amino acid residue Asn361 of AAV9 with reference to SEQ ID NO: 41 (or a corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP3 after amino acid residue Asn296 of AAV9 with reference to SEQ ID NO: 42 (or a
10 corresponding amino acid residue of another AAV).

In some embodiments, the antibody fragment is inserted within the variable region VR-VIII of the AAV capsid protein. In some embodiments, the antibody fragment is inserted within the variable region VR-VIII of an AAV9 capsid protein. In some embodiments, the antibody fragment is inserted within the AAV capsid protein after amino acid residue Gln588 of AAV9
15 with reference to SEQ ID NO:40, or a corresponding amino acid residue of another AAV.

In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP1 after amino acid residue Gln588 of AAV9 with reference to SEQ ID NO: 40 (or a corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP2 after amino acid residue Gln451 of
20 AAV9 with reference to SEQ ID NO: 41 (or a corresponding amino acid residue of another AAV). In some embodiments, the antibody fragment is inserted within the AAV capsid protein VP3 after amino acid residue Gln386 of AAV9 with reference to SEQ ID NO: 42 (or a corresponding amino acid residue of another AAV).

In some embodiments, the antibody fragment is inserted within the VP1/VP2 common region
25 of the AAV capsid protein. In some embodiments, the antibody fragment is inserted within the VP1/VP2 common region of VP1. In some embodiments, the antibody fragment is inserted within the VP1/VP2 common region of the VP2 capsid protein.

In some embodiments, the antibody fragment is inserted within the N-terminus of the VP2 capsid protein. In some embodiments, the antibody fragment is inserted within the N-terminus
30 of the VP2 capsid protein of AAV9 before amino acid residue Thr1 with reference to SEQ ID NO:41, or a corresponding amino acid in another AAV. As represented for example in SEQ ID NO: 47, it will be understood that when inserted before the amino acid residue Thr1 of

AAV9 VP2 with reference to SEQ ID NO:41 (amino acid sequence), an additional codon is present at the N-terminal end of the sequence encoding the antibody fragment, said codon being a start codon (e.g. ATG as in SEQ ID NO: 47), to allow the recombinant expression of a VP2 protein wherein the antibody fragment is inserted within (and directly fused to) the N-terminus of a VP2.

In some embodiments, the antibody fragment is inserted within the VP1 unique region of a VP1 capsid protein.

In some embodiments, the capsid protein is from any AAV serotype or subtype of AAV, any forms of AAV such as a naturally-occurring AAV or any variant or any derivative thereof, and/or any engineered (i.e. artificial), or recombinant forms of AAV, or any combination of these. In some embodiments, the capsid protein is VP1 and/or VP2 and/or VP3 from AAV comprising AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination of these.

Among AAVs, the composition and numbering of amino acids in an AAV capsid protein may vary. For example, VP1 from AAV9 may comprise a Gly residue at position 455 of the amino acid sequence given in SEQ ID NO:40. In contrast, VP1 from another AAV may not have a Gly residue at position 455, either because the equivalent Gly residue has a differently numbered position, or because the VP1 in the AAV does not have an equivalent Gly residue. Alignments of capsid proteins from different AAVs may be used to identify corresponding amino acid residues or regions in each variant. For example, sequences of VPs from different AAVs may be aligned using the ClustalW algorithm, or crystal structures of VPs from AAVs may be compared by structural alignment with the Secondary Structure Matching Program (SSM) (DiMattia et al., Journal of Virology 86 (12): 6947-6958 (2012)). Figure 3 shows an alignment of VP1 from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVrh10, and AAV-TT.

As an example, the amino acid residues in VP1 from exemplary AAVs which correspond to Gly455, Asn498, and Gln588 in VP1 from AAV9 (amino acid sequence given in SEQ ID NO: 40) are as follows: Gly454, Asn498, Ser588 of the amino acid sequence given in SEQ ID NO:60 (AAV1); Gly454, Asn498, Ser588 of the amino acid sequence given in SEQ ID NO:65 (AAV6); Gly456, Asn500, Asn590 of the amino acid sequence given in SEQ ID NO:67 (AAV8); Gly453, Asn497, Asn587 of the amino acid sequence given in SEQ ID NO:69 (AAV AAV-TT).

Where an amino acid residue in an AAV capsid protein from a first AAV does not have a corresponding amino acid residue in an AAV capsid protein from a second AAV, a modification according to the present disclosure may be made at or near to an adjacent or nearby amino acid residue. The alternative amino acid in the second AAV may have a comparable location, function, and/or topology to the amino acid residue in the first AAV. For example, where an insertion is disclosed herein at Gly455 (according to VP1 numbering) in an AAV capsid protein from AAV9, the insertion may be made at a Gly residue located in the VR-IV of the GH loop of another AAV, or at an alternative amino acid residue in the VR-IV. The alternative amino acid may have a comparable location, function, and/or topology to Gly455 in AAV9.

In one embodiment, the invention provides an engineered capsid protein comprising an AAV capsid protein and at least two antibody fragments as described in the present disclosure, i.e. the capsid protein (e.g. VP1, VP2, or VP3) is inserted with more than one (two or more) antibody fragments. In some embodiments, the engineered capsid protein is an engineered VP1 comprising an AAV capsid protein VP1 and at least two antibody fragments as described in the present disclosure. In some embodiments, the engineered capsid protein is an engineered VP2 comprising an AAV capsid protein VP2 and at least two antibody fragments as described in the present disclosure. In some embodiments, the engineered capsid protein is an engineered VP3 comprising an AAV capsid protein VP3 and at least two antibody fragments as described in the present disclosure. In some embodiments, the engineered capsid protein comprises an AAV capsid protein and two, three, four, five and up to ten antibody fragments. In some embodiments, the antibody fragments are inserted in more than one site in the AAV capsid protein. In some embodiments, each of the antibody fragments is inserted at a different site in the AAV capsid protein. In some embodiments, two, or three, or four, or five antibody fragments are inserted within the capsid protein, at different sites of insertion, as described above. In some embodiments, the engineered capsid protein is VP1, VP2 or VP3.

In some embodiments, the at least two antibody fragments are identical. In some embodiments, the at least two antibody fragments are different. In some embodiments, at least three antibody fragments are inserted and the at least three antibody fragments are identical, different or are a mixture of antibody fragments that are identical and antibody fragments that are different.

AAV capsids comprising engineered capsid proteins

In one aspect, there is provided an AAV capsid comprising an engineered capsid protein according to the present disclosure. Therefore, in one aspect, the invention provides, an AAV capsid comprising an engineered capsid protein, wherein the engineered capsid protein
5 comprises an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof, optionally wherein the antibody fragment binds an antigen. In another aspect, the invention provides, an AAV capsid comprising an engineered capsid protein, wherein the engineered capsid protein comprises an
10 AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a knob domain of a bovine ultralong CDR-H3, or a portion thereof, optionally wherein the antibody fragment binds an antigen.

In some embodiments, the antibody fragment does not comprise a stalk of an ultralong CDR-H3. In some embodiments, the AAV capsid protein does not comprise a stalk of an ultralong CDR-H3. In some embodiments, the AAV capsid does not comprise a stalk of an ultralong
15 CDR-H3. In some embodiments, the antibody fragment does not comprise a stalk of a bovine ultralong CDR-H3. In some embodiments, the AAV capsid protein does not comprise a stalk of a bovine ultralong CDR-H3. In some embodiments, the AAV capsid does not comprise a stalk of a bovine ultralong CDR-H3.

In an AAV capsid, the three VPs assemble to form a 60-subunit icosahedral-shaped capsid.
20 The VPs are present in an average ratio of about 5:5:50 (VP1:VP2:VP3), although the assembly process is stochastic and the actual ratio may vary slightly (Wörner et al., Nat Commun 12 (1): 1642 (2021)). As described herein, the number of antibody fragments comprised within an AAV capsid may depend on which VP is inserted with the antibody fragment and on the cloning strategy, as described in the Examples, in particular Example 5.

25 Figure 10 illustrates possible configurations of capsid proteins. Knob domains are described for illustrative purposes, as an example of the antibody fragments of the present disclosure.

In some embodiments, only AAV capsid VP1 is inserted with an antibody fragment, i.e., the AAV capsid comprises an engineered capsid protein VP1 according to the present disclosure, while VP2 and VP3 have not been engineered, i.e. do not comprise an antibody fragment. It
30 will be understood that such an AAV capsid comprises about 5 copies of an engineered capsid protein according to the present disclosure (i.e. 5 copies of an engineered capsid protein VP1; in addition to 5 copies of VP2 and 50 copies of VP3).

Figure 10A illustrates the configuration of a capsid comprising 5 antibody fragments, such as knob domains, wherein the antibody fragments are comprised within only VP1 of the AAV capsid. Methods for producing such a capsid may comprise providing two different *cap* genes: 1) a first *cap* gene deleted in a start codon for VP1 (represented in Figure 10A with an asterisk over the start codon, in black), so that no VP1 is produced, while VP2 and VP3 are produced from this first *cap* gene); and 2) a second *cap* nucleic sequence encoding only VP1, comprising the antibody fragment (represented in Figure 10A as a diagonal pattern), and co-expressing both *cap* genes. When the capsid assembles, only VP1 comprises the antibody fragment and about 5 copies of the antibody fragment are comprised within the engineered capsid.

10 In some embodiments, the first *cap* gene deleted in a start codon for VP1 is from AAV9 and comprises SEQ ID NO: 73 (AAV9 VP1 del"). The second *cap* nucleic sequence encoding VP1 only may be a *cap* gene where VP1 has a deletion in the splice acceptor site for VP2 and VP3 (e.g. "VP1 sa del", SEQ ID NO:70). Therefore, the second *cap* nucleic sequence may comprise the sequence of a *cap* gene deleted in the splice acceptor site for VP2 and VP3 and the sequence of an antibody fragment inserted therein. An example of such a sequence of the invention is provided in SEQ ID NO: 74. It will be understood that the antibody fragment within that sequence may be replaced by any other sequence coding for an antibody fragment according to the invention, and that the antibody fragment may be inserted in another insertion site as described in the present disclosure. In some embodiments, the antibody fragment is inserted within the common VP3 region of VP1 as described herein. In some embodiments, the antibody fragment is inserted within the VP1/VP2 common region. In some embodiments, the antibody fragment is inserted within the VP1 unique region.

In some embodiments, only AAV capsid VP2 is inserted with an antibody fragment, i.e., the AAV capsid comprises an engineered capsid protein VP2 according to the present disclosure, while VP1 and VP3 have not been engineered, i.e. do not comprise an antibody fragment. It will be understood that such an AAV capsid comprises about 5 copies of an engineered capsid protein according to the present disclosure (i.e. 5 copies of an engineered capsid protein VP2; in addition to 5 copies of VP1 and 50 copies of VP3).

Figure 10B illustrates the configuration of a capsid comprising 5 antibody fragments, such as knob domains, wherein the antibody fragments are comprised within only VP2 of the AAV capsid. Methods for producing such a capsid may comprise providing two different *cap* genes: 1) a first *cap* gene deleted in a start codon for VP2 (represented in Figure 10B with an asterisk

over the start codon, in black), so that no VP2 is produced, while VP1 and VP3 are produced from this first *cap* gene; and 2) a second *cap* genetic sequence encoding VP2 only comprising the antibody fragment (represented in Figure 10B as a diagonal pattern), and co-expressing both *cap* genes. When the capsid assembles, only VP2 comprises the antibody fragment and about 5 copies of the antibody fragment are comprised within the engineered capsid.

In some embodiments, the first *cap* gene deleted in a start codon for VP2 is from AAV9 and comprises SEQ ID NO: 39 (AAV9 del VP2"). The second *cap* genetic sequence encoding VP2 comprising the antibody fragment may be generated by inserting antibody fragments within VP2 as described in the present disclosure.

For example, the second *cap* genetic sequence encoding VP2 comprising the antibody fragment inserted within the N-terminus of AAV9 VP2 may comprise SEQ ID NO: 72. As another example, the second *cap* genetic sequence encoding VP2 comprising the antibody fragment inserted within the common VP3 region of AAV9 VP2 may comprise SEQ ID NO: 71. It will be understood that the antibody fragment within those sequences may be replaced by any other sequence coding for an antibody fragment according to the invention, and that the antibody fragment may be inserted in another insertion site as described in the present disclosure. In some embodiments, the antibody fragment is inserted within the common VP3 region of VP2 as described herein. In some embodiments, the antibody fragment is inserted within the VP1/VP2 common region. In some embodiments, the antibody fragment is inserted within the N-terminus of VP2.

In some embodiments, only AAV capsid VP3 is inserted with an antibody fragment, i.e., the AAV capsid comprises an engineered capsid protein VP3 according to the present disclosure, while VP1 and VP2 have not been engineered, i.e. do not comprise an antibody fragment. It will be understood that such an AAV capsid comprises about 50 copies of an engineered capsid protein according to the present disclosure (i.e. 50 copies of an engineered capsid protein VP3; in addition to 5 copies of VP1 and 5 copies of VP2). Figure 10C illustrates the configuration of a capsid comprising 50 antibody fragments, such as knob domains, wherein the antibody fragments are comprised within only VP3 of the AAV capsid. Methods for producing such a capsid may comprise providing different *cap* genes: 1) a first *cap* gene comprising a deletion in a splice acceptor for VP2 and VP3 in the *cap* gene (indicated in Figure 10C with an asterisk over the splice acceptor, in white), so that neither VP2 nor VP3 is produced, while VP1 is produced from this *cap* gene, and 2) additional *cap* genetic sequences encoding VP2 and VP3

(either on the same plasmid or on separate plasmids) comprising the antibody fragment (represented in Figure 10C as a diagonal pattern), and co-expressing all *cap* genes.

In some embodiments, the AAV capsid comprises more than one engineered capsid protein according to the present disclosure.

- 5 In some embodiments, the AAV capsid comprises two engineered capsid proteins according to the present disclosure. In some embodiments, the first and second engineered capsid proteins are respectively VP1 and VP2, or VP1 and VP3, or VP2 and VP3.

In some embodiments, the first and second engineered capsid proteins are respectively VP1 and VP2. In such embodiments, the AAV capsid may comprise about 10 copies of an antibody
10 fragment. Figure 10D illustrates the configuration of a capsid comprising 10 antibody fragments, such as knob domains, wherein about 5 antibody fragments are comprised within VP1 and about 5 antibody fragments are comprised within VP2 of the AAV capsid. Methods for producing such a capsid may comprise providing different *cap* genes: 1) a first *cap* gene comprising a deletion in a start codon for VP1 and VP2 in the *cap* gene (indicated in Figure
15 10D with an asterisk over the start codon, in black), so that neither VP1 nor VP2 is produced, while VP3 is produced from this *cap* gene, and 2) additional *cap* genetic sequences encoding VP1 and VP2 (either on the same plasmid or on separate plasmids) comprising the antibody fragment (represented in Figure 10D as a diagonal pattern), and co-expressing all *cap* genes.

In some embodiments, the first and second engineered capsid proteins are respectively VP1
20 and VP3. In such embodiments, the AAV capsid may comprise about 55 copies of an antibody fragment. Figure 10E illustrates the configuration of a capsid comprising 55 antibody fragments, such as knob domains, wherein about 5 antibody fragments are comprised within VP1 and about 50 antibody fragments are comprised within VP3 of the AAV capsid. Methods for producing such a capsid may comprise providing two *cap* genes: 1) a first *cap* gene
25 comprising a deletion in a start codon for VP2 in the *cap* gene (indicated in Figure 10E with an asterisk over the start codon, in black), so that VP2 is not produced, and inserting an antibody fragment in the *cap* gene so that engineered VP1 and VP3 are produced from this *cap* gene, and 2) additional *cap* gene encoding VP2 (represented in Figure 10E as a diagonal pattern), and co-expressing both *cap* genes.

- 30 In some embodiments, the first and second engineered capsid proteins are respectively VP2 and VP3. In such embodiments, the AAV capsid may comprise about 55 copies of an antibody fragment. Figure 10F illustrates the configuration of a capsid comprising 55 antibody

fragments, such as knob domains, wherein about 5 antibody fragments are comprised within VP2 and about 50 antibody fragments are comprised within VP3 of the AAV capsid. Methods for producing such a capsid may comprise providing different *cap* genes: 1) a first *cap* gene comprising a deletion in a start codon for VP1 and VP2 in the *cap* gene (indicated in Figure 10F with an asterisk over the start codon, in black), so that neither VP1 nor VP2 are produced, and inserting an antibody fragment in the *cap* gene so that engineered VP3 is produced from this *cap* gene, and 2) additional *cap* gene encoding VP1 and VP2 (either on the same plasmid or on separate plasmids), VP1 that does not contain an antibody fragment and VP2 inserted with an antibody fragment (represented in Figure 10F as a diagonal pattern), and co-expressing all *cap* genes.

Alternative methods for producing such a capsid may comprise providing: 1) a first *cap* gene comprising a deletion in a splice acceptor for VP2/VP3 in the *cap* gene (indicated in Figure 10H with an asterisk over the splice acceptor, in white, e.g. “VP1 sa del” as described herein), so that only VP1 is produced from this *cap* gene, and 2) additional *cap* gene encoding VP2 and VP3 (either on the same plasmid or on separate plasmids), both containing an antibody fragment (e.g. a knob domain as represented in Figure 10H as a diagonal pattern), and co-expressing all *cap* genes.

An example of this capsid may be produced using a first nucleotide sequence comprising “AAV9 VP1 sa del” (SEQ ID NO:70) encoding a first AAV9 *cap* gene in which a VP2/VP3 splice acceptor site is deleted, and a second nucleotide sequence comprising a second AAV9 *cap* gene in which an antibody fragment is comprised and in which the VP1 start codon is deleted. The second nucleotide may comprise SEQ ID NO:75. It will be understood that nucleotide sequences encoding other antibody fragments as described in the present disclosure may be substituted for the nucleotide sequence encoding the antibody fragment used as an example in SEQ ID NO:75, and that the antibody fragment may be inserted in another insertion site as described in the present disclosure. e.g. at different positions within the common VP3 region.

In some embodiments, the AAV capsid comprises three engineered capsid proteins according to the present disclosure. In some embodiments, the first, second and third engineered capsid proteins are respectively VP1, VP2 and VP3.

Figure 10G illustrates the configuration of a capsid comprising 60 antibody fragments (such as knob domains, wherein the antibody fragments are comprised within each of VP1, VP2, and

VP3 of the AAV capsid. Methods for producing such a capsid may comprise a step of providing a *cap* gene comprising a nucleic sequence comprising a sequence encoding an antibody fragment, such as a knob domain (represented as in Figure 10G as a diagonal pattern), so that each of VP1, VP2, and VP3 comprising an antibody fragment are produced.

5 An exemplary AAV capsid comprising about 60 copies of an antibody fragment may be produced using nucleotide sequences encompassed by the invention such as AAV9 Gly455 B1 (SEQ ID NO:43), AAV9 Asn498 B1 (SEQ ID NO:44), or AAV9 Gln588 B1 (SEQ ID NO:45). It will be understood that other nucleotide sequences encoding other antibody fragments as described in the present disclosure may be used, and that the antibody fragment may be inserted
10 in another insertion site as described in the present disclosure, to generate engineered capsids according to the invention. In some embodiments, the antibody fragment is inserted within the common VP3 region of VP1, VP2 and VP3 as described herein. It will be understood that the number of antibody fragments comprised within the AAV capsid may reflect not only whether the antibody fragment is comprised within VP1, VP2, and/or VP3, but also the number of
15 antibody fragments comprised within the VP. For example, 2 antibody fragments may be comprised within VP1, but not within VP2 or VP3, with the resulting capsid comprising 10 antibody fragments in total, all comprised within VP1.

Accordingly, a plurality of antibody fragments may be comprised within VP1, but not VP2 or VP3. In certain embodiments, a plurality of antibody fragments may be comprised only within
20 VP2 but not within VP1 or VP3, or comprised only within VP3 but not within VP1 or VP2, or comprised within VP1 and VP2 but not VP3, or comprised within VP1 and VP3 but not VP2, or comprised within VP2 and VP3 but not VP1, or comprised within each of VP1, VP2, and VP3. A plurality of antibody fragments may be comprised within the same site in an AAV capsid protein, or the plurality of antibody fragments may be comprised within different sites
25 in the AAV capsid protein.

In certain embodiments, each of the antibody fragments in the plurality of antibody fragments is the same antibody fragment. In some embodiments, at least 2 of the antibody fragments in the plurality of antibody fragments are different, e.g., comprising a different sequence and/or different binding properties. An AAV capsid may comprise 2 or more different antibody
30 fragments. It will be understood that in such embodiments wherein the antibody fragments bind to different antigens or targets, there is provided multi-specific engineered capsids, capable of binding more than one antigen or target. In some embodiments, the invention provides a AAV

capsid which comprise two antibody fragments which bind to a different target or antigen of interest (i.e., there is provided bi-specific engineered capsids). It will be understood that in such embodiments, each antibody fragment may be a bovine ultralong CDR-H3, or a portion thereof which binds to an antigen, a knob domain of a bovine ultralong CDR-H3, or a portion thereof which binds to an antigen, and that the number of copies of each antibody fragment and insertion site may differ and the AAV capsid may be generated according to the present disclosure.

Figure 11 shows an illustration of antibody fragments inserted within AAV capsids, wherein two exemplary knob domains have different sequences. The total number of knob domains may be varied.

In some embodiments, 2 different antibody fragments are comprised within VP1 and VP2 respectively but not VP3, or comprised within VP1 and VP3 but not VP2, or comprised within VP2 and VP3 but not VP1.

In some embodiments, a first antibody fragment is comprised within VP1 and a second antibody fragment is comprised within VP2. The resulting AAV capsid may comprise 5 copies of the first antibody fragment and 5 copies of the second antibody fragment, such as knob domains (as illustrated in Fig. 11A). Methods for producing this capsid may comprise a step of providing 1) a *cap* gene deleted in start codons for VP1 and VP2 (indicated in Figure 11A with an asterisk over the start codons, in black), so that no VP1 or VP2 are produced from the *cap* gene, while VP3 is produced 2) a sequence encoding VP1 comprising a first antibody fragment (represented in Figure 11A as a diagonal pattern) and a second sequence encoding VP2 comprising a second antibody fragment (represented in Figure 11A as a stippled pattern) on the same plasmid or on separate plasmids.

In some embodiments, a first antibody fragment is comprised within VP2 and a second antibody fragment is comprised within VP3. The resulting AAV capsid may comprise 5 copies of the first antibody fragment and 50 copies of the second antibody fragment, such as knob domains (as illustrated in Fig. 11B). Methods for producing this capsid may comprise a step of providing 1) a *cap* gene deleted in start codons for VP1 and VP2 (indicated in Figure 11B with an asterisk over the start codons, in black) and inserted with a first antibody fragment, so that no VP1 or VP2 are produced from the *cap* gene, while an engineered VP3 is produced 2) a sequence encoding VP1 and a second sequence encoding VP2 comprising a second antibody

fragment (represented in Figure 11B as a diagonal pattern) on the same plasmid or on separate plasmids.

In some embodiments, a first antibody fragment is comprised within VP1 and a second antibody fragment is comprised within VP3. The resulting AAV capsid may comprise 5 copies of the first antibody fragment and 50 copies of the second antibody fragment, such as knob domains (as illustrated in Fig. 11C). Methods for producing this capsid may comprise a step of providing 1) a *cap* gene deleted in the splice acceptor for VP2 and VP3 (indicated in Figure 11C with an asterisk over the splice acceptor, in white) and inserted with a first antibody fragment, so that no VP2 or VP3 are produced from the *cap* gene, while an engineered VP1 is produced, 2) a sequence encoding VP2 and a second sequence encoding VP1 comprising a second antibody fragment (represented in Figure 11C as a stippled pattern) on the same plasmid or on separate plasmids.

Alternative methods for producing such a capsid may comprise providing: 1) a first *cap* gene comprising a deletion in the start codon for VP1 and VP2 (indicated in Figure 11D with an asterisk over the start codon, in black), and an antibody fragment inserted within VP3 so that only engineered VP3 is produced from this *cap* gene, and 2) a sequence encoding VP2, and a sequence encoding VP1 comprising an antibody fragment, (either on the same plasmid or on separate plasmids), and co-expressing all genes. When the capsid assembles, it comprises 50 copies of the first antibody fragment and 5 copies of the second antibody fragment.

The AAV capsid may comprise AAV capsid proteins from one or more AAVs. In some embodiments, the AAV capsid may comprise capsid proteins from AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof. VP1 and/or VP2 and/or VP3 may each be selected from AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof. VP1, VP2, and VP3 may be selected from the same AAV, or may be selected from different AAVs. In some embodiments, VP1 comprises a sequence from a first AAV and is inserted within an antibody fragment, while VP2 and VP3 do not comprise an antibody fragment, and may be from another AAV. In other embodiments, each of VP1, VP2, and VP3 may comprise an antibody fragment, and VP1 is from a first AAV, VP2 is from a second AAV, and VP3 is from a third AAV.

Standard methods may be used for analyzing the configuration of the capsids such as x-ray crystallography, NMR spectroscopy, or cryogenic electron microscopy (cryo EM). Such methods may be useful in the context of the disclosure to assess the assembly and three-dimensional structure of the capsid and the antibody fragment, which may be visualized with high-resolution using advanced software.

Nucleic acids, vectors, rAAV's particles and host cells

In one aspect, the invention provides a nucleic acid encoding an engineered capsid protein or a capsid according to the present invention.

In some embodiments, the nucleic acid encoding an engineered capsid protein or a capsid comprises a first nucleotide sequence encoding a capsid protein and a second nucleotide sequence, wherein the second nucleotide sequence encodes an antibody fragment. The first nucleotide sequence may be contiguous with the second nucleotide sequence, e.g., attached via linkage by phosphodiester bonds. In some embodiments, the nucleic acid comprises at least a third nucleotide sequence encoding a linker, wherein at least one linker is positioned between a capsid protein and an antibody fragment as disclosed in the present disclosure.

In some embodiments, the first nucleotide sequence and the second nucleotide sequence and optionally at least the third nucleotide sequence, are in reading frame with each other. This configuration may facilitate transcription and translation of a gene product (e.g., a protein) comprising the capsid protein and the antibody fragment.

In some embodiments, the first nucleotide sequence comprises an AAV *cap* gene that encodes at least one AAV capsid protein selected from VP1, VP2, and VP3.

The first nucleotide sequence (e.g., the AAV *cap* gene) may comprise a mutation in any one of VP1, VP2, and/or VP3, as compared to the nucleotide sequence of a reference AAV *cap* gene. For example, the start codon of the nucleotide encoding any one of VP1, VP2, or VP3 may comprise a mutation. In some embodiments, the start codon from which synthesis of VP1 starts comprises a mutation that eliminates translation of VP1. In some embodiments, the start codon from which synthesis of VP2 starts comprises a mutation that eliminates translation of VP2. In some embodiments, the start codon from which synthesis of VP3 starts comprises a mutation that eliminates translation of VP3. An exemplary nucleotide sequence is given in SEQ ID NO:39, in which the nucleotide sequence encoding VP2 of AAV9 comprises a mutation in the

start codon. A mutation may comprise a deletion in a start codon or a deletion in a splice acceptor site for any one of VP1, VP2, or VP3.

In some embodiments, the first nucleotide sequence encodes at least one AAV capsid protein comprising VP1, VP2, or VP3, wherein each of VP1 and/or VP2 and/or VP3 are from AAV
5 comprising AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof.

In some embodiments, the nucleic acid encoding an engineered capsid protein or a capsid comprises a nucleotide sequence which encodes an antibody fragment as described in the present disclosure. Examples of suitable sequences are provided in Table 1 (SEQ ID NO: 1, 3,
10 5, 7, 9, 11). In some embodiments, the nucleic acid encoding an engineered capsid protein or a capsid comprises a nucleotide sequence which encodes an antibody fragment which comprises the sequence SEQ ID NO:2, 4, 6, 8, 10, 12, 26, 27, 30, 31, 34 or 35, or any variant thereof as described herein.

A nucleotide sequence for the AAV9 Rep/Cap is given in SEQ ID NO:38. The VP2 start codon
15 beginning at nucleotide 2424 is indicated in bold, underlined text.

A nucleotide sequence for the AAV9 Rep/Cap comprising a mutation in the VP2 start codon is given in SEQ ID NO: 39. The VP2 start codon beginning at nucleotide 2424 is indicated in underlined text. In this sequence, a G to C mutation removed the start codon in VP2.

Amino acid sequences for AAV9 VP1 (SEQ ID NO: 40), AAV9 VP2 (SEQ ID NO: 41), and
20 AAV9 VP3 (SEQ ID NO: 42) are provided. In each sequence, a Gly (G), an Asn (N), and a Gln (Q) residue are indicated in bold, underlined text. According to the numbering in SEQ ID NO: 40, the residues are Gly455, Asn498, and Gln588. According to the numbering in SEQ ID NO: 41, the residues are Gly318, Asn361, and Gln451. According to the numbering in SEQ ID NO:42, the residues are Gly253, Asn296, and Gln386. Corresponding acid nucleic sequences
25 are provided in SEQ ID NO: 48, 49, 50 respectively.

Nucleotide sequences comprising a sequence encoding a capsid protein inserted with an antibody fragment include sequences presented in Table 2. In some embodiments, the nucleic acid encoding an engineered capsid protein or a capsid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 43-59 and SEQ ID NO: 70-94.

30 In another aspect, the invention provides a vector comprising the nucleic acid encoding an engineered capsid protein according to the present invention. The vector may be a plasmid or

a viral vector (i.e. an AAV vector). Therefore, there is provided an AAV vector comprising a nucleic acid encoding an engineered capsid protein, wherein the engineered capsid protein comprises an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof. There is also provided an AAV
5 vector comprising a nucleic acid encoding an engineered capsid protein, wherein the engineered capsid protein comprises an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a knob domain of a bovine ultralong CDR-H3, or a portion thereof. In some embodiments, the antibody fragment binds to an antigen.

Another aspect of the present disclosure relates to a recombinant AAV (rAAV) particle
10 comprising an engineered capsid protein and/or an engineered capsid and/or a nucleic acid and/or a vector, as described herein. Therefore, there is provided a recombinant AAV (rAAV) particle comprising an engineered capsid protein or an engineered capsid comprising an engineered capsid protein, wherein the engineered capsid protein comprises an AAV capsid protein and an antibody fragment which is a bovine ultralong CDR-H3 or a portion thereof.
15 There is also provided a rAAV particle comprising an engineered capsid protein or an engineered capsid comprising an engineered capsid protein, wherein the engineered capsid protein comprises an AAV capsid protein and an antibody fragment which is a knob domain of a bovine ultralong CDR-H3 or a portion thereof. In some embodiments, the engineered capsid protein, or engineered capsid, or rAAV,particle does not comprise a stalk of a bovine
20 ultralong CDR-H3. In some embodiments, the antibody fragment binds to an antigen.

In some embodiments, the rAAV particle comprises a nucleic acid sequence that is not of AAV origin (e.g., a nucleic acid heterologous to AAV). The heterologous nucleic acid may be a nucleic acid sequence of interest. In some embodiments, the heterologous nucleic acid is a
25 single stranded DNA. In some embodiments, the heterologous nucleic acid is a transgene. In some embodiments, the recombinant AAV (rAAV) particle comprises a transgene. In some embodiments, the transgene is encapsulated within the rAAV particle. In some embodiments, the transgene is comprised in a rAAV vector inside the rAAV particle. It will be understood that the rAAV may provide a vector to deliver a transgene into a host cell, wherein the transgene
30 is delivered and expressed in the host cell.

In some embodiments, the transgene encodes a peptide, a polypeptide or a nucleic acid molecule. In some embodiments, the nucleic acid molecule is a small interfering RNA (siRNA), small or short hairpin RNA (shRNA), microRNA (miRNA).

The nucleic acid constructs, AAV vectors and viral particles described herein may be prepared by standard means known in the art. Thus, well established methods of production (including transfection, packaging and purification methods) can be used to prepare a suitable vector preparation.

- 5 In certain embodiments, the rAAV vectors and particles further comprise at least one genetic element, such as an additional nucleic acid encoding a promoter, an intron, an inverted terminal repeats (ITR), a poly A, or a stuffer sequence. A nucleic acid construct of the invention and further comprising at least one ITR flanking said nucleic acid construct in 5' and/or 3', preferably at least two ITRs (one at each end of the nucleic acid construct of the invention, i.e. a 5' ITR and a 3' ITR). An ITR sequence acts in *cis* to provide a functional origin of replication and allows for integration and excision of the nucleic acids construct from the genome of a cell. One or more of the ITRs may be obtained from viral genomes, such as AAV genomes, having different serotypes or may be a chimeric or mutant ITR. An example of a mutant ITR is one having a deletion of a *trs* (terminal resolution site). This deletion allows for continued replication of the genome to generate a single-stranded genome which contains both coding and complementary sequences i.e., a self-complementary viral genome (e.g. self-complementary AAV genome). This allows for bypass of DNA replication in the target cell, and so enables accelerated transgene expression. Preferred ITR sequences are those of from AAV genomes, such as AAV2, AAV9 and variants thereof.
- 15
- 20 In another aspect, the present disclosure provides a host cell comprising a nucleic acid construct of the invention or a vector of the invention.

- The present invention also provides a host cell which produces an AAV viral particle as disclosed herein. Any suitable host cell may comprise a nucleic acid construct of the invention or a vector of the invention. Further, any suitable host cell can be used to produce a viral particle of the invention. In the context of the present invention as a whole, the host cell is preferably an insect cell or a mammalian cell. Non-limiting examples of such cells are sf9, HEK293 (including e.g., HEK293F, HEK293S or HEK293T), BHK or CHO cells. In a non-limiting example, the host cell may comprise a nucleic acid construct or a vector of the invention, and further comprise an additional nucleic acid construct or a vector providing the minimal additional genome sequences needed for packaging of the nucleic acid construct in the viral particle (such as in the form of an AAV helper plasmid providing other essential genes).
- 25
- 30

In one aspect, the disclosure provides a method for producing an AAV particle comprising an engineered capsid protein and a transgene, said method comprising:

- a) providing a first vector comprising a nucleotide sequence encoding an AAV capsid protein, a second nucleotide sequence encoding the antibody fragment; wherein the first nucleotide sequence and the second nucleotide sequence are genetically fused optionally via a nucleotide sequence coding for a linker;
- b) providing a second vector comprising the transgene
- c) providing a third, Helper vector
- d) transfecting a host cell with the first, second and third vector;
- e) recovering the AAV particle from the host cell.

Optionally, the method comprises providing at least an additional vector encoding an AAV capsid. It will be understood that the number of nucleotide sequences will vary depending on the cloning strategy as described in the present disclosure. Different cloning strategies are described herein as examples and may be used to control which of VP1, VP2, and/or VP3 comprise the antibody fragment in the rAAV. For example, when the antibody fragment is comprised within the common VP3 region, each of VP1, VP2, and VP3 as translated from the AAV *cap* gene will comprise the antibody fragment.

In contrast, if the AAV *cap* gene comprises a mutation resulting in that VP1 and/or VP2 and/or VP3 is not produced then, providing an additional nucleotide sequence (either in a same or different vector) encoding said VP(s) is necessary to produce all VPs (VP1, VP2 and VP3) in the rAAV. Therefore, the cloning/ transfection strategy may require transfection of multiple nucleotide sequences encoding all the AAV capsid proteins (engineered, i.e. comprising an antibody fragment, or non-engineered). Also, the transfection of 2, 3, 4, or 5 nucleotide sequences may be used, for example, to maintain a high production efficiency.

Accordingly, the method may comprise providing n nucleotide sequences, each encoding an AAV capsid protein (engineered or not), wherein n is an integer, for example 1, 2, 3, 4, 5, 6, 7, 8, 9, 10; and transfecting the cell line with the n nucleotide sequences. When n is 2 or higher, the nucleotide sequences may be provided together, e.g., combined into one nucleotide sequence or together on one plasmid. Alternatively, the 2 or more nucleotide sequences may be provided separately from one another, e.g., on separate plasmids.

Amino acid sequences and nucleic acid sequences coding for antibody fragments for use in the present invention can be obtained by methods previously described, for example in WO2021191424. The method may comprise the following steps:

- a) immunising a bovine with an immunogenic composition, and;
- 5 b) isolating total RNA from PBMC or secondary lymphoid organ, or antigen-specific memory B-cells, and;
- c) amplifying the cDNA of the ultralong CDR-H3, and;
- d) sequencing an ultralong CDR-H3 or portion thereof;

wherein the immunogenic composition comprises an antigen of interest or immunogenic portions thereof, or DNA encoding the same.

Methods for immunizing a bovine according to step a) and isolating total RNA according to step b) are well known in the art.

Methods for amplifying the cDNA of the ultralong CDR-H3 are also well known in the art. Advantageously, at step c) a method for amplifying directly the cDNA of ultralong CDR-H3 and discriminate from standard CDR-H3 may be used. The method may comprise a primary polymerase chain reaction (PCR) with primers flanking CDR-H3, annealing to the conserved framework 3 and framework 4 of the VH, to amplify all CDR-H3 sequences, irrespective of their length or amino acid sequence. The method may additionally comprise a second round of PCR with stalk primers to specifically amplify ultralong sequences from the primary PCR.

In one embodiment, the method for amplifying the cDNA of CDR-H3 comprises:

- 1) a primary PCR using primers flanking CDR-H3, annealing to the conserved framework 3 and framework 4 of the VH, to amplify all CDR-H3 sequences, and
 - 2) a second round of PCR using stalk primers to specifically amplify ultralong sequences from the primary PCR.
- Step d) may be performed according to methods well known in the art such as direct nucleotide sequencing. The antibody fragment and portion thereof, for example the knob domain, may be defined as described in the present disclosure and its sequence isolated.

Additionally, sequences coding antibody fragments for use in the present invention may be derived from libraries, as described for example in WO2021191424.

Following assembly of a rAAV comprising a capsid protein that comprises an antibody fragment, the rAAV may be screened for functional properties and/or for manufacturability (e.g., viral production). For example, the rAAV may be screened to determine if the antibody fragment retains its binding affinity for its target, or to an antigen of interest, if the capsid protein has assembled, titers of rAAV produced, and/or transduction efficiency. In some
5 embodiments, functional properties and/or manufacturability may be determined for a plurality of rAAVs, wherein each rAAV comprises a capsid protein that comprises an antibody fragment or combination of antibody fragments, so that an rAAV with desired properties may be selected.

10 In certain embodiments, screening may be performed by introducing antibody fragments into one or more sites on a rAAV. Diverse antibody fragments may be inserted into the site(s) on rAAV to generate an rAAV library, for example, by inserting nucleic acids, each encoding a different antibody fragment into an AAV cap gene and generating a library plasmid pool. The insertion may be at specific sites in the *cap* gene, or may be at random sites in the *cap* gene. In
15 some embodiments, all 60 subunits of the capsid gene express an antibody fragment, for example, each of the 60 subunits may be engineered to express a copy of the same antibody fragment. In some embodiments, fewer than 60 subunits of the capsid gene express an antibody fragment. The library plasmid pool may be used to transfect a producer cell line, such as HEK293 cells, in order to produce the rAAV display library. The rAAV library may be
20 screened for binding to a target and/or transduction efficiency of target cells in order to select a rAAV with desired properties resulting from the insertion of antibody fragment(s) according to the present disclosure.

In some embodiments, there is provided methods for purifying a rAAV particle as described in the present disclosure, said method comprising the steps of a) providing a composition of
25 rAAVs comprising an antibody fragment e.g. obtained according to methods described in the present disclosure, wherein the composition comprises rAAVs and other components such as impurities, b) loading the composition obtained in step a) onto a binding substrate to the antibody fragment such that the rAAVs comprising the antibody fragment are retained on the substrate, c) recovering the rAAVs, comprising the antibody fragment thereby isolating the
30 rAAV from other components in the AAV preparation. The binding substrate may be a column or one or more magnetic beads on which a target of the antibody fragment has been immobilized. The rAAV preparation may be a crude lysate or may be a purified or partially purified, for example, after filtration or chromatography steps.

A further aspect of the present disclosure relates to a method for improving transduction of a target cell by a rAAV, wherein the rAAV comprises a capsid protein comprising an antibody fragment. The antibody fragment may bind to at least one specific molecule on a target cell (e.g., an antigen or a target molecule), and the rAAV may transduce the target cell specifically, selectively, or preferentially as compared to the control rAAV (i.e. the corresponding AAV which does not comprise the engineered capsid protein, i.e. which does not comprise an antibody fragment). The AAV capsid protein may comprise VP1, VP2, or VP3. Each of VP1 and/or VP2 and/or VP3 may be from an AAV, wherein the AAV comprises AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof.

The nucleic acid constructs, vectors and/or AAV viral particles herein described can be formulated into pharmaceutical compositions. Therefore, herein provided is a pharmaceutical composition comprising a nucleic acid construct of the invention, a vector of the invention and/or an AAV viral particle of the invention together with a pharmaceutically acceptable carrier, excipient, and/or diluent.

The pharmaceutical composition of the invention may comprise a pharmaceutically acceptable excipient, carrier, buffer, stabilizer, and/or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material may be determined by the skilled person according to the route of administration.

The pharmaceutical composition may be provided in liquid form. Liquid pharmaceutical compositions generally include a liquid carrier such as water, or physiological saline solution. For injection at the site of affliction, the active ingredient will be in the form of an aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection, Hartmann's solution. Preservatives, stabilizers, buffers, antioxidants and/or other additives may be included, as required.

Dosages and dosage regimes can be determined within the normal skill of the medical practitioner responsible for administration of the composition.

In one aspect, there is provided an AAV vector or AAV particle or pharmaceutical compositions as described herein, for use as a medicament.

In another embodiment, the present invention also encompasses the use of the nucleic acid constructs, vectors, viral particles and/or pharmaceutical compositions described herein for the treatment or the prevention of a disease or condition in a patient.

Herein provided is a nucleic acid construct of the invention, a vector of the invention, an AAV viral particle of the invention, and/or a pharmaceutical composition of the invention for use in a method of treating or preventing a disease or a condition in a patient in need thereof. The present invention further provides a method of treating or preventing a disease or condition in a patient in need thereof, said method comprising administering to the patient a therapeutically effective amount of a nucleic acid construct, a vector, a viral particle, and/or a pharmaceutical composition of the invention. The present invention also provides the use of a nucleic acid construct, a vector, a viral particle, and/or a pharmaceutical composition of the invention for the manufacture of a medicament for the treatment or prevention of a disease or condition in a patient in need thereof.

15 **Brief Description of the Figures**

Figure 1 shows an illustration of engineered viruses comprising capsid proteins inserted with an antibody fragment, such as a knob domain (Fig. 1A, 1C), or a bovine ultralong CDR-H3 (Fig. 1B, 1D), optionally via a linker (Fig. 1E-1F respectively).

Figure 2 shows an alignment of VP1, VP2, and VP3 amino acid sequences from AAV9 (top, middle and bottom line respectively).

Figure 3 shows an alignment of an AAV capsid protein from representative AAVs.

Figure 4 shows a diagram for the design of AAVs Cap constructs for the production of capsid proteins inserted with antibody fragments within the N-terminus of VP2. (AAV9 Del VP2: AAV9 Cap construct with a VP2 start codon mutation, as indicated with '*'; AAV9 VP2 N-term antibody fragment construct. The antibody fragment is shown as white rectangle).

Figure 5 shows a diagram for the design of AAVs Cap constructs for the production of capsid proteins inserted with antibody fragments within the common VP3 region. The antibody fragment is shown as white rectangle.

Figure 6 shows silver staining of VP2-N-term antibody fragments engineered rAAV that were purified from AAV9 affinity column. 2 µl of Precision Plus Protein Unstained Protein

Standards (Bio-rad Laboratories) were used in the first lane. 5 μ l (about 5E10 viral particles) of purified rAAV were loaded on to each lane.

Figure 7 shows a Western blot of VP2-N-term antibody fragment engineered rAAV that were purified from AAV9 affinity column. 10 μ l of Chameleon Duo pre-stained protein ladder (Li-Cor Biosciences) were used in the first and last lane. In Figure 7A, 2 μ l (about 2E10 viral particles) of purified rAAV were loaded on to each lane. In Figure 7B, 10 μ l of rAAV elution from C5 protein immobilized on Pierce NHS-activated magnetic beads were loaded on to each lane. In Figure 7C, 10 μ l of rAAV elution from uncoated magnetic beads that capped with ethanolamine as control were loaded on to each lane.

- Figure 8 shows transduction of VP2-N-term antibody fragments engineered rAAV in Ad-293 HEK GPI-C5 cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 8A shows rAAV with MOI=5E4 vg/cell, $p<0.0001$ for one-way ANOVA test, $p<0.0001$ for multiple comparisons between engineered rAAV and AAV9; Figure 8B: rAAV with MOI=1E5 vg/cel, $p<0.0001$ for one-way ANOVA test, $p<0.0001$ for multiple comparisons between engineered rAAV and AAV9; Figure 8C: rAAV with MOI=2E5 vg/cell, $p<0.0001$ for one-way ANOVA test, $p<0.0001$ for multiple comparisons between engineered rAAV.

Figure 9 shows binding to human C5 (Fig.9A). No binding was detected in wild-type AAV9 or to human homologue C3b (Fig.9B).

- Figure 10 illustrates engineering strategies of antibody fragments inserted within AAV capsids, showing that the number of antibody fragments inserted within AAV capsids such as knob domains as illustrated, may be varied.

- Figure 11 illustrates engineering strategies where two antibody fragments are inserted within AAV capsids, wherein two exemplary knob domains have different sequences (i.e. bispecific AAV capsids). The number of combined knob domains may be varied.

- Figure 12 shows transduction of various configurations of engineered rAAV9 in HEK293 cells expressing the target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 12A shows transduction of rAAV9 comprising the antibody fragment B3.2 at a loop insertion site in VP2 and VP3. Figure 12A showed $p<0.0001$ for two-way ANOVA test, $p<0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of

1E4 vg/cell and 1E5 vg/cell respectively, and $p=0.0154$ at dose of 1E3 vg/cell. Figure 12B shows transduction of rAAV9s comprising the antibody fragment B3.2 at a loop insertion site in VP2, and Figure 12C shows transduction of rAAV9 comprising the antibody fragment B3.2 at a loop insertion site on VP1. Figure 12B showed $p<0.0001$ for two-way ANOVA test, $p<0.0001$ for multiple comparisons between Ad-293 HEK GPI-C5 cells and Ad-293 HEK WT cells at dose of 1E3vg/cell, 1E4 vg/cell and 1E5 vg/cell respectively. Figure 12C showed $p<0.0001$ for two-way ANOVA test, $p<0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E4 vg/cell and 1E5 vg/cell respectively. Figure 12D shows the transduction of wild-type AAV9. Figure 12D showed $p=0.0029$ for two-way ANOVA test, $p=0.0019$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E3 vg/cell and $p=0.004$ at 1E4vg/cell. Solid circle: Ad-293 HEK GPI-C5 cells; empty circle: Ad-293 HEK WT cells. Figure 12E shows a Western blot for VP proteins.

Figure 13 shows transduction of an engineered rAAV9s in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 13A shows transduction of rAAV9 comprising the antibody fragment D3.1 at the N-terminus of VP2. Figure 13A showed $p<0.0001$ for two-way ANOVA test, $p<0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell. Figure 13B shows transduction of rAAV9 comprising the antibody fragment D3.2 at a loop insertion site. Figure 13B showed $p<0.0001$ for two-way ANOVA test, $p<0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E4vg/cell and 1E5 vg/cell respectively. Figure 13C shows transduction of wild-type AAV9. Figure 13C showed $p=0.01$ for two-way ANOVA test, $p=0.0005$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell.

Figure 14 shows transduction of an engineered rAAV9s in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 14A shows transduction of rAAV9 comprising the antibody fragment E3.1 at the N-terminus of VP2. Figure 14A showed $p<0.0001$ for two-way ANOVA test, $p<0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell, 0.0085 for 1E4vg/cell, and 0.0176 for 1E3vg/cell. Figure 14B shows transduction of rAAV9 comprising the antibody fragment E3.2 at a loop insertion site. Figure 14B showed $p<0.0001$ for two-way ANOVA test,

p<0.0001 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell and 0.0097 for 1E4vg/cell. Figure 14C shows transduction of wild-type AAV9. Figure 14C showed p<0.0001 for two-way ANOVA test, p<0.0001 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5
5 vg/cell.

Figure 15 shows transduction of an engineered rAAV9s in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 15A shows transduction of rAAV9 comprising the antibody fragment G3.1 at the N-terminus of VP2. Figure 15A showed
10 p<0.0001 for two-way ANOVA test, p<0.0001 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell, 0.0070 for 1E4vg/cell, and 0.0202 for 1E3vg/cell. Figure 15B shows transduction of rAAV9 comprising the antibody fragment G3.2 at a loop insertion site. Figure 15B showed p=0.001 for two-way ANOVA test, p=0.0003 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT
15 cells at dose of 1E5 vg/cell. Figure 15C shows transduction of wild-type AAV9. Figure 17C showed p=0.001 for two-way ANOVA test, p<0.0001 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell.

Figure 16 shows transduction of an engineered rAAV1 in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested
20 for each dose. Error bars show the SD of four replicates. Figure 16A shows transduction of rAAV1 comprising the antibody fragment B3.1 at the N-terminus of VP2. Figure 16A showed p=0.1315 for two-way ANOVA test. Figure 16B shows transduction of wild-type AAV1. Figure 16B showed p=0.0003 for two-way ANOVA test, p<0.0001 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E4 vg/cell and 0.04
25 at 1E5vg/cell.

Figure 17 shows transduction of an engineered rAAV6 variant ("AAV6var") in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 17A shows transduction of rAAV6var comprising the antibody fragment B3.1 at the N-terminus of
30 VP2. Figure 17A showed p<0.0001 for two-way ANOVA test, p<0.0001 for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of 1E5 vg/cell and 1E4vg/cell. Figure 17B shows transduction of AAV6var that does not comprise an

antibody fragment. Figure 17B showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E5$ vg/cell.

Figure 18 shows transduction of an engineered rAAV8s in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 18A shows transduction of rAAV8 comprising the antibody fragment B3.1 at the N-terminus of VP2. Figure 18A showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E5$ vg/cell and $1E4$ vg/cell. Figure 18B shows transduction of rAAV8 comprising the antibody fragment B3.2 at a loop insertion site. Figure 18B showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E5$ vg/cell, $1E4$ vg/cell and $1E3$ vg/cell. Figure 18C shows transduction of AAV8 wild-type (i.e. not inserted with an antibody fragment). Figure 18C showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E5$ vg/cell. Figure 18D shows a Western blot for VP proteins.

Figure 19 shows transduction of an engineered rAAV8 in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 19A shows transduction of rAAV8 comprising the antibody fragment E3.1 at the N-terminus of VP2. Figure 19A showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E6$ vg/cell and 0.0017 at $1E5$ vg/cell. Figure 19B shows transduction of AAV8 wild-type. Figure 19B showed $p = 0.0002$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E6$ vg/cell.

Figure 20 shows transduction of an engineered rAAVTT in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 20A shows transduction of rAAVTT comprising the antibody fragment B3.1 at the N-terminus of VP2. Figure 20A showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E6$ vg/cell and $1E5$ vg/cell. Figure 20B shows transduction of AAVTT that does not comprise an antibody fragment. Figure 20B

showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E6$ vg/cell.

Figure 21 shows transduction of an engineered rAAVTT in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 21A shows transduction of rAAVTT comprising the antibody fragment G3.1 at the N-terminus of VP2. Figure 21A showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E6$ vg/cell and 0.0012 at $1E5$ vg/cell. Figure 21B shows transduction of AAVTT, not engineered. Figure 21B showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $1E6$ vg/cell and 0.0082 at $1E5$ vg/cell.

Figure 22 shows transduction of an engineered rAAV6 variant ("AAV6var") in HEK293 cells expressing a target of an antibody fragment, as compared to HEK293 WT control cells. Four replicates were tested for each dose. Error bars show the SD of four replicates. Figure 22A shows transduction of rAAV6var comprising the antibody fragment D3.1 at the N-terminus of VP2. Figure 22A showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $5E5$ vg/cell. Figure 22B shows transduction of AAV6var that has not been inserted with an antibody fragment. Figure 22B showed $p < 0.0001$ for two-way ANOVA test, $p < 0.0001$ for multiple comparisons between Ad-293 HEK target cells and Ad-293 HEK WT cells at dose of $5E5$ vg/cell and 0.0002 at $1E5$ vg/cell.

Figure 23 shows a Western blot of various engineered rAAV9s

Figure 24 shows a Western blot of engineered rAAVs from different serotypes.

25 **Example 1: Engineering of recombinant AAV (rAAV)**

AAV capsid proteins were engineered to generate capsid proteins, and rAAVs, comprising an antibody fragment, wherein the antibody fragment comprised a bovine ultralong CDR-H3, or a knob domain of a bovine ultralong CDR-H3.

Sequences of bovine ultralong CDR-H3 and portions thereof, for example knob domains, were obtained according to methods previously described in WO2021191424. Cows were

immunized using various (7) immunogens, to generate antibody fragments described herein, that bind to 7 different targets. In the following example, letters A to G respectively refer to antibody fragments that bind to each one of the 7 targets. For one same target, e.g. “A”, number 1 (i.e. “A1”) refers to a knob domain of a bovine ultralong CDR-H3, number 2 (i.e. “A2”) refers to the whole bovine ultralong CDR-H3. Number 3 and variants (e.g. 3.1, 3.2) each refer to an antibody fragment which is a knob domain, fused to at least one linker. Number 4 and variants (e.g. 4.1, 4.2) each refer to an antibody fragment which is a bovine ultralong CDR-H3, fused to at least one linker. Sequences of antibody fragments including as exemplified herein are represented in **Table 1**. Linkers are indicated in italics.

10 The antibody fragments were inserted within capsid proteins of an AAV according to the present disclosure. In one approach, capsid protein VP2 of AAV9 was engineered by inserting an antibody fragment within the N-terminus of VP2. In this approach, firstly, a deletion in VP2 was made in a first *cap* gene of AAV (“AAV9 del VP2”), so that VP2 was not produced from this *cap* gene, while VP1 and VP3 were produced. Then, a second *cap* gene was provided, where VP2 comprised an antibody fragment. When expressed together, the two *cap* genes produced VP1 and VP3 that did not comprise an antibody fragment (e.g. a knob domain) whereas VP2 did. The corresponding design of Rep/Cap constructs is illustrated in Figure 4.

In another approach, antibody fragments were inserted within the common VP3 region of a *cap* gene of AAV9, resulting in engineered rAAV9 comprising VP1, VP2 and VP3 proteins that all comprise an antibody fragment. The corresponding design of Rep/Cap constructs is shown in Figure 5.

Table 2 provides sequences for AAV9 VP1 (SEQ ID NO: 40), AAV9 VP2 (SEQ ID NO: 41), and AAV9 VP3 (SEQ ID NO: 42). Corresponding nucleic acid sequences are provided in SEQ ID NO: 48-50. Insertion sites Gly455, Asn498, Gln588 in the common VP3 region are according to the numbering in SEQ ID NO: 40 (VP1) and are indicated in bold, underlined text. Corresponding Gly (G), Asn (N), and Gln (Q) residues in VP2 are Gly318, Asn361, and Gln451 according to the numbering in SEQ ID NO: 41, and Gly253, Asn296, and Gln386 in VP3 according to the numbering in SEQ ID NO:42. Unless otherwise specified, in the context of the present disclosure, insertion sites Gly455, Asn498, Gln588 in the common VP3 region will refer to those residues at those exact positions in VP1 with reference to SEQ ID NO:40 or any corresponding position in VP2 or VP3 as described herein.

DNA sequences encoding engineered capsid proteins inserted with antibody fragments were synthesized de-novo and cloned in backbone plasmids according to standard cloning methods allowing the production of recombinant expression vector.

Sequences of Rep/Cap of AAV9, and engineered capsid proteins, are represented in Table 2.

- 5 In SEQ ID NO:38 (Wild-type Rep/Cap), the VP2 start codon beginning at nucleotide 2424 is indicated in bold, underlined text. A nucleotide sequence for the AAV9 Rep/Cap comprising a mutation in the VP2 start codon is given in SEQ ID NO: 39. The VP2 start codon beginning at nucleotide 2424 is indicated in bold, underlined text. In this sequence, a G to C mutation (ACG>ACC) removed the start codon in VP2. This mutant will be referred to as “AAV9 del
10 VP2”.

- Representative sequences of AAV9 (Rep/Cap) comprising an antibody fragment inserted at the N-terminal of VP2 is provided in SEQ ID NO: 47. Representative sequences of AAV9 (Rep/Cap) comprising an antibody fragment inserted after Gly455, Asn498, Gln588 in the common VP3 region are provided respectively in SEQ ID NO: 43, 44 and 45. Isolated
15 sequences of VP1, VP2 and VP3 inserted at those positions in the common VP3 region are represented in SEQ ID NO: 51-53 (respectively VP1, VP2, VP3 Gly455), SEQ ID NO: 54-56 (respectively VP1, VP2, VP3 Asn498), SEQ ID NO: 57-59 (respectively VP1, VP2, VP3 Gln588). In those sequences, the antibody fragment A1 or B1 (knob domain) was represented. It will be understood that the engineered capsid proteins comprising alternative antibody
20 fragments have been generated using similar methods of insertions and sequences can be readily obtained by replacing the A1 or B1 sequence with any other sequence coding for an antibody fragment, including as exemplified herein and/or as provided in Table 1.

- Rep/cap were expressed under the control of a CMV promoter. SEQ ID NO: 46 provides an example of a nucleotide sequence including a CMV promoter indicated in lower case text, an
25 antibody fragment indicated in bold, underlined text; AAV9 VP2 indicated in double underlined text; and SV40 polyA is in italic lowercase text.

The AAV helper plasmid pALD-X80 was purchased from Aldevron, LLC. The cis payload plasmid is ssAAV ITR-CMV-eGFP-SV40-ITR.

- Linkers used in the following examples include a linker comprising the sequence GGGGG
30 (SEQ ID NO: 95) (the corresponding nucleotide sequence used in the example comprises ggtggaggcgggggt (SEQ ID NO:96), or a linker comprising the sequence GGGGSGGGGS (or “GGGGSX2” or “G4Sx2” (SEQ ID NO: 97). Corresponding nucleotide sequences used in the

example comprise the sequence ggt gga ggc ggg agt gga ggt ggc ggg agt (SEQ ID NO: 98) and ggc ggg ggt gga agt ggc gga ggt gga agt (SEQ ID NO: 99). For example, when two linkers are used, each GGGGSX2 linker located at each N-or C-terminal end of the antibody fragment, one linker may be encoded by sequence SEQ ID NO: 98 and the other linker may be encoded by sequence SEQ ID NO: 98 or SEQ ID NO: 99. Alternatively, when two linkers are used, each GGGGSX2 linker located at each N-or C-terminal end of the antibody fragment, one linker may be encoded by sequence SEQ ID NO: 99 and the other linker may be encoded by sequence SEQ ID NO: 99.

As described above, antibody fragments exemplified are represented in Table 1, including sequences of antibody fragments genetically fused to at least one linker as described herein (represented in *italic*).

Recombinant AAV (rAAV) vectors comprising engineered capsid proteins according to the present disclosure were produced by standard transfection methods suitable for producing rAAV, for example as described below.

Expi293 cells (Thermo Fisher) were passaged using Expi293 Expression Media (Thermo Fisher) in shake flasks. The Expi293 cells were cultured on an orbital shaker in an Eppendorf incubator. To set up the production flasks, a shake flask was inoculated the day before transfection. Viable cell density was calculated using a Vi-Cell Blu (Beckman Coulter).

A transfection complex was created for each flask as follows for the production flask with a 30 mL working volume. Separately, the Cis plasmid, the engineered Rep/Cap plasmid (AAVs), and the helper plasmid (pALD-X80) (for AAVs inserted with antibody fragments in the common VP3 region); or, the Cis plasmid, the helper plasmid (pALD-X80), and the engineered Rep/Cap plasmids (AAVs), including the AAV9 Del VP2 containing plasmid (for AAVs inserted with antibody fragments in the VP2 N-terminal region only) were diluted in Opti-PRO serum free media, then added Polyethylenimine (PEI) MAX (Polysciences Inc), vortexed and incubated for at room temperature. Transfection complexes were then added to shake flasks containing cells. Cells were cultured with the transfection mixture with constant agitation.

After growth, flasks were spiked with the concentrated AAV lysis buffer and Benzonase (MilliporeSigma). This mixture was incubated with constant agitation. The mixture was clarified by centrifugation. Samples were stored at -80°C until further analysis.

rAAV genome copies titer determination

The rAAV sample was thawed at room temperature, then briefly vortexed and centrifuged for one minute. After this, each sample was added to an individual well of a 96-well PCR plate. The plate was then transferred to a Bio-Rad thermal cycler and was heated for, followed by heat inactivation, then cooled.

- 5 The rAAV sample was mixed with a ddPCR master mix composed of Supermix for Probes (No dUTP; Bio-Rad Laboratories), forward primer GAACCGCATCGAGCTGAA (SEQ ID NO: 100), reverse primer TGCTTGTCGGCCATGATATAG (SEQ ID NO: 101), Probe 6-Fam/Zen/3'IBFQ: ATCGACTTCAAGGAGGACGGCAAC (SEQ ID NO: 102), and DNase-free water. This primer set targets eGFP transgene region. Each sample was run in duplicate
10 in a 96-well PCR plate.

The plate was placed into the Bio-Rad QX-200 droplet generator and droplets were generated per the manufacturer's instructions. After droplet generation, the plate was heat-sealed with a foil covering and placed into a Bio-Rad thermocycler programmed to run one step of 95°C for 10 minutes, followed by 40 cycles of 95°C for 30 seconds and 56°C for 1 minute, then one step
15 of 98°C of 10 minutes, final hold at 4°C. For PCR amplification steps, the ramping was set at 2.5°C/seconds.

Once complete, the plate was placed into a Bio-Rad QX200 droplet for droplet reading per the manufacturer's instruction. The concentration of vector genomes (VG/mL) was quantified using the following formula:

- 20 VG/ML: $X = [(aY)(1000/b)]$, where:

X is VG/mL;

a is volume of the ddPCR reaction (25 µl);

Y is the ddPCR readout in copies per microliter;

b is the volume of vector in the ddPCR (5 µL);

- 25 Assay acceptance criteria were defined as follows: The %CV between the replicates must be ≤15%; each reaction well must have ≥1,000 accepted droplets.

Viral particle quantification by ELISA

The viral particle titer was determined by AAV9 titration ELISA kit (PROGEN) according to the manufacturer's instructions and mouse monoclonal ADK9 antibody was used for both the

capture and detection steps. Washes in the provided 1X Assay Buffer (ASSB) were performed between each step using a AquaMax 4000 microplate washer (Molecular Devices). Samples were detected with a SpectraMax M5e plate reader (Molecular Devices). Capsid titers were then interpolated from the standard curve based on reading results.

5 **Results**

VG and VP titers obtained for rAAV9 engineered-VP2 N-term inserted antibody fragments are represented in **Table 3**. Wild-type AAV9 and AAV9 del VP2 (VP2 deleted) were used as controls.

10 VG and VP titers obtained for rAAV9 engineered with antibody fragments inserted within the common VP3 region (Gly455, Asn498) are represented in **Table 4**. Wild-type AAV9 was used as control. In this experiment, the productivity of AAV9 Gln588 inserted antibody fragment was lower (data not shown).

15 The results indicate that rAAV9 with engineered capsid proteins (comprising a viral vector with a CMV promoter driven eGFP transgene) could be successfully produced without significant impact on capsid assembly.

Table 3: Productivity of rAAV9 engineered-VP2 N-term inserted antibody fragments

rAVV tested	VG titer (lysates)	VP titer (lysates)	Full% (lysates)
AAV9 VP2 N-term B1	4.85E+10	3.55E+11	14%
AAV9 VP2 N-term B2	5.35E+10	3.35E+11	16%
AAV9 VP2 N-term B3.1	5.10E+10	3.00E+11	17%
AAV9 VP2 N-term B4.1	4.40E+10	2.80E+11	16%
AAV9 VP2 N-term C1	1.15E+11	7.53E+11	15%
AAV9 VP2 N-term C2	1.48E+11	9.77E+11	15%
AAV9 VP2 N-term C3.1	9.68E+10	5.69E+11	17%
AAV9 VP2 N-term C4.1	8.62E+10	5.30E+11	16%
AAV9 VP2 N-term A1	9.81E+10	5.85E+11	17%

rAAV tested	VG titer (lysates)	VP titer (lysates)	Full% (lysates)
AAV9 VP2 N-term A2	7.87E+10	4.78E+11	17%
AAV9 VP2 N-term A3.1	8.80E+10	4.92E+11	18%
AAV9 VP2 N-term A4.1	5.69E+10	3.43E+11	17%
AAV9 Del VP2	4.25E+10	1.75E+11	24%
AAV9	1.20E+11	6.50E+11	18%

Table 4: Productivity of rAAV9 engineered-common VP3 region inserted antibody fragments

rAAV tested	VG titer (lysates)	VP titer (lysates)	Full% (lysates)
AAV9 Gly455 A1	1.10E+09	1.20E+09	92%
AAV9 Gly455 A2	1.10E+09	8.50E+08	129%
AAV9 Gly455 B1	2.50E+09	3.20E+09	78%
AAV9 Gly455 B2	1.30E+09	1.80E+09	72%
AAV9 Gly455 C1	1.20E+09	3.90E+09	31%
AAV9 Gly455 C2	1.30E+09	3.40E+09	38%
AAV9 Asn498 A1	6.50E+08	1.20E+08	542%
AAV9 Asn498 A2	6.50E+08	1.70E+08	382%
AAV9 Asn498 B1	8.40E+08	1.10E+09	76%
AAV9 Asn498 C1	5.40E+08	1.10E+08	491%
AAV9 Asn498 C2	5.80E+08	1.40E+08	414%
AAV9	6.60E+10	4.90E+11	13%

Example 2: Assessment of the productivity of purified rAAV particles

Some rAAV samples obtained as described in the previous example were further purified as follows. The rAAV samples were purified with POROS CaptureSelect AAV9 Affinity Resin (ThermoFisher) packed onto Tricorn 10/50 column. The chromatography system was run with AKTA Avant 150 (Cytiva). The column was first equilibrated. The clarified upstream sample was loaded on to the column. The column was washed with equilibration buffer. The column was then eluted with elution buffer. The pH was adjusted. The column can be reused after stripping, cleaning, and equilibration. The loaded materials, flow through, and eluates were collected for further analysis.

rAAV genome copies titer and viral particle quantification were determined using the methods described in the previous example. Recovery rates were also calculated according to standard methods.

In addition, silver staining of VP protein gel was performed; about 5.0×10^{10} purified virus particles for each sample was heated at 95°C in the presence of loading buffer and separated on a NuPAGE Novex Bis-Tris 4-12% protein mini-gel (Invitrogen). The gel was stained based on the chemical reduction of silver ions to metallic silver on a protein band using SilverQuest silver staining kit (Invitrogen) according manufacturer's protocol. Once the gel was stained, it was imaged using Chemidoc Imager (Bio-Rad Laboratories) with Imaging Lab software (Bio-Rad Laboratories). The results are presented in Table 5 and 6.

Table 5: Productivity of various purified rAAV9 engineered-VP2 N-term antibody fragments

rAAV	VG titer (purified)	VP titer (purified)	Full% (purified)
AAV9 VP2 N-term B1	1.62×10^{12}	2.32×10^{13}	7%
AAV9 VP2 N-term B2	1.69×10^{12}	2.21×10^{13}	7.6%
AAV9 VP2 N-term B3.1	1.43×10^{12}	2.17×10^{13}	6.5%
AAV9 VP2 N-term B4.1	1.73×10^{12}	2.32×10^{13}	7.5%
AAV9	1.36×10^{12}	1.46×10^{13}	9.3%

Table 6: Recovery rates of rAAV9 engineered-VP2 N-term antibody fragments

rAAV	Conditions	Load	Eluate	Flow through
AAV9 VP2 B1	VG total	9.32E+13	6.87E+13	1.64E+12
	VG recovery	NA	74%	2%
	VP total	9.36E+14	6.77E+14	8.62E+12
	VP recovery	NA	72%	1%
	Full %	10%	10%	19%
AAV9 VP2 B2	VG total	8.47E+13	5.85E+13	1.24E+12
	VG recovery	NA	69%	2%
	VP total	7.55E+14	6.81E+14	1.26E+13
	VP recovery	NA	90%	2%
	Full %	11%	9%	10%
AAV9 VP2 B3.1	VG total	7.60E+13	5.38E+13	1.10E+12
	VG recovery	NA	71%	2%
	VP total	6.36E+14	5.74E+14	7.11E+12
	VP recovery	NA	90%	1%
	Full %	12%	9%	15%
AAV9 VP2 B4.1	VG total	9.12E+13	7.42E+13	2.07E+12
	VG recovery	NA	81%	3%
	VP total	7.93E+14	6.38E+14	1.05E+13
	VP recovery	NA	81%	2%
	Full %	12%	12%	20%
AAV9	VG total	7.54E+13	5.77E+13	1.39E+12
	VG recovery	NA	77%	2%
	VP total	5.99E+14	4.84E+14	6.27E+12
	VP recovery	NA	81%	1%
	Full %	13%	12%	22%

The results show that engineered rAAV9 viral particles having a nucleic acid comprising a CMV promoter driving expression of eGFP transgene could be successfully produced and purified with AAV9 affinity column.

5

The recovery rates for engineered rAAV9 were close to rAAV9 with high titer yield. In addition, figure 6 shows the expected sizes for the antibody fragment N-term fused VP2 and the wild-type AAV9 VP2, in addition to intact VP1 and VP3.

Example 3: Assessment of antigen binding of the engineered rAAVs

- 5 In this example, the binding properties of the antibody fragments when inserted within rAAV were assessed. One of the target C5 was used as an illustration.

Purified C5 protein was immobilized on Pierce NHS-activated magnetic beads (ThermoFisher). Uncoated magnetic beads that were capped with ethanolamine were used as control for immunoprecipitation. 10 μ L beads were incubated with 10 μ L rAAV (21 °C, 30
10 min). Then the beads were washed 5 times with 200 μ L PBS+ 0.05%Tween20, one time with 200 μ L PBS, and eluted in 50 μ L SDS-PAGE loading buffer (Li-Cor Biosciences) at 95°C.

AAV capsid proteins were screened in binding ELISAs according to the following method: 96-well ELISA plates (Nunc Maxisorp) were coated with a 2 μ g/mL solution of either human C5 (purified in house) or human C3b (Complement Technologies, Inc), in carbonate-bicarbonate
15 buffer (Sigma). All washing steps comprised four wash cycles with PBS, 0.05% Tween 20. Blocking Buffer was PBS, 10% Sea Block (Thermo Fisher [v/v]). Two-fold, eleven-point serial dilutions of AAV capsid were prepared in Assay Buffer (PBS, 0.05% Tween 20 (v/v), 10 % Sea Block [v/v]), to give a final assay range of 10 nM – 10 pM. For detection, a 1/20 dilution
20 of a biotinylated anti-AAV9 antibody (Progen ADK9) secondary antibody was used with a 1/1000 dilution of Streptavidin HRP (Thermo). Finally, ‘One-Step’ 3,3',5,5' Tetramethylbenzidine (Thermo Scientific) was used to reveal. The reaction was stopped with the addition of a 2% (w/v) NaF solution and the OD was measured at 630 nm. Data were fitted using a 4-parameter logistic fit (GraphPad Prism 8.1.1) to derive apparent equilibrium dissociation constants (K_D *app*).

- 25 The rAAV purified with AAV9 affinity column were denatured with SDS-PAGE loading buffer (Li-Cor Biosciences) at 95°C, size fractionated by 4-15% SDS–PAGE Bio-Rad mini-Protean TGX precast Gels, transferred to 0.2 μ m PVDF membrane with Trans-Blot Turbo semi-dry transfer system (Bio-Rad Laboratories) and detected with the B1 antibody fragment that recognizes a common epitope at the C terminus of AAV VP proteins. The magnetic beads
30 pull-down eluted proteins were separated on an SDS-PAGE gel and observed by Western blotting in same condition. The membrane was first blocked with Intercept blocking buffer TBS (Li-Cor Biosciences) for 1 hour at room temperature, then stained with primary antibody

(Anti-AAV VP1/VP2/VP3 mouse monoclonal B1, Progen, 1:250) with Intercept antibody diluent T20 TBS (Li-Cor Biosciences) over night at 4°C. Then the membrane was washed with TBS-Tween20 for 3 times with 10 min each, then stained with Secondary Antibody (IRDye 800CW Donkey anti-Mouse IgG, Li-Cor Biosciences, 1:2000) in Intercept antibody
 5 diluent T20 TBS (Li-Cor Biosciences) for 1h at RT, then washed with TBS-Tween20 for 3 times with 10 min each. Proteins were visualized using a Li-Cor Odyssey CLx far red imager (Li-Cor Biosciences).

The total signal for each VP band was measured by Image Studio software (Li-Cor Biosciences), and VP2 components for rAAV were quantified as percentage of VP2 signal
 10 density compared to total VP1/VP2/VP3 signal density as shown in Table 7.

Table 7: VP2 components of various VP2-N-term antibody fragments engineered rAAV9

VP2 component	% VP2 components (purified by AAV9 column)	% VP2 components (C5 pull-down)
AAV9 VP2 N-term B1	26%	42%
AAV9 VP2 N-term B2	28%	41%
AAV9 VP2 N-term B3.1	20%	35%
AAV9 VP2 N-term B4.1	23%	31%
AAV9	20%	

Engineered rAAVs were capable of binding its target as confirmed by its successful pull-down with C5 protein immobilized magnetic beads (Figure 7B) but not with uncoated control
 15 magnetic beads (Figure 7C). The engineered rAAVs all bound to human C5 but not human C3 (Figure 9). No binding to either C5 or C3 was detected for the wild type AAV9 capsid. For binding to C5, K_D *app* values, derived from $n=2$ experiments, were as follows: B1 = 386 pM (range 233-640 pM), B2 = 720 pM (range 360 pM – 1.44 nM), B3.1 = 958 pM (range 700 pM – 1.311 nM), B4.1 = 1.37 nM (range 859 pM – 2.183 nM). Binding to human C5 as measured
 20 by OD at 630 nm is shown in Figure 9A.

Finally, the results indicate that an exemplary antibody fragment (i.e. a knob domain) may be selected due to its binding properties to a target of interest, that would not be expected to normally interact with a AAV that does not comprise the antibody fragment, therefore conferring new binding properties to the engineered rAAV.

5 **Example 4 – Assessment of the ability of the engineered rAAVs to transduce cells**

- The human-derived Ad-293 HEK (Agilent Technologies) were passaged in DMEM + 10% FBS + 1% Penicillin/Streptomycin (all from Thermo Fisher Scientific). The plasmid with CMV promoter driven expression of C5 protein fused to C-terminal of glycosyl phosphatidylinositol (GPI) as anchor membrane protein were used for transient transfection. Ad-293 HEK GPI-C5
- 10 cells were prepared as cells were transfected using X-tremeGene 360 Transfection reagent (Roche) according to the manufacturer's protocol. For a 75 cm² plate, 12 µg of GPI-C5 plasmid and 24 µl X-tremeGene transfection reagent (Roche) were added and mixed in 1200µl of Opti-MEM medium (Thermo fisher). The transfection reagent/DNA complex were incubated for 15 minutes at room temperature, then added to cells by dropwise.
- 15 The cells were cultured in DMEM + 10% FBS + 1% Penicillin/Streptomycin (all from Thermo Fisher Scientific). After 48 hours of transfection, the cells were seeded into 96-well plates in 100µl volume with density of 5E5 cells per ml and rAAV with serial dilution of multiplicity of infection (MOIs, vg/cell) were added to each well. Four replicates were tested for each dose. After 48h of transduction, cells were washed with PBS, then resuspend with FACS buffer
- 20 (PBS+3%FBS). Cells were run with Guava easyCyte HT flow cytometer (Illuminex) and eGFP expression were analyzed with Guavasoft 4.0 (Illuminex). For cell gating, forward and side scatter density plots (FSC vs SSC) were used based on size and granularity. Forward scatter height versus forward scatter area density plots (FSC-H vs FSC-A) were further used to exclude doublets.
- 25 All statistical analyses were performed, and graph were generated with GraphPad Prism 8 software (GraphPad Prism software, Inc.). A p value <0.05 was considered to be statistically significant. In the case of one group variable, one-way ANOVA was used. In the case of two group variables, two-way ANOVA was used. A *post hoc* analysis with a Bonferroni's test for multiple comparisons was tested.
- 30 Engineered AAVs specifically bind HEK cells enriched with GPI-C5 membrane anchor protein. The engineered rAAV showed enhanced transduction in Ad-293 HEK GPI-C5 cells as compared with rAAV9 (Figure 8).

Example 5: Engineered AAVs with variable number of copies of an antibody fragment

In the preceding examples, the engineered rAAVs comprised either 60 insertions wherein each of VP1, VP2, and VP3 comprised the insertion of an antibody fragment within the common VP3 region, or 5 insertions wherein the VP2 protein only comprised an insertion of an antibody fragment at its N-terminus.

Alternatively, the engineered capsid proteins were configured to produce rAAVs comprising either 5 copies or 55 copies of an antibody fragment. To produce various configurations of VP1, VP2, and VP3 insertions in the following experiments, the cloning strategies included:

Cloning strategy 1: Insertion of an antibody fragment within VP2 and VP3, but not in VP1 as illustrated in Figure 10F or 10H. This strategy was designed to result in about 55 insertions on the resulting engineered rAAV. Firstly, a deletion in the splice acceptor site for VP2 and VP3 was made in a first *cap* gene of AAV (called “VP1 sa del”, SEQ ID NO:70), so that neither VP2 nor VP3 were produced from this *cap* gene, while VP1 was produced. Then, a second *cap* gene was provided, where VP1 was deleted (“VP1 del”) and which comprised an antibody fragment. Together, the two *cap* genes produced VP1 that did not comprise an antibody fragment, VP2 and VP3 both comprising an antibody fragment.

Cloning strategy 2: Insertion of an antibody fragment to VP2, but not VP1 or VP3 as illustrated in Figure 10B. The antibody fragment was fused to/or inserted within the N-terminus of VP2. This strategy results in about 5 insertions of the antibody fragment in the rAAV vector. Firstly, a deletion in VP2 was made in a first *cap* gene of AAV (AAV9 del VP2), so that VP2 was not produced from this *cap* gene, while VP1 and VP3 were produced. Then, a second *cap* gene was provided, where VP2 comprised an antibody fragment. Together, the two *cap* genes produced VP1 and VP3 that did not comprise a knob domain whereas VP2 did.

Cloning strategy 3: Insertion of an antibody fragment within VP1, but not VP2 or VP3 (Figure 10A). This strategy was designed to result in about 5 insertions on the resulting rAAV. Firstly, a deletion in VP1 was made in a first *cap* gene of AAV (e.g., by deleting the start codon for VP1, as in a construct called “VP1 del”, e.g., SEQ ID NO:73), so that VP1 was not produced from this *cap* gene, while VP2 and VP3 were produced. Then, a second *cap* gene was provided, where VP1 had a deletion in the splice acceptor site for VP2 and VP3 (“VP1 sa del”, SEQ ID NO:70) and where VP1 comprised an antibody fragment. Together, the two *cap* genes produced VP2 and VP3 that did not comprise a knob domain whereas VP1 did.

Table 8 summarizes the productivity data for decorated AAV9s produced according to each of the cloning strategies 1, 2, and 3. The antibody fragment used in this example was B3.2, which comprises the knob domain B1 plus a linker GGGGSGGGGS at its C-terminal end and at its N-terminal end, that was added to the insertion site Gly455 in AAV9 VPs within the common VP3 region (of either VP2/VP3, or only VP2 or only VP1).

Results of the transduction assay in HEK293 cells are summarized in Figure 12A-D. Figure 12E shows a Western blot for VP proteins, illustrating a change in size of VPs following addition of the antibody fragment B3.2, as compared to a wild-type, i.e. not engineered, AAV9 control.

Consistent with properties of engineered rAAV9 with 60 copies of antibody fragments, the results demonstrate that engineered rAAV capsids with different numbers of antibody fragments could be successfully produced at high titers, and insertions could be added to VP1 alone, to VP2 alone, or to VP2 and VP3. All of the engineered AAV9s were capable of transducing HEK cells expressing the target of knob domain at higher percentages than a control AAV9 showing that the engineered AAV9s gained specificity to the target of the knob domain.

Table 8: Productivity of engineered AAVs

Cloning strategy	Rep/Cap Plasmids	Expected number of knob domains	VG titer	VP titer	Full%
1	AAV9 VP1 sa del	55	3.65E+10	7.29E11	5.01%
	VP1 del AAV9 Gly455 B3.2 (SEQ ID NO:75)				
2	AAV9 Del Vp2	5	3.05E11	5.88E12	5.19%
	AAV9 VP2 Gly 455 B3.2 (SEQ ID NO:71)				
3	AAV9 VP1 del	5	1.66E11	2.54E12	6.54%

Cloning strategy	Rep/Cap Plasmids	Expected number of knob domains	VG titer	VP titer	Full%
	VP1 sa del AAV9 Gly455 B3.2 (SEQ ID NO:74)				
Control	AAV9	-	1.37E11	2.26E12	6.06%

Example 6: Engineering of recombinant AAV9s using other antibody fragments

In this example, additional antibody fragments were tested (D, E, F, and G derived antibody fragments).

- 5 Engineered rAAV9s were produced and purified as described in the previous examples 1 and 2. Cloning strategy corresponding to cloning strategy 2 of example 5 was used to produce engineered rAAV9 either at the N-terminus of VP2, or at Gly455 in the loop insertion site of VP2 only. Transduction assays were conducted in HEK293.

Viral genome titers, capsid titers, and percentage of full capsids are summarized in Table 9.

- 10 Results of the transduction assay are shown in Figures 13-15.

The results demonstrate that the engineered rAAV9s comprising antibody fragments could be successfully produced at high titers. Overall, the engineered rAAV9s transduced HEK293 cells expressing the target of the knob domains to a greater extent when compared to transduction of control WT HEK293 cells.

- 15 **Table 9: Productivity of the engineered rAAVs**

Rep/Cap Plasmids	VG titer (lysates)	VP titer (lysates)	Full% (lysates)
AAV9 VP2 N-term E3.1 (SEQ ID NO: 76) (E1 knob + GGGGG linker)	3.20E+11	5.88E+12	5.44%
AAV9 VP2 N-term F3.1 (SEQ ID NO: 77) (F1 knob + GGGGG linker)	2.82E+11	3.70E+12	7.62%

Rep/Cap Plasmids	VG titer (lysates)	VP titer (lysates)	Full% (lysates)
AAV9 VP2 N-term G3.1 (SEQ ID NO: 78) (G1 knob + GGGGG linker)	3.21E+11	3.39E+12	9.47%
AAV9 VP2 Gly 455 E3.2 (SEQ ID NO: 79) (E1 knob + GGGGSX2-GGGGSX2 linkers)	2.27E+11	3.70E+12	6.14%
AAV9 VP2 Gly 455 F3.2 (SEQ ID NO: 80) (F1 knob + GGGGSX2-GGGGSX2 linkers)	1.36E+11	2.10E+12	6.48%
AAV9 VP2 Gly 455 G3.2 (SEQ ID NO: 81) (G1 knob + GGGGSX2-GGGGSX2 linkers)	9.64E+10	1.79E+12	5.39%
AAV9 VP2 N-term D3.1 (D1 knob + GGGGG linker)	2.18E+11	2.23E+12	9.78%
AAV9 VP2 Gly 455 D3.2 (D1 knob + GGGGSX2-GGGGSX2 linkers)	8.82E+10	1.38E+12	6.39%
AAV9	2.34E+11	2.98E+12	11.38%

Example 7: Engineered AAVs from different AAV serotypes

A selection of exemplary AAV variants was evaluated for their ability to be inserted with antibody fragments. Variants AAV1, an exemplary AAV6 variant (called “AAV6var”), AAV8, AAV9, and AAV true type (AAV-TT) were tested. These variants are representatives of the AAV clades A, B, E, and F.

Production of engineered AAVs, purification, titer determination, and transduction assays were carried out as described herein. AAV8 titration ELISA kit (PROGEN), AAV1 titration ELISA kit (PROGEN), AAV6 titration ELISA kit (PROGEN) and AAV2 titration ELISA kit (PROGEN) were used for AAV8, AAV1, AAV6var and AAVTT respectively for their viral particle quantification. Antibody fragments comprising either 1 copy of linker GGGGG (SEQ

ID NO: 95) or 2 copies of linker GGGGSGGGGS (SEQ ID NO: 97) were cloned into the exemplary AAV variants using cloning strategy as described herein to insert an antibody fragment within VP2 only. Representative sequences of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh.10, AAV-TT capsid proteins are provided in SEQ ID NO: 60-69 respectively.

Results

Viral genome titers, capsid titers, and percentage of full capsids are summarized in Tables 10-12. Results of the transduction assays are shown in Figures 16-22. Western blots in Figure 23 and Figure 24 show the change in size of VP fragments in various rAAVs following the addition of the antibody fragments comprising knob domains that bind to the targets.

Transduction results from experiments with engineered AAV variants were consistent with results from experiments with engineered AAV9. Overall, for all the AAV serotypes tested, the engineered AAVs transduced HEK293 cells expressing the target of the antibody fragments to a greater extent when compared to transduction of control WT HEK293 cells.

Table 10. Productivity of engineered AAV8s

Rep/Cap Plasmid	VG titer	VP titer	Full%
AAV8 VP2 N-term B1 (SEQ ID NO: 83)	2.84E+11	2.5E12	11.57%
AAV8 Del Vp2 (SEQ ID NO: 82)			
AAV8 VP2 N-term B2 (SEQ ID NO: 84)	2.74E+11	2.5E12	10.99%
AAV8 Del Vp2			
AAV8 VP2 N-term B3.1 (B1 + GGGGG linker) (SEQ ID NO: 85)	1.98E+11	2.2E12	9.07%
AAV8 Del Vp2			
AAV8 Vp2 N-term B4.1 (SEQ ID NO: 86)	1.91E+11	1.8E+13	10.45%

Rep/Cap Plasmid	VG titer	VP titer	Full%
(B2 + GGGGG linker)			
AAV8 Del Vp2			
AAV8 Del Vp2	4.49E+10	2.26E11	19.86%
AAV8 VP2 Gly 455 B3.2 (SEQ ID NO: 87)			
(B1 knob + GGGGSX2-GGGGSX2 linkers)			
AAV8	3.96E+11	2.8E+13	13.90%

Table 11. Productivity of engineered AAVs from different serotypes

Rep/Cap Plasmid	VG titer	VP titer	Full%
AAV1 Del Vp2 (SEQ ID NO: 88)	1.45E+11	8.10E+11	17.9%
AAV1 VP2 N-term B3.1 (SEQ ID NO: 89)			
(B1 knob + GGGGG linker)			
AAVTT Del Vp2 (SEQ ID NO: 90)	1.45E11	2.70E+12	8.3%
AAVTT VP2 N-term B3.1 (SEQ ID NO: 91)			
(B1 knob + GGGGG linker)			
AAV6var del Vp2	1.17E+11	1.41E+12	8.30%
AAV6var VP2 N-term B3.1			
(B1 knob + GGGGG linker)			
AAV1	1.49E+11	6.40E+11	23.28%
AAVTT	3.49E11	5.90E+12	5.9%
AAV6var	8.11E+10	7.46E+11	10.87%

Table 12. Productivity of engineered AAVs from different serotypes

Rep/Cap Plasmid	VG titer (lysates)	VP titer (lysates)	Full% (lysates)
AAV6var del Vp2	1.35E+11	1.21E+12	11.16%
AAV6var VP2 N-term D3.1 (D1 knob + GGGGG linker)			
AAV1 Del Vp2	2.34E+11	8.90E+11	26.29%
AAV1 VP2 N-term F3.1 (SEQ ID NO: 92) (F1 knob + GGGGG linker)			
AAVTT Del Vp2	2.63E+11	2.29E+12	11.48%
AAVTT VP2 N-term G3.1 (SEQ ID NO: 93) (G1 knob + GGGGG linker)			
AAV8 Del Vp2	3.26E+11	3.64E+12	8.96%
AAV8 VP2 N-term E3.1 (SEQ ID NO: 94) (E1 knob + GGGGG linker)			
AAV6var	9.65E+10	6.01E+11	16.06%
AAV1	3.04E+11	1.15E+12	26.43%
AAVTT	6.89E+11	6.65E+12	10.36%
AAV8	3.76E+11	2.26E+12	16.64%

Taken together, these data show that the capsid proteins described herein may be engineered to comprise knob domains and antibody fragments comprising knob domains which vary in size, sequence, specificity, and in number. The knob domains or antibody fragments comprising knob domains may be inserted at different sites in the AAV capsid protein, may be

inserted in different VP proteins, and may be inserted into different AAV (e.g., AAVs from different clades).

Sequences of antibody fragments and capsid proteins useful in the context of the present disclosure are represented in Table 1 and Table 2 respectively (Amino acid sequences (aa.) and nucleotide sequences (nt.)).

Table 1: Antibody fragments (*Linkers are indicated in italic*)

Sequence identifier and name	Sequence
SEQ ID NO: 1 A1 (nt.)	GTCTGTCCTGACGGATTAACTGGGGCTACGGCTGTGCCCGCG GTTCTTCCCGGTTTTGTACTAGACATGATTGGTGTGTTATGAC GAGAGAGCCGATTCACACACATACGGGTTCTGCACTGGCAATC GGGTGACGAAC
SEQ ID NO: 2 A1 (aa)	VCPDGFNWGYGCAAGSSRFCTRHDWCCYDERADSHTYGFCTGN RVTN
SEQ ID NO: 3 A2 (nt.)	ACTGTGCAGCAGAAAACACACCAGGTCTGTCCTGACGGATTAA ACTGGGGCTACGGCTGTGCCGCCGGTCTTCCCGGTTTTGTACT AGACATGATTGGTGTGTTATGACGAGAGAGCCGATTCACACA CATACGGGTTCTGCACTGGCAATCGGGTGACGAACACCTACGA GTTCCATGCCGATGCC
SEQ ID NO: 4 A2 (aa)	TVQQKTHQVCPDGFNWGYGCAAGSSRFCTRHDWCCYDERADSH TYGFCTGNRVNTNTEFHADA
SEQ ID NO: 5 B1 (nt.)	GGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCGG GATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGA GATACGGGGGATACGGCGCAGACTCGGGAGTGGGATCG
SEQ ID NO: 6 B1 (aa)	GCPPGYKSGVDCSPGSECKWGCYAVDGRRYGGYGADSGVGS
SEQ ID NO: 7 B2 (nt.)	GTGCACCAGCGGACCATCAAGTCCGGTTGTCCCCCTGGCTACA AGAGCGGCGTGGACTGCTCCCCGGGATCGGAGTGCAAATGGG GATGCTACGCCGTGGATGGACGGAGATACGGGGGATACGGCG CAGACTCGGGAGTGGGATCGACCTACACCCACGAGTTCTACGT G
SEQ ID NO: 8 B2 (aa)	VHQRTIKSGCPPGYKSGVDCSPGSECKWGCYAVDGRRYGGYGA DSGVGSTYTHEFYV
SEQ ID NO: 9 C1 (nt.)	TCATGCCCTGATGGATTCTTATCGTTCATGGGACGACTTCTG CTGTCCAATGGTGGGGAGATGTCTCGCACCCAGAAACACG
SEQ ID NO: 10 C1 (aa)	SCPDGFSYRSWDDFCCPMVGRCLAPRNT
SEQ ID NO: 11 C2 (nt.)	ACTTCTGTCCTTCAATCCACTAAGCCACAGAAATCATGCCCTG ATGGATTCTTATCGTTCATGGGACGACTTCTGCTGTCCAATG GTGGGGAGATGTCTCGCACCCAGAAACACGTATACCACGGAGT TCACAATTGAGGCA
SEQ ID NO: 12 C2 (aa)	TSVLQSTKPQKSCPDGFSYRSWDDFCCPMVGRCLAPRNTYTTEFT IEA

Sequence identifier and name	Sequence
SEQ ID NO: 13 A3 (aa)	GGGGGVCPDGFNWGYGCAAGSSRFCTRHDWCCYDERADSHTY GFCTGNRVTNNGGGGG
SEQ ID NO: 14 A4 (aa)	GGGGGTVQQKTHQVCPDGFNWGYGCAAGSSRFCTRHDWCCYD ERADSHTYGFCTGNRVTNTYEFHADAGGGGG
SEQ ID NO: 15 B3 (aa)	GGGGGGCPPGYKSGVDCSPGSECKWGCYAVDGRRYGGYGADSG VGSGGGGG
SEQ ID NO: 16 B4 (aa)	GGGGGVHQRTIKSGCPPGYKSGVDCSPGSECKWGCYAVDGRRY GGYGADSGVGSTYTHEFYVGGGGG
SEQ ID NO: 17 C3 (aa)	GGGGGSCPDGFSYRSWDDFCCPMVGRCLAPRNTGGGGG
SEQ ID NO: 18 C4 (aa)	GGGGGTSVLQSTKPQKSCPDGFSYRSWDDFCCPMVGRCLAPRNT YTTEFTIEAGGGGG
SEQ ID NO: 19 A3.1 (aa)	VCPDGFNWGYGCAAGSSRFCTRHDWCCYDERADSHTYGFCTGN RVTNNGGGGG
SEQ ID NO: 20 A4.1 (aa)	TVQQKTHQVCPDGFNWGYGCAAGSSRFCTRHDWCCYDERADSH TYGFCTGNRVTNTYEFHADAGGGGG
SEQ ID NO: 21 B3.1 (aa)	GCPPGYKSGVDCSPGSECKWGCYAVDGRRYGGYGADSGVGS GGG
SEQ ID NO: 22 B3.2 (aa)	GGGGSGGGGSGCPPGYKSGVDCSPGSECKWGCYAVDGRRYGGY GADSGVGSGGGGSGGGGS
SEQ ID NO: 23 B4.1 (aa)	VHQRTIKSGCPPGYKSGVDCSPGSECKWGCYAVDGRRYGGYGA DSGVGSTYTHEFYVGGGGG
SEQ ID NO: 24 C3.1 (aa)	SCPDGFSYRSWDDFCCPMVGRCLAPRNTGGGGG
SEQ ID NO: 25 C4.1 (aa)	TSVLQSTKPQKSCPDGFSYRSWDDFCCPMVGRCLAPRNTYTTEFT IEAGGGGG
SEQ ID NO: 26 E1 (aa)	TCPDGYTRTNYYCRRDGC GSWCNGAERQQPCIRGPCCCDLT
SEQ ID NO: 27 E2 (aa)	TTVHQQTHQDQTCPDGYTRTNYYCRRDGC GSWCNGAERQQPCI RGPCCCDLTYRTAYEYHVET
SEQ ID NO: 28 E3.1 (aa)	TCPDGYTRTNYYCRRDGC GSWCNGAERQQPCIRGPCCCDLTGGG GG
SEQ ID NO: 29 E3.2 (aa)	GGGGSGGGGSTCPDGYTRTNYYCRRDGC GSWCNGAERQQPCIRG PCCCDLTGGGGSGGGGS
SEQ ID NO: 30 F1 (aa)	RCPDGSIWRCGDVDGATCSLRPCACWVQKSSGGWYCKPPAT
SEQ ID NO: 31 F2 (aa)	TTVHQRTREGRCPDGSIWRCGDVDGATCSLRPCACWVQKSSGG WYCKPPATGATYEHVET

Sequence identifier and name	Sequence
SEQ ID NO:32 F3.1 (aa)	RCPDGSIWRCGDVDGATCSLRPCACWVQKSSGGWYCKPPATGG GGG
SEQ ID NO:33 F3.2 (aa)	GGGGSGGGGSRCPDGSIWRCGDVDGATCSLRPCACWVQKSSGG WYCKPPATGGGGSGGGGS
SEQ ID NO:34 G1 (aa)	SCPDGYTYRFGSVCTARWRNDCFRDADYDYQRRCYTTSLPQT
SEQ ID NO:35 G2 (aa)	TTVHQETRQSCPDGYTYRFGSVCTARWRNDCFRDADYDYQRR YTTSLPQTYIYEHSIHA
SEQ ID NO:36 G3.1 (aa)	SCPDGYTYRFGSVCTARWRNDCFRDADYDYQRRCYTTSLPQTGG GGG
SEQ ID NO:37 G3.2 (aa)	GGGGSGGGGSSCPDGYTYRFGSVCTARWRNDCFRDADYDYQRR CYTTSLPQTGGGGSGGGGS

Table 2: Rep/Cap sequences

Sequence identifier and name	Sequence
SEQ ID NO: 38 AAV9 Rep/Cap (nt.)	GTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGGCAGACCA TGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCA TTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCCGGGGTTTTAC GAGATTGTGATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCCGGC ATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTG CCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTG ACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGT GTGAGTAAGGCCCGGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGA GAGAGCTACTTCCACATGCACGTGCTCGTGGAACCACCGGGGTGAAA TCCATGGTTTTGGGACGTTTCCTGAGTCAGATTTCGCGAAAACTGATTC AGAGAATTTACCGCGGGATCGAGCCGACTTTGCCAACTGGTTCGCGG TCACAAAGACCAGAAATGGCGCCGGAGGCGGGAACAAGGTGGTGGAT GAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCC AGTGGGCGTGGAATAATATGGAACAGTATTTAAGCGCCTGTTTGAATCT CACGGAGCGTAAACGGTTGGTGCGCAGCATCTGACGCACGTGTCGCA GACGCAGGAGCAGAACAAAGAGAATCAGAATCCCAATTCTGATGCGCC GGTGATCAGATCAAAAACCTCAGCCAGGTACATGGAGCTGGTCGGGTG GCTCGTGGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGG ACCAGGCCTCATACTCCTTCAATGCGGCCTCCAACCTCGCGGTCCCA AATCAAGGCTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAA AACCGCCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTT CAGCAATCGGATTTATAAAATTTTGGAATAAACGGGTACGATCCCCA ATATGCGGCTTCCGTCTTTCTGGGATGGGCCACGAAAAAGTTTCGGCAA GAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGGAAGACCAA CATCGCGGAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAAA

Sequence identifier and name	Sequence
	CTGGACCAATGAGAACTTTCCCTTCAACGACTGTGTCGACAAGATGGT GATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCGTGGAGTCGG CCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCA AGTCCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACCTCCAACAC CAACATGTGCGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCA GCAGCCGTTGCAAGACCGGATGTTCAAATTTGAACTACCCGCCGTCTG GATCATGACTTTGGGAAGGTCACCAAGCAGGAAGTCAAAGACTTTTTTC CGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTC AAAAAGGGTGGAGCCAAGAAAAGACCCGCCCCAGTGACGCAGATAT AAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGTC AGACGCGGAAGCTTCGATCAACTACGCAGACAGGTACCAAAACAAATG TTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGC GAGAGACTGAATCAGAATTCAAATATCTGCTTCACTACGGTGTCAAAA GACTGTTTAGAGTGCTTTCCCGTGTGAGAATCTCAACCCGTTTCTGTCTG TCAAAAAGGCGTATCAGAACTGTGCTACATTCATCACATCATGGGAA AGGTGCCAGACGCTTGCACTGCTTGCAGCTGGTCAATGTGGACTTGG ATGACTGTGTTTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCCG ATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCTG CGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCA ACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTACAAAATA CCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTCAACGCAGC AGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAA GGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTT CCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGCAACCTCGGGCG AGCAGTCTTCCAGGCCAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTT GAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCA GTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGC ACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTGGCGACACAGA GTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTC AGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAAGTGGC AGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTCGGGAAATG GCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCAACAGCAC CCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTACAAGCAAAT CTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGG CTACAGCACCCCCTGGGGGTATTTTGACTTCAACAGATTCCACTGCCAC TTCTCACCACTGACTGGCAGCGACTCATCAACAACAACCTGGGGATTTC CGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTAGGTCAAAGAG GTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCTTACCAGC ACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTACGTGCTCG GGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCAT GATTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGT GGGTCGTTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGAAATGCTA AGAACGGGTAAACAACCTCCAGTTCAGCTACGAGTTTGAGAACGTACCT TTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACTAATGAAT CCTCATCGACCAATACTTGTACTATCTCTCAAAGACTATTAACGGTT CTGGACAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGACCCAGCA ACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGCTACCGAC AACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTG CTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGAT GAATCCTGGACCTGCTATGGCCAGCCACAAGAAGGAGGAGGACCGTTT CTTCCCTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACGTGAAGA GACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAAT

Sequence identifier and name	Sequence
	TAAAACTACTAACCCTGGTAGCAACGGAGTCCTATGGACAAGTGGCCAC AAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAA ACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACC TGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTC ACCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCCCTCCTCA GATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTC AACAAAGGACAAGCTGAACCTCTTCATCACCCAGTATTCTACTGGCCAA GTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCG CTGGAACCCGGAGATCCAGTACACTTCCAACCTATTACAAGTCTAATAAT GTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCA TTGGCACCAGATACCTGACTCGTAATCTGTAATTGCTTGAATCAATA AACCGTTTAATTCGTTTCAGTTGAACCTTTGGTCTCTGCG *VP2 start codon is in upper case single underlined.
SEQ ID NO: 39 AAV9 del VP2 (nt.)	GTCCTGTATTAGAGGTCACGTGAGTGTGTTTGGCAGATTTTGGCAGACCA TGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCA TTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCCGGGGTTTAC GAGATTGTGATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCCGGC ATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTG CCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTG ACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGT GTGAGTAAGGCCCCGGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGA GAGAGCTACTTCCACATGCACGTGCTCGTGGAACCACCGGGGTGAAA TCCATGGTTTTGGGACGTTTCTGAGTCAGATTTCGCAAAAACCTGATTC AGAGAATTTACCGCGGGATCGAGCCGACTTTGCCAACTGGTTTCGCGG TCACAAAGACCAGAAATGGCGCCGGAGGCGGGAACAAGGTGGTGGAT GAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCC AGTGGGCGTGGACTAATATGGAACAGTATTTAAGCGCCTGTTTGAATCT CACGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTGCA GACGCAGGAGCAGAACAAAGAGAATCAGAATCCCAATTCTGATGCGCC GGTGATCAGATCAAAAACCTCAGCCAGGTACATGGAGCTGGTCGGGTG GCTCGTGGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGG ACCAGGCCTCATACATCTCCTTCAATGCGGCCTCCAACTCGCGGTCCCA AATCAAGGCTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAA AACCGCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTTC CAGCAATCGGATTTATAAAATTTTGGAATAAACGGGTACGATCCCCA ATATGCGGCTTCCGTCTTTCTGGGATGGGCCACGAAAAAGTTTCGGCAA GAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGGAAGACCAA CATCGCGGAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAAA CTGGACCAATGAGAACTTTCCCTTCAACGACTGTGTGACAAAGATGGT GATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCTGGAGTCGG CCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCA AGTCCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACCTCCAACAC CAACATGTGCGCCGTGATTGACGGGAACCTCAACGACCTTCGAACACCA GCAGCCGTTGCAAGACCGGATGTTCAAATTTGAACTACCCGCCGTCTG GATCATGACTTTGGGAAGGTACCAAGCAGGAAGTCAAAGACTTTTTTC CGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTC AAAAAGGTGGAGCCAAGAAAAGACCCGCCCCCAGTGACGCAGATAT AAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGTC AGACGCGGAAGCTTCGATCAACTACGCAGACAGGTACCAAAACAAATG TTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGC GAGAGACTGAATCAGAATTCAAATATCTGCTTCACTACGGGTGTCAA GACTGTTTAGAGTGCTTTCCCGTGTCAGAATCTCAACCCGTTTCTGTGCG

Sequence identifier and name	Sequence
	TCAAAAAGGCGTATCAGAAACTGTGCTACATTCATCACATCATGGGAA AGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATGTGGACTTGG ATGACTGTGTTTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCCG ATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCCG CGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCA ACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTACAAATA CCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTCAACGCAGC AGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAA GGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTT CCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCG AGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGTT GAGGAAGCGGCTAAG<u>ACCG</u>CTCCTGGAAGAAGAGGCCTGTAGAGCA GTCTCCTCAGGAACCGACTCCTCCGCGGGTATTGGCAAATCGGGTGC ACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTGGCGACACAGA GTCAGTCCCAGACCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTC AGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAAGTGGC AGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTCGGGAAATTG GCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCACCAGCAC CCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTACAAGCAAAT CTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGG CTACAGCACCCCCTGGGGGTATTTTGACTTCAACAGATTCCACTGCCAC TTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGATT CGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTAGGTCAAAGAG GTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCTTACCAGC ACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTACGTGCTCG GGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCAT GATTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGT GGGTCGTTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCAATGCTA AGAACGGGTAAACAACCTCCAGTTCAGCTACGAGTTTGAGAACGTACCT TTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACTAATGAAT CCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATTAACGGTT CTGGACAGAATCAACAAACGCTAAAATTCAAGTGTGGCCGGACCCAGCA ACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGCTACCGAC AACACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTG CTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGAT GAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTT CTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAAGTGAAGA GACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAAT TAAAACTACTAACCCGGTAGCAACGGAGTCCTATGGACAAGTGGCCAC AAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAA ACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACC TGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTC ACCCTTCTCCGCTGATGGGAGGGTTTGAATGAAGCACCCGCTCCTCA GATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCAACGGCCTTC AACAAGGACAAGCTGAACCTTTTCATCACCCAGTATTCTACTGGCCAA GTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCG CTGGAACCCGGAGATCCAGTACACTTCCAATATTACAAGTCTAATAAT GTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCA TTGGCACCAAGATACCTGACTCGTAATCTGTAATTGCTTGTTAATCAATA AACCGTTTAATTCGTTTCAGTTGAACCTTGGTCTCTGCG *VP2 start codon is in upper case single underlined, and G was mutated to C in the VP2 start codon.

Sequence identifier and name	Sequence
SEQ ID NO: 40 AAV9 CAP VP1 (aa)	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLP GYKYLGPNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNH ADAEFQERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRP VEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDPQPIGEPPAAPSG VGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDREVTTSTRT WALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSR DWQRLINNNWGRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFT DSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYC LEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHQSOLDRLMNPLIDQYLYYL SKTINGSGQNQQTLKFSVAGPSNMAVQGRNYPGPSYRQQRVSTTVTQNN NSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQG TGRDNDVADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWV QNQGILPGMVWQDRDVYLQGPWAKIPHTDGNFHPSPMLGGFGMKHPPP QILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWN PEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL * Gly (G) 455, Asn (N) 498, Gln (Q) 588 are indicated in bold, underlined text.
SEQ ID NO: 41 AAV9 CAP VP2 (aa)	TAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDPQPI GEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWL GDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDF NRFHCHFSRPRDWQRLINNNWGRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQ AVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHQSOLDRLMNP LIDQYLYYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYPGPSYRQQR VSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPL SGSLIFGKQGTGRDNDVADKVMITNEEEIKTTNPVATESYGQVATNHQSA QAQAQTGWVQNQGILPGMVWQDRDVYLQGPWAKIPHTDGNFHPSPML GGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWEL QKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL * Corresponding Gly (G) 455, Asn (N) 498, Gln (Q) 588 with reference to SEQ ID NO:40 are indicated in bold, underlined text.
SEQ ID NO: 42 AAV9 CAP VP3 (aa)	MASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDREVTTSTRTWALP TYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSRPRDWQ RLINNNWGRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDY QLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFP SQMLRTGNNFQFSYEFENVPFHSSYAHQSOLDRLMNPLIDQYLYYLSKTIN GSGQNQQTLKFSVAGPSNMAVQGRNYPGPSYRQQRVSTTVTQNNNSEFA WPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRD NVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQ GILPGMVWQDRDVYLQGPWAKIPHTDGNFHPSPMLGGFGMKHPPPQILI KNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQ YTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL * Corresponding Gly (G) 455, Asn (N) 498, Gln (Q) 588 with reference to SEQ ID NO:40 are indicated in bold, underlined text.
SEQ ID NO: 43 Rep/Cap AAV9- Gly455 Antibody fragment (nt.)	GTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGCACACCA TGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCA TTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCCGGGGTTTTAC GAGATTGTGATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCCGGC ATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTG CCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTG ACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGT GTGAGTAAGGCCCGGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGA

Sequence identifier and name	Sequence
	GAGAGCTACTTCCACATGCACGTGCTCGTGGAACCACCGGGGTGAAA TCCATGGTTTTGGGACGTTTCCTGAGTCAGATTTCGCGAAAAACTGATTC AGAGAATTTACCGCGGGATCGAGCCGACTTTGCCAAACTGGTTCGCGG TCACAAAGACCAGAAATGGCGCCGGAGGCGGGAACAAGGTGGTGGAT GAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCC AGTGGGCGTGGAATAATATGGAACAGTATTTAAGCGCCTGTTTGAATCT CACGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTCGCA GACGCAGGAGCAGAACAAAGAGAATCAGAATCCCAATTCTGATGCGCC GGTGATCAGATCAAAAACCTCAGCCAGGTACATGGAGCTGGTCCGGTG GCTCGTGGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGG ACCAGGCCTCATACATCTCCTTCAATGCGGCCTCCAACTCGCGGTCCCA AATCAAGGCTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAA AACCGCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTTC CAGCAATCGGATTTATAAAATTTTGGAACTAAACGGGTACGATCCCCA ATATGCGGCTTCCGTCTTTCTGGGATGGGCCACGAAAAAGTTCGGCAA GAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGGAAGACCAA CATCGCGGAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAAA CTGGACCAATGAGAACTTCCCTTCAACGACTGTGTCGACAAGATGGT GATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCTGTGGAGTCGG CCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCA AGTCCTCGGCCCAGATAGACCCGACTCCCGTGATCGTCACCTCCAACAC CAACATGTGCGCCGTGATTGACGGGAAGTCAACGACCTTCGAACACCA GCAGCCGTTGCAAGACCGGATGTTCAAATTTGAACTACCCGCGCTGTG GATCATGACTTTGGGAAGGTACCAAGCAGGAAGTCAAGACTTTTTTC CGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTC AAAAAGGGTGGAGCCAAGAAAAGACCCGCCCCCAGTGACGCAGATAT AAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGTC AGACGCGGAAGCTTCGATCAACTACGCAGACAGGTACCAAAACAAATG TTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGC GAGAGACTGAATCAGAATTCAAATATCTGCTTCACTACGGTGTCAAA GACTGTTTAGAGTGCTTTCCTGTGTCAGAATCTCAACCCGTTTCTGTG TCAAAAAGGCGTATCAGAACTGTGCTACATTCATCAGATCATGGGAA AGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATGTGGACTTGG ATGACTGTGTTTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCCG ATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTTCG CGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCA ACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATA CCTTGGACCCGGCAACGGAAGTTCGACAAGGGGGAGCCGGTCAACGCAGC AGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAA GGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTT CCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCG AGCAGTCTTCCAGGCCAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTT GAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCA GTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGC ACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTGGCGACACAGA GTCAGTCCCAGACCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTC AGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAGTGGC AGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTCGGGAAATTG GCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCACCAGCAC CCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTACAAGCAAAT CTCAACAGCACATCTGGAGGATCTTCAAATGACAAGCAGCTACTTCGG CTACAGCACCCCCTGGGGGTATTTTGAATCAACAGATTCCACTGCCAC

Sequence identifier and name	Sequence
	TTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTCATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCAAAATGCTAAGAACGGGTAAACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATTAAACGGTTCTGGAGGTTGTCCCCCTGGCTACAAGAGCGGGCTGGAGTGTCTCCCGGGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGATACGGGGGATACGGCGCAGACTCGGGAGTGGGATCGCAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCCTTTGTCTGATCTTTAATTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACC CGGTAGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTC CGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCCAAGGACCCATTGGGCCAAAATTCTCACACGGACGGCAACTTTCACCCCTTCTCCGCTGATGGGAGGGTTTGGAAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAA CACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACCTTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATC GAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACATTACAAGTCTAATAATGTTGAATTTGCTGTT AATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCCAGATACCTGACTCGTAATCTGTAATTGCTTGTTAATCAATAAACCGTTTAATTCGTTTCAGTTGAACTTTGGTCTCTGCG *Antibody fragment (Knob domain B1 as an example) is indicated in bold, underlined text.
SEQ ID NO: 44 Rep/Cap AAV9-Asn498 Antibody fragment (nt.)	GTCCTGTATTAGAGGTCACGTGAGTGTGTTTGGCGACATTTTGGCGACACCA TGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCA TTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCCGGGGTTTTAC GAGATTGTGATTAAGTCTCCCGAGCGACCTTGACGAGCATCTGCCCGGC ATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTG CCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTG ACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGT GTGAGTAAGGCCCCGGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGA GAGAGCTACTTCCACATGCACGTGCTCGTGGAACCACCGGGGTGAAA TCCATGGTTTTGGGACGTTTCCTGAGTCAGATTCGCGAAAAAAGTATTC AGAGAATTTACCGCGGGATCGAGCCGACTTTGCCAAACTGGTTCGCGG TCACAAAGACCAGAAAATGGCGCCGGAGGCGGGAACAAGGTGGTGGAT GAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCC AGTGGGCGTGACTAATATGGAACAGTATTTAAGCGCCTGTTTGAATCT CACGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTCGCA GACGCAGGAGCAGAACAAGAGAATCAGAATCCCAATTCTGATGCGCC GGTGATCAGATCAAAAACCTTCAGCCAGGTACATGGAGCTGGTCCGGGTG GCTCGTGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGG

Sequence identifier and name	Sequence
	ACCAGGCCTCATACATCTCCTTCAATGCGGCCTCCAACCTCGCGGTCCCA AATCAAGGCTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAA AACCGCCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTTC CAGCAATCGGATTTATAAAATTTTGGAACTAAACGGGTACGATCCCCA ATATGCGGCTTCCGTCTTTCTGGGATGGGCCACGAAAAAGTTTCGGCAA GAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGGAAGACCAA CATCGCGGAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAAA CTGGACCAATGAGAACTTTCCCTTCAACGACTGTGTGACAAAGATGGT GATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCGTGGAGTCGG CCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCA AGTCCTCGGCCCAGATAGACCCGACTCCCGTGATCGTCACTCCAACAC CAACATGTGCGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCA GCAGCCGTTGCAAGACCGGATGTTCAAATTTGAACTACCCGCCGTCTG GATCATGACTTTGGGAAGGTCACCAAGCAGGAAGTCAAAGACTTTTTTC CGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTC AAAAAAGGTGGAGCCAAGAAAAGACCCGCCCCCAGTGACGCAGATAT AAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGTC AGACGCGGAAGCTTCGATCAACTACGCAGACAGGTACCAAAACAAATG TTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGC GAGAGACTGAATCAGAATTCAAATATCTGCTTCACTCACGGTGTCAAA GACTGTTTAGAGTGCTTTCCCGTGTCAGAATCTCAACCCGTTTCTGTCTG TCAAAAAGGCGTATCAGAACTGTGCTACATTCATCACATCATGGGAA AGGTGCCAGACGCTTGCACTGCTTGCAGCTTGGTCAATGTGGACTTGG ATGACTGTGTTTCTGAACAATAAATGACTTAAACCAAGTATGGCTGCCG ATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCTG CGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCA ACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTACAAATA CCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTCAACGCAGC AGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAA GGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTT CCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCG AGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTT GAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCA GTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGC ACAGCCCGCTAAAAAGAGACTCAATTTCCGTGAGTGGCGACACAGA GTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTC AGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAAGTGGC AGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTCTCGGGAAATTG GCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCAACCAGCAC CCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTACAAGCAAAT CTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGG CTACAGCACCCCCTGGGGGTATTTTGACTTCAACAGATTCCACTGCCAC TTCTCACACGTGACTGGCAGCGACTCATCAACAACAACCTGGGGATTTC CGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAG GTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCTTACCAGC ACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTACGTGCTCG GGTCGGCTCACGAGGGCTGCCTCCCGCGGTTCCAGCGGACGTTTTTCAT GATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGT GGGTCGTTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGAAATGCTA AGAACGGGTAACAACCTTCAGTTCAGCTACGAGTTTGAGAACGTACCT TTCCATAGCAGCTACGCTCACAGCCAAGCCTGGACCGACTAATGAAT CCACTCATCGACCAATACTTGTAATCTCTCAAAGACTATTAACGGTT

Sequence identifier and name	Sequence
	CTGGACAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGACCCAGCA ACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGCTACCGAC AACAACGTGTCTCAACCACTGTGACTCAAAACAACAAC<u>GGTTGTCCCC</u> <u>CTGGCTACAAGAGCGGCGTGGACTGCTCCCCGGGATCGGAGTGCA</u> <u>AATGGGGATGCTACGCCGTGGATGGACGGAGATACGGGGGATACG</u> <u>GCGCAGACTCGGGAGTGGGATCG</u>GAGCGAATTGCTTGGCCTGGAGCT TCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTG CTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGG ATCTTTAATTTTGGCAAACAAGGAAGTGGAGAGACAACGTGGATGC GGACAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCC GGTAGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACAGAGTGC CCAAGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCC GGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTG GGCCAAAATTCTCACACGGACGGCAACTTTCACCTTCTCCGCTGATG GGAGGGTTTGGAAATGAAGCACCCGCCTCCTCAGATCCTCATAAAAAC ACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTG AACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCG AGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATC CAGTACACTTCCAATACTATTACAAGTCTAATAATGTTGAATTTGCTGTTA ATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACC TGACTCGTAATCTGTAATTGCTTGTTAATCAATAAACCGTTTAATTTCGTT TCAGTTGAACTTTGGTCTCTGCG *Knob domain B1 as an example of antibody fragment is indicated in bold, underlined text.
SEQ ID NO: 45 Rep/Cap AAV9 - Gln588 Antibody fragment	GTCCTGTATTAGAGGTCACGTGAGTGTGTTTGGCAGATTTTGGCAGACCA TGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCA TTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCCGGGGTTTAC GAGATTGTGATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCCGGC ATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTG CCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTG ACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGT GTGAGTAAGGCCCGGAGGCCCTTTTCTTTGTGCAATTTGAGAAGGGA GAGAGCTACTTCCACATGCACGTGCTCGTGGAACCACCGGGGTGAAA TCCATGGTTTTGGGACGTTTCTGAGTCAGATTCGCGAAAAACTGATTC AGAGAATTTACCGCGGGATCGAGCCGACTTTGCCAACTGGTTCGCGG TCACAAAGACCAGAAATGGCGCCGAGGCGGGAACAAGGTGGTGGAT GAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCC AGTGGGCGTGGACTAATATGGAACAGTATTTAAGCGCCTGTTTGAATCT CACGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTCGCA GACGCAGGAGCAGAAACAAAGAGAATCAGAATCCCAATTCTGATGCGCC GGTGATCAGATCAAAAACCTCAGCCAGGTACATGGAGCTGGTCGGGTG GCTCGTGGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGG ACCAGGCCTCATACTCTCCTTCAATGCGGCCTCCAACCTCGCGGTCCCA AATCAAGGCTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAA AACCGCCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTT CAGCAATCGGATTTATAAAATTTTGGAACTAAACGGGTACGATCCCCA ATATGCGGCTTCCGTCTTTCTGGGATGGGCCACGAAAAAGTTTCGGCAA GAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGGAAGACCAA CATCGCGGAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAA CTGGACCAATGAGAACTTTCCCTTCAACGACTGTGTCGACAAGATGGT GATCTGGTGGGAGGAGGGAAGATGACCGCCAAGGTCTGAGTGGTGG CCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCA

Sequence identifier and name	Sequence
	AGTCCTCGGCCCAGATAGACCCGACTCCCGTGATCGTCACCTCCAACAC CAACATGTGCGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCA GCAGCCGTTGCAAGACCGGATGTTCAAATTTGAACTCACCCGCCGTCTG GATCATGACTTTGGGAAGGTCACCAAGCAGGAAGTCAAAGACTTTTTTC CGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTC AAAAAGGGTGGAGCCAAGAAAAGACCCGCCCCCAGTGACGCAGATAT AAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGTC AGACGCGGAAGCTTCGATCAACTACGCAGACAGGTACCAAAACAAATG TTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGC GAGAGACTGAATCAGAATTCAAATATCTGCTTCACTCACGGTGTCAA GACTGTTTAGAGTGCTTTCCCGTGTCAGAATCTCAACCCGTTTCTGTCG TCAAAAAGGCGTATCAGAAACTGTGCTACATTCATCACATCATGGGAA AGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATGTGGACTTGG ATGACTGTGTTTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCCG ATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCCG CGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCA ACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTACAAATA CCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTCAACGCAGC AGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAA GGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTT CCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCG AGCAGTCTTCCAGGCCAAAAGAGGGCTTCTTGAACCTCTTGGTCTGGTT GAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCA GTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGC ACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTGGCGACACAGA GTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTC AGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAGTGGC AGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTCGGGAAATTG GCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCACCAGCAC CCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTACAAGCAAAT CTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGG CTACAGCACCCCCTGGGGGTATTTTGACTTCAACAGATTCCACTGCCAC TTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGATTTC CGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAG GTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCTTACCAGC ACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGTGCTCG GGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCAT GATTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGT GGGTCTGTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCAAAATGCTA AGAACGGGTAAACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTACCT TTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACTAATGAAT CCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATTAACGGTT CTGGACAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGACCCAGCA ACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGCTACCGAC AACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTG CTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGAT GAATCCTGGACCTGCTATGGCCAGCCACAAGAAGGAGAGGACCGTTT CTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAAGTGAAGA GACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAAT TAAACTACTAACCCGGTAGCAACGGAGTCCTATGGACAAGTGGCCAC AAACCACAGAGTGCCCAA GGTTGTCCCCCTGGTACAAGAGCGGC GTGGA CTGCTCCCCGGGATCGGAGTGCAAATGGGGATGCTACGCC

Sequence identifier and name	Sequence
	<u>GTGGATGGACGGAGATACGGGGGATACGGGCGCAGACTCGGGAGTGGGATCGGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATAC</u> <u>TTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCA</u> <u>TTTGGGCCAAAATTCCTCACACGGACGGCAACTTTCACCCCTTCTCCGCT</u> <u>GATGGGAGGGTTTGGAATGAAGCACCCGCTCCTCAGATCCTCATCAA</u> <u>AAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAA</u> <u>GCTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAG</u> <u>ATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGA</u> <u>GATCCAGTACACTTCCAATAATTACAAGTCTAATAATGTTGAATTTGCT</u> <u>GTTAATACTGAAGGTGTATATAGTGAACCCCGCCCATTTGGCACCAGA</u> <u>TACCTGACTCGTAATCTGTAATTGCTTGTTAATCAATAAACCGTTTAAT</u> <u>TCGTTTCAGTTGAACTTTGGTCTCTGCG</u> *Knob domain B1 as an example of antibody fragment is indicated in bold, underlined text.
SEQ ID NO: 46 AAV9 Cap CMV-Knob-VP2	cgttacataactacgtaaatggcccgctggctgacgcccacgacccccgccattgacgtcaataatgacgtat gttcccatagtaacccaatagggaactttccattgacgtcaatgggtggagtatttacggtaaactgccacttgccagt acatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcattatgccagt acatgaccttatgggaactttctacttggcagtcacatctacgtattagtcacgtattaccatgggtgatcggtttggcag tacatcaatggcgctggatagcgggttgactcacggggaattccaagtctccacccattgacgtcaatgggagttgttt tggcaccaaatcaacgggaactttccaaaatgctgtaacaactccgccccattgacgcaaatggcggtaggcggtgta cgggtgggaggtctatataagcagagctggttagtgaaaccgtcagatcCTCGAGGCCGCCACCATGT <u>CATGCCCTGATGGATTCTCTTATCGTTTCATGGGACGACTTCTGCTG</u> <u>TCCAATGGTGGGAGATGTCTCGCACCCAGAAACACGACCGCTCCT</u> <u>GGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCC</u> <u>GCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAAAAGAGACTCAAT</u> <u>TTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCTCAACCAATC</u> <u>GGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTT</u> <u>CAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGATGGA</u> <u>GTGGGTAGTTCCTCGGGAATTTGGCATTGCGATTCCCAATGGCTGGGG</u> <u>GACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCACCTAC</u> <u>AACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTGGAGGATCT</u> <u>TCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGGGGGTATTTTG</u> <u>ACTTCAACAGATTCCAAGTCCACTTCTCACCACGTGACTGGCAGCGACT</u> <u>CATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAACTTCAAGCT</u> <u>CTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGAC</u> <u>CATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGA</u> <u>CTATCAGCTCCCGTACGTGCTCGGGTTCGGCTCACGAGGGCTGCCTCCCG</u> <u>CCGTTCCCAGCGGACGTTTTTCATGATTTCCTCAGTACGGGTATCTGACGC</u> <u>TAAATGATGGAAGCCAGGCCGTGGGTTCGTTTCGTCCTTTTACTGCCTGGA</u> <u>ATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAACCTTCCAGTTCAGC</u> <u>TACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCCAA</u> <u>AGCCTGGACCGACTAATGAATCCAATCATCGACCAATACTTGTACTATC</u> <u>TCTCAAAGACTATTAACGGTTCTGGACAGAATCAACAAACGCTAAAAAT</u> <u>TCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAACTACA</u> <u>TACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTC</u> <u>AAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCT</u> <u>CAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCA</u> <u>CAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTTT</u> <u>GGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAAGTCAT</u> <u>GATAACCAACGAAGAAGAAATTAATACTACTAACCCGGTAGCAACGG</u> <u>AGTCCTATGGACAAGTGGCCACAAACCACAGAGTGCCCAAGCACAGG</u> <u>CGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTT</u>

Sequence identifier and name	Sequence
	<u>GGCAGGACAGAGATGTGTACCTGCAAGGACCCATTG</u><u>GGGCCAAAATTC</u> <u>CTCACACGGACGGCAACTTTACCCCTTCTCCGCTGATGGGAGGGTTTGG</u> <u>AATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCT</u> <u>GCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATC</u> <u>ACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTG</u> <u>CAGAAAGGAAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCC</u> <u>AACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTG</u> <u>TATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCT</u> <u>GTAAGAAATTCTCTAG</u><i>atccagacatgataagatacattgatgagtttgacaaaccacaactag</i> <i>aatgcagtgaataaatgctttatttggaaatttggatgctattgctttatttgaaccattataagctgcaataaaca</i> <i>agtt</i> *CMV promoter is in lower case text, knob domain (C1 as an example) is in bold underlined text, AAV9 VP2 is in uppercase double underlined text and SV40 polyA is in italic lowercase text.
SEQ ID NO: 47 AAV9 Cap VP2 antibody fragment	<u>ATGTCATGCCCTGATGGATTCTCTTATCGTTCATGGGACGACTTCT</u> <u>GCTGTCCAATGGTGGGGAGATGTCTCGCACCCAGAAACACGACCG</u> CTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGACT CCTCCGCGGGTATTGGCAAATCGGGTGACAGCCCGCTAAAAAGAGAC TCAATTTCCGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCCTCAAC CAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAAT GGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGA TGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGGCTG GGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCACC TACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTGGAGGA TCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGGGGGTATT TTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCG ACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAACTTCAA GCTCTTCAACATTAGGTCAAAGAGGTTACGGACAACAATGGAGTCAA GACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTC AGACTATCAGTCCCGTACGTGCTCGGGTCGGCTCACGAGGGCTGCCTC CCGCCGTTCCAGCGGACGTTTTTCATGATTCCTCAGTACGGGTATCTGA CGCTTAATGATGGAAGCCAGGCCGTGGGTGCTTCGTCCTTTTACTGCCT GGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAACCTCCAGTT CAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGC CAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTTGTACT ATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACAAACGCTAA AATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAACT ACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGA CTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGC TCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGC CACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTT TTGGCAAACAAGGAACCTGGAAGAGACAACGTGGATGCGGACAAAGTC ATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGTAGCAACG GAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCAAGCACAG GCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTT TGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTGCGGCCAAAATT CCTCACACGGACGGCAACTTTACCCCTTCTCCGCTGATGGGAGGGTTTG GAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTAC CTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCAT CACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCT GCAGAAGGAAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTT CCAACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGG

Sequence identifier and name	Sequence
	TGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA Knob domain C1 as an example of an antibody fragment is in bold underlined text
SEQ ID NO: 48 AAV9 Cap VP1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTG AAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGC CTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTC CACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACT GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTACGT CAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGAC GTTTTTCATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTCGTTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATT AACGGTTCTGGACAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGA CCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGC TACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGC GAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATA GCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGG ACCGTTTCTTCTCTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACT GGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGA AGAAATTAAAACTACTAACCCGGTAGCAACGGAGTCTATGGACAAGT GGCCACAAACCACAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGG TTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATG TGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCA ACTTTCACCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCC TCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCCAACG GCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTG GCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAAACAGC AAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACCTATTACAAGTCT AATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCC GCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA
SEQ ID NO: 49 AAV9 Cap VP2	ACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCG GACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAAAAG

Sequence identifier and name	Sequence
	AGACTCAATTTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCT CAACCAATCGGAGAACCTCCCCGAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAA TGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTG CCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGG GGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACT GGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGG GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTGTCCTTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAAC TTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGTACTATCTCTCAAAGACTATTAACGGTCTGAGACAGAATCAACA AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTCTTTGTCTGGATC TTTAATTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCA AGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGG TATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGC CAAAATTCCTCACACGGACGGCAACTTTCACCCCTTCTCCGCTGATGGGA GGGTTTGAATGAAGCACCCGCTCCTCAGATCCTCATCAAAAACACA CCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAAC TCTTTCATACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGT GGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAATAATTACAAGTCTAATAATGTTGAATTTGCTGTTAATA CTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA
SEQ ID NO: 50 AAV9 Cap VP3	ATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCC GATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAATGG CTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCC ACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTGGA GGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGGGGG TATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGC AGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAACT TCAAGCTCTTCAACATTAGGTCAAAGAGGTTACGGACAACAATGGAG TCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGG ACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGAGGGCT GCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGGGT TCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTGTCCTTTTAC TGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAATTCC AGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCA CAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTT GTACTATCTCTCAAAGACTATTAACGGTCTGAGACAGAATCAACAAC

Sequence identifier and name	Sequence
	GCTAAAATTCAGTGTGGCCGACCCAGCAACATGGCTGTCCAGGGAAG AAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCAC TGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCT TGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGG CCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCTTT AATTTTTGGCAAACAAGGAAGTGGGAAGAGACAACGTGGATGCGGACA AAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGTAG CAACGGAGTCTTATGGACAAGTGGCCACAAACCACCAGAGTGCCCAAG CACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTA TGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCA AAATTCTCACACGGACGGCAACTTTCACCTTCTCCGCTGATGGGAGG GTTTGAATGAAGCACCCGCTCTCTCAGATCTCATCAAAAACACACCT GTACCTGCGGATCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCT TTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGG GAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTA CACTTCCAATAATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACT GAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGACT CGTAATCTGTAA
SEQ ID NO: 51 AAV9 Cap VP1- Gly455 antibody fragment	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTG AAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGC CTGTAGAGCAGTCTCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGCACAGCCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTC CACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAAC GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTTCAGGT CAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGAC GTTTTCATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTCTGTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCAATCATCGACCAATACTTGTACTATCTCTCAAAGACTATT AACGGTTCTGGAGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACT <u>GCTCCCCGGGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATG</u> <u>GACGGAGATACGGGGGATACGGCGCAGACTCGGGAGTGGGATCGC</u> AGAATCAACAAACGCTAAAATTCAGTGTGGCCGACCCAGCAACATGG CTGTCCAGGGAAGAAACTACATACCTGGACCCAGCTACCGACAACAAC GTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGC

Sequence identifier and name	Sequence
	CTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCC TGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCC TTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACTGGAAGAGACAAC GTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAC TACTAACCCGGTAGCAACGGAGTCCTATGGACAAGTGGCCACAAACCA CCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAG GAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAG GACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTCACCCCTC TCCGCTGATGGGAGGGTTTGAATGAAGCACCCGCCTCCTCAGATCCTC ATCAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAG GACAAGCTGAACCTCTTCATCACCCAGTATTCTACTGGCCAAAGTCAGCG TGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAAC CCGGAGATCCAGTACACTTCCAACCTATTACAAGTCTAATAATGTTGAAT TTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCAC CAGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 52 AAV9 Cap VP2 - Gly455 antibody fragment	ACGGCTCCTGGAAGAAGAGGCGCTGTAGAGCAGTCTCCTCAGGAACCG GACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCCGCTAAAAAG AGACTCAATTTCCGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCCT CAACCAATCGGAGAACCTCCCCGAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAA TGGCTGGGGGACAGAGTCATCACCAACCAGCACCCGAACCTGGGCCCTG CCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGG GGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACTGACT GGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCAGCGGACGTTTTATGATTCTCAGTACGG GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTGTCCTTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAAC TTCCAGTTCAGTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGTACTATCTCTCAAAGACTATTAACGGTTCTGGAGGTTGTCCCC <u>CTGGCTACAAGAGCGCGTGGACTGCTCCCCGGGATCGGAGTGCA</u> <u>AATGGGGATGCTACGCCGTGGATGGACGGAGATACGGGGGATACG</u> <u>GCGCAGACTCGGGAGTGGGATCGCAGAATCAACAACGCTAAAATT</u> CAGTGTGGCCGACCCAGCAACATGGCTGTCCAGGGAAGAACTACAT ACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCA AAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTC AATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCAC AAAGAAGGAGAGGACCGTTTCTTCTTTGTCTGGATCTTTAATTTTTG GCAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATG ATAACCAACGAAGAAGAAATTAACCTACTAACCCGGTAGCAACGGA GTCCTATGGACAAGTGGCCACAAACCACAGAGTGCCCAAGCACAGGC GCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTG GCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCC TCACACGGACGGCAACTTTTACCCCTTCTCCGCTGATGGGAGGGTTTGA ATGAAGCACCCGCTCCTCAGATCCTCATCAAAAACACACCTGTACCTG

Sequence identifier and name	Sequence
	CGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCAC CCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCA GAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCA ACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGT ATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG TAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 53 AAV9 Cap VP3- Gly455 antibody fragment	ATGGCTTCAGGTGGTGGCGCACCAAGTGGCAGACAATAACGAAGGTGCC GATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAATGG CTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGGCCCTGCC ACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTGGA GGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGGGGG TATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGC AGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAACT TCAAGCTCTTCAACATTCAAGTCAAAGAGGTTACGGACAACAATGGAG TCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGG ACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGAGGGCT GCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGGGTA TCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTGTCCTTTTAC TGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAACCTTC AGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCA CAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTT GTACTATCTCTCAAAGACTATTAACGGTTCTGGAG<u>AGTTGTCCCCCTGG</u> <u>CTACAAGAGCGGCGTGGACTGCTCCCCGGGATCGGAGTGCAAATG</u> <u>GGGATGCTACGCCGTGGATGGACGGAGATACGGGGGATACGGCGC</u> <u>AGACTCGGGAGTGGGATCGC</u>CAGAATCAACAACGCTAAAATTCACT GTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCT GGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAAC AACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCTGGGCTCTCAATG GACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAG AAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTGGCAA ACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAA CCAACGAAGAAGAAATTAATACTACTAACCCTGGTAGCAACGGAGTCCT ATGGACAAGTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAG ACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAG GACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCAC ACGGACGGCAACTTTCACCTTCTCCGCTGATGGGAGGGTTTGGAAATG AAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCG GATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCC AGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGA AGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAAC ATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATA TAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 54 AAV9 Cap VP1- Asn498 antibody fragment	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTG AAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGCGGCCCTCGAGCACGACAAGGCCTACGACC

Sequence identifier and name	Sequence
	AGCAGCTCAAGGCCGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGC CTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGACAGCCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTC CACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACA GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGT CAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCGGTTCACGCGGAC GTTTTCATGATTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTGCTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATT AACGGTTCTGGACAGAATCAACAAACGCTAAAATTTCAGTGTGGCCGGA CCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGC TACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACGGT <u>TGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCGGGATCG</u> <u>GAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGATACGGG</u> <u>GGATACGGCGCAGACTCGGGAGTGGGATCG</u> AGCGAATTTGCTTGGC CTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCC TGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCC TTTGTCTGGATCTTTAATTTTGGCAAACAAGGAAGTGGAAAGAGACAAC GTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAC TACTAACCCGGTAGCAACGGAGTCCTATGGACAAGTGGCCACAAACCA CCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAG GAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAG GACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTCACCCCTC TCCGCTGATGGGAGGGTTTGGAAATGAAGCACCCGCCTCCTCAGATCCTC ATCAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAG GACAAGCTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCG TGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAAC CCGGAGATCCAGTACACTTCCAACCTATTACAAGTCTAATAATGTTGAAT TTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCAC CAGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 55 AAV9 Cap VP2 Asn498- antibody fragment	ACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCG GACTCCTCCGCGGGTATTGGCAAATCGGGTGACAGCCCCGCTAAAAAG AGACTCAATTTCCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCT CAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAA

Sequence identifier and name	Sequence
	TGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTG CCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCTGG GGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACT GGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGTCTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCAGCGGACGTTTTATGATTCTCAGTACGG GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCTGTCCTTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAAC TTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGTACTATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACA AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAAC<u>GGTTGTCCCCCTGGCTACAAGAGC</u> <u>GGCGTGGACTGCTCCCCGGGATCGGAGTGCAAATGGGGATGCTAC</u> <u>CCCGTGGATGGACGGAGATACGGGGGATACGGCCGACACTCGGG</u> <u>AGTGGGATCGAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCT</u> CAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCA CAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTT GGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAGTCAT GATAACCAACGAAGAAGAAATTAATACTACTAACCCGTAGCAACGG AGTCCTATGGACAAGTGGCCACAAACCACAGAGTGCCCAAGCACAGG CGCAGACCGGCTGGGTTCAAACCAAGGAATACTTCCGGGTATGGTTT GGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTC CTCACACGGACGGCAACTTTACCCCTTCTCCGCTGATGGGAGGGTTTGG AATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCT GCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATC ACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTG CAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCC AACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTG TATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCT GTAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 56 AAV9 Cap VP3 - Asn498 antibody fragment	ATGGCTTCAGGTGGTGGCGCACCAAGTGGCAGACAATAACGAAGGTGCC GATGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGG CTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCC ACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTGGA GGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCTGGGGG TATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGC AGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAACT TCAAGCTCTTCAACATTCAAGTCAAAGAGGTTACGGACAACAATGGAG TCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGG ACTCAGACTATCAGTCTCCCGTACGTGCTCGGGTCGGCTCACGAGGGCT GCCTCCCGCCGTTCCAGCGGACGTTTTATGATTCTCAGTACGGGTA TCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCTGCTCTTTAC TGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAATTCC AGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCA CAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTT

Sequence identifier and name	Sequence
	GTACTATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACAAAC GCTAAAATTTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAG AAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCAC TGTGACTCAAAACAACAAC<u>GGTTGTCCCCCTGGCTACAAGAGCGGC</u> <u>GTGGACTGCTCCCCGGGATCGGAGTGCAATGGGGATGCTACGCC</u> <u>GTGGATGGACGGAGATACGGGGGATACGGCGCAGACTCGGGAGT</u> <u>GGGATCG</u>AGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAAT GGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAA GAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTTGGCA AACAAGGAAGTGAAGAGACAACGTGGATGCGGACAAAGTCATGATA ACCAACGAAGAAGAAATTAATACTAACCCTAGCAACGGAGTCC TATGGACAAGTGGCCACAAACCACAGAGTGCCCAAGCACAGGCGCA GACCGGCTGGGTTCAAAACCAAGGAATACTCCGGGTATGGTTTGGCA GGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCA CACGGACGGCAACTTTCACCCTTCTCCGCTGATGGGAGGGTTTGAATG AAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCG GATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCC AGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGA AGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAAC ATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATA TAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 57 AAV9 Cap VP1 - Gln588 antibody fragment	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTG AAGGAATTTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAGAGGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGC CTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGACAGCCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCAATCAC CACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTC CACTGCCACTTCTCACACGTGACTGGCAGCGACTCATCAACAACAAC GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTACAGGT CAAAGAGGTTACGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGAC GTTTTCATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTGCTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCACTCATCGACCAATACTTGTAATCTCTCAAAGACTATT AACGGTCTGGACAGAATCAACAAACGCTAAAATTTCAGTGTGGCCGGA

Sequence identifier and name	Sequence
	CCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGC TACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGC GAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATA GCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGG ACCGTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACT GGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGA AGAAATTAATACTACTAACCCGGTAGCAACGGAGTCCTATGGACAAGT GGCCACAAACCACCAGAGTGCCCAAGGTTGTCCCCCTGGCTACAAGA <u>GCGGCGTGGACTGCTCCCCGGGATCGGAGTGCAATGGGGATGCT</u> <u>ACGCCGTGGATGGACGGAGATACGGGGGATACGGCGCAGACTCGG</u> <u>GAGTGGGATCGGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGA</u> ATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGA CCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTACCCCTTCTC CGCTGATGGGAGGGTTTGAATGAAGCACC CGCCTCCTCAGATCCTCA TCAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGG ACAAGCTGAACCTTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGT GGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACC CGGAGATCCAGTACACTTCCAAC TATTACAAGTCTAATAATGTTGAATT TGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACC AGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example of an antibody fragment is in uppercase bold single underlined
SEQ ID NO: 58 AAV9 Cap VP2 - Gln588 antibody fragment	ACGGCTCCTGGAAGAAGAGGGCCTGTAGAGCAGTCTCCTCAGGAACCG GACTCCTCCGCGGTATTGGCAAATCGGGTGACAGCCCGCTAAAAAG AGACTCAATTTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCT CAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAA TGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTG CCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGG GGGTATTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACT GGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGTCTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGG GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCTGCTCTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAAC TTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGACTATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACA AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATC TTTAATTTTTGGCAAACAAGGAAGTGAAGAGACAACGTGGATGCGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCA <u>AGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCGGG</u> <u>ATCGGAGTGCAATGGGGATGCTACGCCGTGGATGGACGGAGATA</u>

Sequence identifier and name	Sequence
	<u>CGGGGGATACGGCGCAGACTCGGGAGTGGGATCGGCACAGGCGCA</u> GACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCA GGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCA CACGGACGGCAACTTTCACCCTTCTCCGCTGATGGGAGGGTTTGAATG AAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCG GATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCC AGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGA AGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAAC ATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATA TAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example is in uppercase bold single underlined
SEQ ID NO: 59 AAV9 Cap VP3 - Gln588 antibody fragment	ATGGCTTCAGGTGGTGGCGCACCAAGTGGCAGACAATAACGAAGGTGCC GATGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGG CTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCC ACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTGGA GGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGGGGG TATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGC AGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAACT TCAAGCTCTTCAACATTCAAGGTCAAAGAGGTTACGGACAACAATGGAG TCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGG ACTCAGACTATCAGCTCCCGTACGTGCTCGGGTGGCTCACGAGGGCT GCCTCCCGCCGTTCCCAGCGGACGTTTTTCATGATTCTCAGTACGGGT TCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTGCTTCGTCTTTTAC TGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAACCTCC AGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCA CAGCCAAAGCCTGGACCGACTAATGAATCCAATCATCGACCAATACTT GTACTATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACAAAC GCTAAAATTCAAGTGTGGCCGACCCAGCAACATGGCTGTCCAGGGAAG AAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCAC TGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCT TGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGG CCAGCCACAAAGAAGGAGAGGACCGTTTCTTCTTGTCTGGATCTTT AATTTTGGCAAACAAGGAACCTGGAAGAGACAACGTGGATGCGGACA AAGTCATGATAACCAACGAAGAAGAAATTAATACTAATCCCGGTAG CAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCAAG <u>GTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCGGGAT</u> <u>CGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGATACG</u> <u>GGGGATACGGCGCAGACTCGGGAGTGGGATCGGCACAGGCGCAGA</u> CCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGG ACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACA CGGACGGCAACTTTCACCCTTCTCCGCTGATGGGAGGGTTTGAATGA AGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGG ATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCA GTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAA GGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAAC TTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATAT AGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *Knob domain B1 as an example is in uppercase bold single underlined
SEQ ID NO: 60 AAV1 CAP	MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLP GYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHA DAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPV

Sequence identifier and name	Sequence
	EQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPQPPLGEPPATPAA VGPTTMASSGGAPMADNNEGADGVGNASGNWHCDSTWLGDREVTTSTR TWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSPR DWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFS DSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYC LEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQYLYYL NRTQNQSGSAQNKDLLFSRSGSPAGMSVQPKNWLPGPCYRQQRVSKTKTD NNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDEKFFPMSGVMIFGK ESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDV HAMGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKNPPP QILIKNTPVPANPPAEFSATKFAFITQYSTGQVSVEIEWELQKENS KRWN EVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL
SEQ ID NO: 61 AAV2 CAP	MAADGYLPDWLEDTLSEGIQWWKLKPGPPPKPAERHKDDSRGLVLP YKYLGPFGNLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD AEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVE HSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPPLGQPPAAPSG LGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTR TWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRD WQRLINNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANNLSTVQVFTD SEYQLPYVLGSAHQGCLPPFPADVFMVMPQYGYLTLNNGSQAVGRSSFYCL EYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLS RTNTPSGTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN NSEYSWTGATKYHLNGRDSL VNPGPAMASHKDDEEKFFPQSGVLIFGKQG SEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNT QGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIL IKNTPVPANPSTTFSAAKFAFITQYSTGQVSVEIEWELQKENS KRWN PEIQYTSNYNKS VNVDFTVDNTNGVYSEPRPIGTRYLTRNL
SEQ ID NO: 62 AAV3 CAP	MAADGYLPDWLEDNLSEGIQWWALKPGVPQPKANQQHQDNRRGLVLP GYKYLGPFGNLDKGEPVNEADAAALEHDKAYDQQLKAGDNPYLKYNHA DAEFQERLQEDTSFGGNLGRAVFQAKKRILEPLGLVEEAAKTAPGKKGAV DQSPQEPDSSSGVGKSGKQPKARKRLNFGQTGDSESVDPQPPLGEPPAAPT LGSNTMASGGGAPMADNNEGADGVGNSSGNWHCDSTWLGDREVTTSTR TWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRD WQRLINNNWGFPRKLSFKLFNIQVRGVTQNDGTTIANNLSTVQVFTD SEYQLPYVLGSAHQGCLPPFPADVFMVMPQYGYLTLNNGSQAVGRSSFYCL EYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLN RTQGTTS GTTNQSRLLSQAGPQSMQLARNWLPGPCYRQQRLSKTANDN NNSNFPWTAASKYHLNGRDSL VNPGPAMASHKDDEEKFFPMHGNLIFGK EGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANNLQSSNTAPTTGT VNHQGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPP QIMIKNTPVPANPPTTFSPAKFAFITQYSTGQVSVEIEWELQKENS KRWN PEIQYTSNYNKS VNVDFTVDNTNGVYSEPRPIGTRYLTRNL
SEQ ID NO: 63 AAV4 CAP	MTDGYLPDWLEDNLSEGVREWWALQPGAPKPKANQQHQDNARGLVLP YKYLGPFGNLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHA DAEFQQLQGDTSFGGNLGRAVFQAKKRVLEPLGLVEQAGETAPGKKRP LIESPQQPDSTGIGKKGKQPAKKLVFEDETGAGDGPPEGSTSGAMSDDS EMRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWV LPTYNNHLYKRLGESLQSNYNGFSTPWGYFDFNRFHCHFSPRDWQRLIN NNWGMRPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELP YVMDAQEGSLPPFPNDVFMVMPQYGYCGLVTGNTSQQQTDNRNAYFCLEY FPSQMLRTGNNFEITYSFEKVPFHSMYAHSQSLDRLMNPLIDQYLWGLQST

Sequence identifier and name	Sequence
	TTGTTLNAGTATTNFTKLRPNTFSNFKKNWLPGPSIKQQGFSKTANQNYKI PATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPK QNGNTATVPGTLIFTSEEELAAATNATDMDWGNLPGGDQSNSNLPTVDRL TALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPI FIKNTVPANPATTFSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEV QFTSNYGOQNSLLWAPDAAGKYTEPRAIGTRYLTHHL
SEQ ID NO: 64 AAV5 CAP	MSFVDHPPDWLEEVGEGLEFLGLEAGPPKPKPNQQHQDQARGLVLPGY NYLGPNGLDGRGEPVNRADDEVAREHDISYNEQLEAGDNPYLKYNHADA EQEKLADDTSGGNLKGAVFQAKKRVLEPFGLVEEGAKTAPTGRIDDHF PKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAPASSLGADTMSAGGG PLGDNNQGADGVGNASGDWHCDSTWMDRVTKSTRTWLPSYNNHQ YREIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVG NGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSSSFFCLEYFPSKMLR TGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLYRFVSTNNTGGVQ FNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRME LEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGN MLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEI VPGSVW MERDVYLQGPWAKIPETGAHFHPSPAMGGFGLKHPPPMMLIKNTVPNGNI TSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQ FVDFAPDSTGEYRTRTPIGTRYLTRPL
SEQ ID NO: 65 AAV6 CAP	MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLP GYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHA DAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPFGLVEEGAKTAPGKKRPV EQSPQEPDSSSGIGKTGQPAKKRLNFGQTGDSESVDPQPLGEPPATPAA VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITSTR TWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSR DWQRLINNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFS DSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYC LEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYYL NRTQNQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTD NNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDKFFPMMSGVMIFGK ESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNLQSSSTDPATGDV HVMGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPP QILIKNTVPANPPAEFSATKFASFITQYSTGQVSVIEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL
SEQ ID NO: 66 AAV7 CAP	MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDNRGLVLP GYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHA DAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPAKKRPV EPSPQRSPTSSTGIGKKGQPPARKRLNFGQTGDSESVDPQPLGEPPAAPSS VGSGTVAAGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITSTR TWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDFNRFHCHFSR DWQRLINNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTIQVFS SEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQSVGRSSFYCLE YFPSQMLRTGNNFEFSYSFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYYLAR TQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQN NNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDDEDRFFPSSGVLIFGK TGATNKTTLNVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQVV NNQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLGGFGLKHPP QILIKNTVPANPPEVFTPAKFASFITQYSTGQVSVIEIEWELQKENS KRWNPEIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

Sequence identifier and name	Sequence
SEQ ID NO: 67 AAV8 CAP	MAADGYLPDWLEDNLSEGIREWWALKPGAPKPKANQQKQDDGRGLVLP GYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLQAGDNPYLRYNHA DAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPV EPSPQRSPDSSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPPLGEPPAAPSG VGPNTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDREVITSTR TWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDNRFHCHFS RDWQRLINNNWGFRPKRLSFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFT DSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYC LEYFPSQMLRTGNNFQFTYTFEDVPHSSYAHQSGLDRLMNPLIDQYLYYL SRTQTTGGTANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQ NNNSNFAWTAGTKYHLNGRNSLANPGIAMATHKDDEERFFPSNGILIFGK QNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQQQNTAPQIGT VNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPLMGGFGLKHPP PQILIKNTPVPADPPTTFNQSKLNSFITQYSTGQVSVEIEWELQKENS KRWN PEIQYTSNYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL
SEQ ID NO: 68 AAVrh.10 CAP	MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDGRGLVLP GYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHA DAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPV EPSPQRSPDSSSTGIGKKGQQPAKKRLNFGQTGDSESVDPQPPIGEPPAGPSG LGSGTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDREVITSTR TWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDNRFHCHFS RDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFT DSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYC LEYFPSQMLRTGNNFESYQFEDVPHSSYAHQSGLDRLMNPLIDQYLYYL SRTQSTGGTAGTQQLLFSQAGPNNMSAQAKNWLPGPCYRQQRVSTTSL NNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDEERFFPSGVL MFG KQGAGKDNVDYSSVMLTSEEEIKTTNPVATEQYGVVADNLQQQNAAPIV GAVNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPLMGGFGLKH PPPQILIKNTPVPADPPTTFSQAKLASFITQYSTGQVSVEIEWELQKENS KR WNPEIQYTSNYKSTNVDFAVNTDGTYSEPRPIGTRYLTRNL
SEQ ID NO: 69 AAV-TT CAP	MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLP YKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD AEFQERLKEDTSFGGNLGRAVFQAKKRILEPLGLVEEPVKTA PGKKRPVEH SPAEPDSSSGTGKSGQQPARKRLNFGQTGDADSVDPQPPLGQPPAAPSGLG TNTMASGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITSTRTW ALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCHFS PRDWQ RLINNNWGFRPKRLSFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEY QLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYF PSQMLRTGNNFTFSYTFEDVPHSSYAHQSGLDRLMNPLIDQYLYYLSRTN TPSGTTTMSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTAADNNNSD YSWTGATKYHLNGRDSLVPNGPAMASHKDDEEKYFPQSGVLIFGKQDSG KTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQSGNTQAATSDVNTQG VLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIK NTPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYT SNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL
SEQ ID NO: 70 AAV9 VP1 sa del	ATGGCTGCCGATGGTTATCTTCCTGATTGGCTCGAGGACAACCTTAGTG AAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCG

Sequence identifier and name	Sequence
	ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGC CTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTC CACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAAC GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGT CAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCGGTTCCAGCGGAC GTTTTCATGATTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTCTGTTCTGCTCTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATT AACGGTTCTGGACAGAATCAACAAACGCTAAAATTCAAGTGTGGCCGGA CCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGC TACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGC GAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATA GCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGG ACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACT GGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGA AGAAATTAATACTACTAACCCGGTAGCAACGGAGTCCTATGGACAAGT GGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGG TTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATG TGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCA ACTTTACCCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCC TCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCCAACG GCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTG GCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAAACAGC AAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACCTATTACAAGTCT AATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCC GCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *VP2/VP3 splicing acceptor AG is mutated to TG, indicated in bold, underlined text.
SEQ ID NO: 71 AAV9 Cap VP2-Gly455 antibody fragment (GGGSX2-GGGSX2 linker)	ATGACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAA CCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGACAGCCCGCTAAA AAGAGACTCAATTTCCGGTCAGACTGGCGACACAGAGTCAGTCCCAGAC CCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCT CTTACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAA GGTGCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCC CAATGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACA TCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCT GGGGGTATTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGA

Sequence identifier and name	Sequence
	CTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACT CAACTTCAAGCTCTTCAACATTCAAGGTCAAAGAGGTTACGGACAACAA TGGAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTT CACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGA GGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTTCATGATTCCTCAGTAC GGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTGTCCT TTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAA CTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTAC GCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAA TACTTGTACTATCTCTCAAAGACTATTAACGGTTCTGGAGggtggaggcggggt ggagggtggcggggtGTTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTG CTCCCCGGGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGG ACGGAGATACGGGGGATACGGCGCAGACTCGGGAGTGGGATCGggc gggggtggaagtggcggggtggaagtAATCAACAACGCTAAAATTCAGTGTGGC CGGACCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACC CAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAA CAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGT AATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGA GAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTTGGCAAACAAG GAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAAC GAAGAAGAAATTAATACTACTAACCCGGTAGCAACGGAGTCCTATGGA CAAGTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGG CTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAG AGATGTGTACCTGCAAGGACCCATTGGGGCCAAATTCCTCACACGGA CGGCAACTTTCACCTTCTCCGCTGATGGGAGGGTTTGAATGAAGCAC CCGCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTC CAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTC TACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAA ACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACCTATTACA AGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGA ACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *Antibody fragment (e.g. B1) is indicated in upper case, bold and single underlined. GGGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:72 AAV9 VP2 Cap N-term antibody fragment (GGGGG linker)	ATGGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCG GGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGA TACGGGGGATACGGCGCAGACTCGGGAGTGGGATCGggtggaggcgggg gtACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACC GGACTCCTCCGCGGGTATTGGCAAATCGGGTGACAGCCCGCTAAAAA GAGACTCAATTTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCC TCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAA TGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTG CCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGG GGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACT GGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCCAGCGGACGTTTTTCATGATTCTCAGTACGG

Sequence identifier and name	Sequence
	GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTTCGTCCTTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAAC TTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGTACTATCTCTCAAAGACTATTAACGGTCTGACAGAAATCAACA AACGCTAAAATTCAAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATC TTTAATTTTTGGCAAACAAGGAAGTGGAAAGAGACAACGTGGATCGGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCTGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCA AGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGG TATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTGGGC CAAAATTCCTCACACGGACGGCAACTTTCACCCCTTCTCCGCTGATGGGA GGGTTTGAATGAAGCACCCGCCTCTCAGATCCTCATCAAAAACACA CCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAAC TCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGT GGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAATACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATA CTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA * Antibody fragment (e.g. B1) is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:73 AAV9 VP1 del	<u>TG</u>AGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTG AAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAGAAGAGGC CTGTAGAGCAGTCTCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCC CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGCCCTGCCCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTCT CACTGCCACTTCTACACACGTGACTGGCAGCGACTCATCAACAACAAC GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTACAGGT CAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCAGCGGAC GTTTTTCATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTGCTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT

Sequence identifier and name	Sequence
	AATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATT AACGGTTCTGGACAGAATCAACAAACGCTAAAATTTCAGTGTGGCCGGA CCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAGC TACCGACAACAACGTGTCTCAACCACTGTGACTCAAACAACAACAGC GAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATA GCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGG ACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTTGGCAAAACAAGGAAC GGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGA AGAAATTAATACTACTAACCCGGTAGCAACGGAGTCTATGGACAAGT GGCCACAAACCACAGGAGTGCCCAAGCACAGGCGCAGACCGGCTGGG TTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATG TGTAACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGACGGCA ACTTTCACCCCTTCTCCGCTGATGGGAGGGTTTGGGAATGAAGCACCCGCC TCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCCAACG GCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTG GCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGC AAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACATTACAAGTCT ATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCC GCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA <i>*VP1 ATG start codon is mutated to TGA, indicated in bold, underlined text.</i>
SEQ ID NO:74 AAV9 VP1 sa del + AAV9 Gly455 antibody fragment (GGGGSX2- GGGSX2 linkers)	ATGGCTGCCGATGGTTATCTTCC<u>TG</u>ATTGGCTCGAGGACAACCTTAGTG AAGGAATTTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAATAACCTTGGACCCGGCAACGGACTCGACAAAGGGGAGCCGG TCAACGCAGCAGACGCGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGC CTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCCGTTCAGACTG GCGACACAGAGTCAGTCCAGACCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCG ACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGGCCCTGCCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCCGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTCT CACTGCCACTTCTCACACCGTGACTGGCAGCGACTCATCAACAACAAC GGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGT CAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTACGGACTCAGACTATCAGCTCCCGTAC GTGCTCGGGTCCGCTCACGAGGGCTGCCTCCCGCCGTTCCAGCGGAC GTTTTCATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTGCTTCGTCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATT AACGGTTCTGGA<u>gctggaggcgggagtgaggcgccggagtGTTTGTCCCCCTGGC</u> <u>TACAAGAGCGCGGTGGACTGCTCCCCGGGATCGGAGTGCAAAATGCG</u> <u>GATGTGCTACGCGTGGATGGACGGAGATACGGGGATACGGCGCA</u> <u>GACTCGGGAGTGGGATCG</u><u>gccccggctggaatgcccgaagctggaat</u>AATCAACA

Sequence identifier and name	Sequence
	AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATC TTTAATTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCA AGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGG TATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTGGGG CAAAATTCCTCACACGGACGGCAACTTTCACCCTTCTCCGCTGATGGGA GGGTTTGGAAATGAAGCACCCGCTCCTCAGATCCTCATCAAAAAACA CCTGTACCTGCGGATCCTCCAACGGCTTCAACAAGGACAAGCTGAAC TCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGT GGGAGCTGCAGAAGGAAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAATTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATA CTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA *VP2/VP3 splicing acceptor AG is mutated to TG, indicated in upper case, bold, and double underlined text. * Antibody fragment (e.g. B1) is indicated in upper case, bold and single underlined. GGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:75 AAV9 VP1 del + AAV9 Gly455 Antibody fragment (GGGGSX2- GGGSX2 linkers)	<u>TG</u>AGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTG AAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCCTCAACCCA AGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGG GTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG TCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAGAAGAGGC CTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCA AATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCCGGTCAGACTG GCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCG CAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC ACCAAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCTC GGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCAC CACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAATCACCTCTA CAAGCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGC CTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTCAACAGATTC CACTGCCACTTCTCACACGTGACTGGCAGCGACTCATCAACAACAAC GGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGT CAAAGAGGTTACGACAACAATGGAGTCAAGACCATCGCCAATAACCT TACCAGCACGGTCCAGGTCTTACGGACTCAGACTATCAGTCCCGTAC GTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGAC GTTTTATGATTCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCC AGGCCGTGGGTGCTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCA AATGCTAAGAACGGGTAACAACCTTCCAGTTCAGCTACGAGTTTGAGAA CGTACCTTTCCATAGCAGTACGCTCACAGCCAAAGCCTGGACCGACT AATGAATCCACTCATCGACCAATACTTGTAATCTCTCAAGAGACTAT AACGGTTCTGGAG<u>ggtggagggcgagtgagtggtggcgagtggtGGTTGTCCCCCTGGC</u>

Sequence identifier and name	Sequence
	<u>TACAAGAGCGGCGTGGACTGCTCCCCGGGATCGGAGTGCAAATGG</u> <u>GGATGCTACGCCGTGGATGGACGGAGATACGGGGGATACGGCGCA</u> <u>GACTCGGGAGTGGGATCG</u><u>ggcggggtggaagtggcgagggtggaagt</u>AATCAACA AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATC TTTAATTTTGGCAAACAAGGAAGTGGAAAGAGACAACGTGGATGCGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAATCCCGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACAGAGTGGCCCA AGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGG TATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTGGGC CAAAATTCCTCACACGGACGGCAACTTTCACCCCTTCTCCGCTGATGGGA GGGTTTGGAAATGAAGCACCCGCTCCTCAGATCCTCATCAAAAACACA CCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAAC TCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGT GGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAATACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATA CTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA *VP1 ATG start codon is mutated to TGA, indicated in upper case, bold, and double underlined text. * Antibody fragment (e.g. B1) is indicated in upper case, bold and single underlined. GGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:76 AAV9 VP2 N-term E3.1 (E1 + GGGGG linker)	<u>ATGACTTGCCCGGATGGGTACACACGGACCAACTACTACTGTCCGA</u> <u>GGGACGGCTGCGGCTCATGGTGCAACGGTGCGGAGCGTCAGCAGC</u> <u>CGTGTATAAGAGGACCCTGTTGTTGTGACCTTACT</u><u>tggtgaggcggggtA</u> CCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGG ACTCCTCCGCGGGTATTGGCAAATCGGGTGACAGCCCGCTAAAAAGA GACTCAATTTGCGTCAGACTGGCGACACAGAGTCAGTCCAGACCCTC AACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTA CAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTG CCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAAT GGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGC CCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCTG GAGGATCTTCAAATGACAACGCCTACTTCGGGTACAGCACCCCCCTGGG GGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTG GCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCAA CTTCAAGCTCTTCAACATTCAAGGTCAAAGAGGTTACGGACAACAATGG AGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCAC GGACTCAGACTATCAGTCTCCCGTACGTGCTCGGGTCGGCTCACGAGGG CTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGGG TATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTGCTTCGTCCTTTT ACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAAC TCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGC TCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATA CTTGTAATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACAA ACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGA AGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACC ACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTT

Sequence identifier and name	Sequence
	CTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTAT GGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATCT TTAATTTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGAC AAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGTA GCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCAA GCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGGT ATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCC AAAATTCCTCACACGGACGGCAACTTTACCCCTTCTCCGCTGATGGGAG GGTTTGAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACAC CTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAAC CTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTG GGAGCTGCAGAAGGAAAAACAGCAAGCGCTGGAACCCGGAGATCCAGT ACACTTCCAACATTACAAGTCTAATAATGTTGAATTTGCTGTTAATAC TGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGAC TCGTAATCTGTAA *E1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:77 AAV9 VP2 N-term F3.1 (F1 + GGGGG linker)	<u>ATGCGATGCCCGGACGGAAGCATATGGAGGTGCGGGGATGTGGAC</u> <u>GGAGCGACGTGTTTATTGCGACCATGCGCCTGTTGGGTTCAGAAAG</u> <u>TCATCAGGTGGCTGGTACTGTAAGCCACCGGCCACT</u><u>ggtggaggcgggggt</u> ACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCG GACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAAAAG AGACTCAATTTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCTT CAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAA TGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTG CCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCTGG GGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACT GGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGTCTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGG GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTGTTCTGCTCTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAAC TTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGTAATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACA AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCTTTGTCTGGATC TTAATTTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCA AGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGG TATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGC CAAAATTCCTCACACGGACGGCAACTTTACCCCTTCTCCGCTGATGGGA GGGTTTGAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACA

Sequence identifier and name	Sequence
	CCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAAC TCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGT GGGAGCTGCAGAAGGAAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAACATTACAAGTCTAATAATGTTGAATTTGCTGTTAATA CTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA *F1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:78 AAV9 VP2 N-term G3.1 (G1 + GGGGG linker)	ATG<u>TCGTGTCCGGATGGATACACATACCGCTTCGGTAGTGT</u><u>TTGTA</u> <u>CAGCACGCTGGCGTAATGACTGCTTTCGCGATGCGGATTATGATTA</u> <u>TCAGCGAAGATGCTACACCACTTCCCTTCCACAGACG</u><u>ggtggagcgggg</u> <u>gt</u>ACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACC GGACTCCTCCGCGGGTATTGGCAAAATCGGGTGCACAGCCCGCTAAAAA GAGACTCAATTTCCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCC TCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTT ACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGT GCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCCCAA TGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTG CCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACATCT GGAGGATCTTCAATGACAACGCCTACTTCGGCTACAGCACCCCTGG GGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACT GGCAGCGACTCATCAACAACAACTGGGGATTCCGGCCTAAGCGACTCA ACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATG GAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCA CGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGAGG GCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTACGG GTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCTGCTCTT TACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAAC TTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACG CTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT ACTTGTAATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACA AACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGG AAGAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAAC CACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCT TCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTA TGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTCTTTGTCTGGATC TTTAATTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGA CAAAGTCATGATAACCAACGAAGAAGAAATTAATACTACTAACCCGGT AGCAACGGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCA AGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGGG TATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGC CAAAATTCCTCACACGGACGGCAACTTTCACCCCTTCTCCGCTGATGGGA GGGTTTGGAAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACA CCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAAC TCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGT GGGAGCTGCAGAAGGAAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAACATTACAAGTCTAATAATGTTGAATTTGCTGTTAATA CTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA *G1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.

Sequence identifier and name	Sequence
SEQ ID NO:79 AAV9 VP2 Gly 455 E3.2 (E1 knob + GGGGSX2-linkers)	ATGACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAA CCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAA AAGAGACTCAATTTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGAC CCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCT CTTACAATGGCTTCAGGTGGTGGCGCACCAAGTGGCAGACAATAACGAA GGTGCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCC CAATGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACA TCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCT GGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGA CTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACT CAACTTCAAGCTCTTCAACATTACAGGTCAAAGAGGTTACGGACAACAA TGGAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTT CACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGA GGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCTCAGTAC GGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCTTCGTCTCT TTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAAACAA CTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTAC GCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAA TACTTGTAATCTCTCAAAGACTATTAACGGTTCTGGA<u>ggaggaggcgagg</u> <u>ggaggaggcgagg</u><u>ACTTGCCCGGATGGGTACACACGGACCAACTACTAC</u> <u>TGTGGAGGGACGGCTGCGGCTCATGGTGCAACGGTGCGGAGCGT</u> <u>CAGCAGCCGTGTATAAGAGGACCCTGTTGTTGTGACCTTACT</u><u>ggcggg</u> <u>ggtggaagtggcggaggtggaagt</u>AAATCAACAAACGCTAAAATTCAGTGTGGCCGG ACCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCCAG CTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAG CGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAAT AGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAG GACCGTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAAC TGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAG AAGAAATTAATACTACTAACCCTGGTAGCAACGGAGTCTATGGACAAG TGGCCACAAACCACAGAGTGCCCAAGCACAGGCGCAGACCGGCTGG GTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGAT GTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGC AACTTTCACCTTCTCCGCTGATGGGAGGGTTTGGAAATGAAGACCCGC CTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCCAAC GGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCAGTATTCTACT GGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAG CAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACATTACAAGTC TAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCC CGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *E1 knob is indicated in upper case, bold and single underlined. GGGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:80 AAV9 VP2 Gly 455 F3.2 (F1 knob + GGGGSX2-linkers)	ATGACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAA CCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAA AAGAGACTCAATTTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGAC CCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCT CTTACAATGGCTTCAGGTGGTGGCGCACCAAGTGGCAGACAATAACGAA GGTGCCGATGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCC CAATGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACA TCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCT

Sequence identifier and name	Sequence
	GGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGA CTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACT CAACTTCAAGCTCTTCAACATTCAAGGTCAAAGAGGTTACGGACAACAA TGGAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTT CACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGA GGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCCCTCAGTAC GGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTTCGTCTT TTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAA CTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTAC GCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAA TACTTGTACTATCTCTCAAAGACTATTAACGGTTCTGGAGggaggaggcgggagt ggaggaggcgggagtCGATGCCCGGACGGAAGCATATGGAGGTGCGGGGA TGTGGACGGAGCGACGTGTTTCATTGCGACCATGCGCCTGTTGGGT TCAGAAGTCATCAGGTGGCTGGTACTGTAAGCCACCGGCGACTggcg ggggtggaagtggcgggaggtggaagtAATCAACAACGCTAAAATTTCAGTGTGGCC GGACCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACCC AGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAAC AGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTA ATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAG AGGACCGTTTTCTTCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGG AACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACG AAGAAGAAATTAATACTACTAACCCGGTAGCAACGGAGTCCTATGGAC AAGTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGC TGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGA GATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGAC GGCAACTTTCACCTTCTCCGCTGATGGGAGGGTTTGAATGAAGCACC CGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCC AACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCT ACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAA CAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACCTATTACAA GTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAA CCCCGCCCATTTGGCACCAGATACCTGACTCGTAATCTGTAA *F1 knob is indicated in upper case, bold and single underlined. GGGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:81 AAV9 VP2 Gly 455 G3.2 (G1 knob + GGGGSX2- GGGGSX2 linkers)	ATGACCGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAA CCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGACAGCCCGCTAAA AAGAGACTCAATTTCCGGTCAGACTGGCGACACAGAGTCAGTCCCAGAC CCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCT CTTACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAA GGTGCCGATGGAGTGGGTAGTTCCTCGGGAAAATTGGCATTGCGATTCC CAATGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACA TCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCT GGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGA CTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCTAAGCGACT CAACTTCAAGCTCTTCAACATTCAAGGTCAAAGAGGTTACGGACAACAA TGGAGTCAAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTT CACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCACGA GGGCTGCCTCCCGCCGTTCCAGCGGACGTTTTTCATGATTCCCTCAGTAC GGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTTCGTCTT TTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAA CTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTAC

Sequence identifier and name	Sequence
	GCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAA TACTTGACTATCTCTCAAAGACTATTAACGGTTCTGGAggtggagcgggagt ggaggtggcgggagtTCGTGTCCGGATGGATACACATAACCGCTTCGGTAGT GTTTGTACAGCACGCTGGCGTAATGACTGCTTTCGCGATGCGGATT ATGATTATCAGCGAAGATGCTACACCACTTCCCTTCCACAGACGgc gggggtggaagtggcggaggtggaagtAATCAACAAACGCTAAAATTCAGTGTGGC CGGACCCAGCAACATGGCTGTCCAGGGAAGAACTACATACCTGGACC CAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAA CAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGT AATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGA GAGGACCGTTTCTTTCTTTGTCTGGATCTTTAATTTTGGCAAACAAG GAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATAACCAAC GAAGAAGAAATTAATACTACTAACCCGGTAGCAACGGAGTCTATGGA CAAGTGGCCACAAACCACAGAGTGCCCAAGCACAGGCGCAGACCGG CTGGGTTCAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAG AGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGA CGGCAACTTTCACCCCTTCTCCGCTGATGGGAGGGTTTGAATGAAGCAC CCGCCTCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTC CAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTC TACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAA ACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACATTACA AGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGA ACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA *G1 knob is indicated in upper case, bold and single underlined. GGGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:82 AAV8 Del Vp2	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTG AGGGCATTCGCGAGTGGTGGGCGCTGAAACCTGGAGCCCCGAAGCCCA AAGCCAACCAGCAAAAGCAGGACGACGGCCGGGGTCTGGTGCTTCCTG GCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAGGGGGAGCCCG TCAACGCGGCGGACGCAGCGGCCCTCGAGCACGACAAGGCCTACGACC AGCAGCTGCAGGCGGGTGACAATCCGTACCTGCGGTATAACCACGCCG ACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAGAAGCGGGTTCTCGAACCTC TCGGTCTGGTTGAGGAAGGCGCTAAGACCGCTCCTGGAAAGAAGAGAC CGGTAGAGCCATCACCCAGCGTTCTCCAGACTCCTCTACGGGCATCGG CAAGAAAGGCCAACAGCCCCGCCAGAAAAAGACTCAATTTTGGTCAGAC TGGCGACTCAGAGTCAGTTCCAGACCCTCAACCTCTCGGAGAACCTCC AGCAGCGCCCTCTGGTGTGGGACCTAATACAATGGCTGCAGGCGGTGG CGCACCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGTAGTTC CTCGGGAAATTGGCATTGCGATTCCACATGGCTGGGCGACAGAGTCAT CACCACCAGCACCCGAACCTGGGGCCCTGCCACCTACAACAACCACT CTACAAGCAAATCTCCAACGGGACATCGGGAGGAGCCACCAACGACA ACACCTACTTCGGCTACAGCACCCCTGGGGGTATTTTGACTTTAACAG ATTCCACTGCCACTTTTACCACGTGACTGGCAGCGACTCATCAACAAC AACTGGGGATTCCGGCCCAAGAGACTCAGCTTCAAGCTCTTCAACATC CAGGTCAAGGAGGTCACGCAGAATGAAGGCACCAAGACCATCGCCAA TAACCTCACAGCACCATCCAGGTGTTTACGGACTCGGAGTACCAGCT GCCGTACGTTCTCGGCTCTGCCACCAAGGGCTGCCTGCCTCCGTTCCCG GCGGACGTGTTTCATGATTCCCCAGTACGGCTACCTAACACTCAACAAC GGTAGTCAGGCCGTGGGACGCTCCTCCTTCTACTGCCTGGAATACTTTC CTTCGCAGATGCTGAGAACCAGCAACAACCTCCAGTTTACTTACACCTT CGAGGACGTGCCTTTCCACAGCAGCTACGCCACAGCCAGAGCTTGA

Sequence identifier and name	Sequence
	CCGGCTGATGAATCCTCTGATTGACCAGTACCTGTACTACTTGTCTCGG ACTCAAACAACAGGAGGCACGGCAAATACGCAGACTCTGGGCTTCAGC CAAGGTGGGCCTAATAACAATGGCCAATCAGGCAAAGAACTGGCTGCCA GGACCCTGTTACCGCCAACAACGCGTCTCAACGACAACCGGGCAAAC AACAAATAGCAACTTTGCCTGGACTGCTGGGACCAAATACCATCTGAAT GGAAGAAATTCATTGGCTAATCCTGGCATCGCTATGGCAACACACAAA GACGACGAGGAGCGTTTTTTTCCAGTAACGGGATCCTGATTTTTGGCA AACAAAATGCTGCCAGAGACAATGCGGATTACAGCGATGTCATGCTCA CCAGCGAGGAAGAAATCAAAACCACTAACCTGTGGCTACAGAGGAAT ACGGTATCGTGGCAGATAACTTGCAGCAGCAAAACACGGCTCCTCAA TTGGAACGTGTCAACAGCCAGGGGGCCTTACCCGGTATGGTCTGGCAGA ACCGGGACGTGTACCTGCAGGGTCCCATCTGGGCCAAGATTCTCACA CGGACGGCAACTTCCACCCGTCTCCGCTGATGGGCGGCTTTGGCTGAA ACATCCTCCGCCTCAGATCCTGATCAAGAACACGCCTGTACCTGCGGAT CCTCCGACCACCTTCAACCAGTCAAAGCTGAACTCTTTCATCACGCAAT ACAGCACCGGACAGGTACGCGTGGAAATTGAATGGGAGCTGCAGAAG GAAAACAGCAAGCGCTGGAACCCCGAGATCCAGTACACCTCCAACCTAC TACAAATCTACAAGTGTGGACTTTGCTGTTAATACAGAAGGCGTGTACT CTGAACCCCGCCCCATTGGCACCCGTTACCTACCCGTAATCTGTAA *VP2 start codon is in upper case, bold and single underlined, and G was mutated to C in the VP2 start codon.
SEQ ID NO:83 AAV8 VP2 N-term B1	ATGGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCG GGATCGGAGTGC AAATGGGGATGCTACGCCGTGGATGGACGGAGA <u>TACGGGGGATACGGCGCAGACTCGGGAGTGGGATCG</u>ACCGCTCCT GGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTCTCCAGACTCC TCTACGGGCATCGGCAAGAAAGGCCAACAGCCCGCCAGAAAAAGACT CAATTTTGGTCAGACTGGCGACTCAGAGTCAGTTCCAGACCCTCAACCT CTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGACCTAATAACAATG GCTGCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGA CGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCACATGGCT GGGCGACAGAGTCATCACACCAGCACCCGAACCTGGGCCCTGCCCAC CTACAACAACCACCTCTACAAGCAAATCTCCAACGGGACATCGGGAGG AGCCACCAACGACAACACCTACTTCGGCTACAGCACCCCTGGGGGTA TTTTGACTTTAACAGATTCCACTGCCACTTTTACCACGTGACTGGCAG CGACTCATCAACAACAACCTGGGGATTCCGGCCCAAGAGACTCAGCTTC AAGCTCTTCAACATCCAGGTCAAGGAGGTACGCAGAATGAAGGCACC AAGACCATCGCCAATAACCTCACCAGCACCATCCAGGTGTTTACGGAC TCGGAGTACCAGCTGCCGTACGTTCTCGGCTCTGCCCACCAGGGCTGCC TGCCTCCGTTCCCGGCGGACGTGTTTATGATTCCCCAGTACGGCTACCT AACACTCAACAACGGTAGTCAAGGCCGTGGGACGCTCCTCCTTCTACTGC CTGGAATACTTTCCTTCGCAGATGCTGAGAACCGGCAACAACCTCCAGT TTACTTACACCTTCGAGGACGTGCCTTTCCACAGCAGCTACGCCACAG CCAGAGCTTGGACCGGCTGATGAATCCTCTGATTGACCAGTACCTGTAC TACTTGTCTCGGACTCAAACAACAGGAGGCACGGCAAATACGCAGACT CTGGGCTTCAGCCAAGGTGGGCCTAATAACAATGGCCAATCAGGCAAAG AACTGGCTGCCAGGACCCTGTTACCGCCAACAACGCGTCTCAACGACA ACCGGGCAAAACAACAATAGCAACTTTGCCTGGACTGCTGGGACCAAA TACCATCTGAATGGAAGAAATTCATTGGCTAATCCTGGCATCGCTATGG CAACACACAAAGACGACGAGGAGCGTTTTTTTCCAGTAACGGGATCC TGATTTTTGGCAAACAATAATGCTGCCAGAGACAATGCGGATTACAGCG ATGTCATGCTCACCAGCGAGGAAGAAATCAAAACCACTAACCTGTGG CTACAGAGGAATACGGTATCGTGGCAGATAACTTGCAGCAGCAAAACA

Sequence identifier and name	Sequence
	CGGCTCCTCAAATTGGAAGTGTCAACAGCCAGGGGGCCTTACCCGGTA TGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTCCCATCTGGGCCA AGATTCCCTCACACGGACGGCAACTTCCACCCGTCTCCGCTGATGGGCG GCTTTGGCCTGAAACATCCTCCGCCTCAGATCCTGATCAAGAACACGCC TGTACCTGCGGATCCTCCGACCACCTTCAACCAGTCAAAGCTGAACTCT TTCATCACGCAATACAGCACCGGACAGGTGACGCTGGAATTTGAATGG GAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCCGAGATCCAGTA CACCTCCAATACTACAAATCTACAAGTGTGGACTTTGCTGTTAATACA GAAGGCGTGTACTCTGAACCCCGCCCCATTGGCACCCGTTACCTCACCC GTAATCTGTAA * B1 (knob domain) is indicated in upper case, bold and single underlined.
SEQ ID NO:84 AAV8 VP2 N-term B2	<u>ATGGTGCACCAGCGGACCATCAAGTCCGGTTGTCCCCCTGGCTACA</u> <u>AGAGCGGCGTGGACTGCTCCCCGGGATCGGAGTGCAAATGGGGAT</u> <u>GCTACGCCGTGGATGGACGGAGATACGGGGGATACGGCGCAGACT</u> <u>CGGGAGTGGGATCGACCTACACCCACGAGTTCTACGTGACCGCTCC</u> TGGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTCTCCAGACTC CTCTACGGGCATCGGCAAGAAAGGCCAACAGCCCGCCAGAAAAAGAC TCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTTCCAGACCCTCAACC TCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGACCTAATACAAT GGCTGCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCG ACGGAGTGGGTAGTTCCCTCGGGAAATTGGCATTGCGATTCCACATGGC TGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCCA CCTACAACAACACCTCTACAAGCAAATCTCCAACGGGACATCGGGAG GAGCCACCAACGACAACACCTACTTCGGCTACAGCACCCCTGGGGGT ATTTTGACTTTAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCA GCGACTCATCAACAACAACCTGGGGATTCCGGCCCAAGAGACTCAGCTT CAAGCTCTTCAACATCCAGGTCAAGGAGGTCACGCAGAATGAAGGCAC CAAGACCATCGCCAATAACCTCACCAGCACCATCCAGGTGTTTACGGA CTCGGAGTACCAGCTGCCGTACGTTCTCGGCTCTGCCACCAGGGCTGC CTGCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCCCAGTACGGCTACC TAACACTCAACAACGGTAGTCAGGCCGTGGGACGCTCCTCCTTCTACTG CCTGGAATACTTTCTTTCGAGATGCTGAGAACCGGCAACAACCTCCAG TTACTTACACCTTCGAGGACGTGCCTTTCCACAGCAGCTACGCCCCA GCCAGAGCTTGGACCGGCTGATGAATCCTCTGATTGACCAGTACCTGTA CTACTTGTCTCGGACTCAAACAACAGGAGGCACGGCAAATACGCAGAC TCTGGGCTTCAGCCAAGGTGGGCCTAATAAATGGCCAATCAGGCAAA GAAGTGGCTGCCAGGACCTGTTACCGCCAACAACGCGTCTCAACGAC AACCGGGCAAAACAACAATAGCAACTTTGCCTGGACTGCTGGGACCAA ATACCATCTGAATGGAAGAAATTCATTGGCTAATCCTGGCATCGCTATG GCAACACACAAAGACGACGAGGAGCGTTTTTTTTCCAGTAACGGGATC CTGATTTTTGGCAAACAAAATGCTGCCAGAGACAATGCGGATTACAGC GATGTCATGCTCACCAGCGAGGAAGAAATCAAAACCACTAACCTGTG GCTACAGAGGAATACGGTATCGTGGCAGATAACTTGCAGCAGCAAAAC ACGGCTCCTCAAATTGGAAGTGTCAACAGCCAGGGGGCCTTACCCGGT ATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTCCCATCTGGGCC AAGATTCCCTCACACGGACGGCAACTTCCACCCGTCTCCGCTGATGGGC GGCTTTGGCCTGAAACATCCTCCGCCTCAGATCCTGATCAAGAACACGC CTGTACCTGCGGATCCTCCGACCACCTTCAACCAGTCAAAGCTGAACTC TTTCATCACGCAATACAGCACCGGACAGGTGACGCTGGAATTTGAATG GGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCCGAGATCCAGT ACACCTCCAATACTACAAATCTACAAGTGTGGACTTTGCTGTTAATAC

Sequence identifier and name	Sequence
	AGAAGGCGTGTA CTCTGAACCCCGCCCCATTGGCACCCGTTACCTCACC CGTAATCTGTAA *B2 (ultralong CDR-H3) is indicated in upper case, bold and single underlined.
SEQ ID NO:85 AAV8 VP2 N-term B3.1 (B1+GGGGG linker)	<u>ATGGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCG</u> <u>GGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGA</u> <u>TACGGGGGATACGGCGCAGACTCGGGAGTGGGATCG</u> <u>ggtggaggcggg</u> <u>gt</u> ACCGCTCCTGGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTC TCCAGACTCCTCTACGGGCATCGGCAAGAAAGGCCAACAGCCCGCCAG AAAAAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTTCCAGA CCCTCAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGACCT AATACAATGGCTGCAGGGCGGTGGCGCACCAATGGCAGACAATAACGA AGGCGCCGACGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATT CACATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGC CCTGCCACCTACAACAACCACTCTACAAGCAAATCTCCAACGGGAC ATCGGGAGGAGCCACCAACGACAACACCTACTTCGGCTACAGCACCCC CTGGGGGTATTTTGA CTTTAACAGATTCCACTGCCACTTTTCACCACGT GACTGGCAGCGACTCATCAACAACA ACTGGGGATTCCGGCCCAAGAGA CTCAGCTTCAAGCTCTTCAACATCCAGGTCAAGGAGGTCACGCAGAAT GAAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCATCCAGGTG TTTACGGACTCGGAGTACCAGCTGCCGTACGTTCTCGGCTCTGCCCACC AGGGCTGCCTGCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCCCAGTA CGGCTACCTAACACTCAACAACGGTAGTCAGGCCGTGGGACGCTCCTC CTTCTACTGCCTGGAATACTTTCCTTCGCAGATGCTGAGAACCGGCAAC AACTTCCAGTTTACTTACACCTTCGAGGACGTGCCCTTTCACAGCAGCT ACGCCACAGCCAGAGCTTGGACCGGCTGATGAATCCTCTGATTGACC AGTACCTGTACTACTTGTCTCGGACTCAAACAACAGGAGGCACGGCAA ATACGCAGACTCTGGGCTTCAGCCAAGGTGGGCCTAATACAATGGCCA ATCAGGCAAAGAACTGGCTGCCAGGACCCTGTTACCGCCAACAACGCG TCTCAACGACAACCGGGCAAAACAACAATAGCAACTTTGCCTGGACTG CTGGGACCAAATACCATCTGAATGGAAGAAATTCATTGGCTAATCCTG GCATCGCTATGGCAACACACAAGACGACGAGGAGCGTTTTTTTCCCA GTAACGGGATCCTGATTTTTGGCAAACAATGCTGCCAGAGACAATG CGGATTACAGCGATGTCATGCTCACCAGCGAGGAAGAAATCAAAACCA CTAACCTGTGGCTACAGAGGAATACGGTATCGTGGCAGATAACTTGC AGCAGCAAAACACGGCTCCTCAAATTGGAAGTGTCAACAGCCAGGGGG CCTTACCCGGTATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTC CCATCTGGGCCAAGATTCTCACACGGACGGCAACTTCCACCCGCTCTCC GCTGATGGGCGGCTTTGGCCTGAAACATCCTCCGCCTCAGATCCTGATC AAGAACACGCCTGTACCTGCGGATCCTCCGACCACCTTCAACCAGTCA AAGCTGAACCTTTTCATCACGCAATACAGCACCGGACAGGTCAGCGTG GAAATTGAATGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCC CGAGATCCAGTACACCTCCA ACTACTACAAATCTACAAGTGTGGACTTT GCTGTTAATACAGAAGGCGTGTACTCTGAACCCCGCCCCATTGGCACCC GTTACCTCACCCGTAATCTGTAA *B1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:86 AAV8 VP2 N-term B4.1	<u>ATGGTGCACCAGCGGACCATCAAGTCCGGTTGTCCCCCTGGCTACA</u> <u>AGAGCGGCGTGGACTGCTCCCCGGGATCGGAGTGCAAATGGGGAT</u> <u>GCTACGCCGTGGATGGACGGAGATACGGGGGATACGGCGCAGACT</u> <u>CGGGAGTGGGATCGACCTACACCCACGAGTTCTACGT</u> <u>ggtggaggcgg</u> <u>ggt</u> ACCGCTCCTGGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGT

Sequence identifier and name	Sequence
	TCTCCAGACTCCTCTACGGGCATCGGCAAGAAAGGCCAACAGCCCGCC AGAAAAAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTTCCA GACCCTCAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGA CCTAATACAATGGCTGCAGGCGGTGGCGCACCAATGGCAGACAATAAC GAAGGCGCCGACGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGAT TCCACATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGG GCCCTGCCCACCTACAACAACCACCTCTACAAGCAAATCTCCAACGGG ACATCGGGAGGAGCCACCAACGACAACACCTACTTCGGCTACAGCACC CCCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTTTCACCAC GTGACTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCCAAGA GACTCAGCTTCAAGCTCTTCAACATCCAGGTCAAGGAGGTCAACGCAGA ATGAAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCATCCAGG TGTTTACGGACTCGGAGTACCAGCTGCCGTACGTTCTCGGCTCTGCCCA CCAGGGCTGCCTGCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCCCAG TACGGCTACCTAACACTCAACAACGGTAGTCAGGCCCGTGGGACGCTCC TCCTTCTACTGCCTGGAATACTTTCCTTCGCAGATGCTGAGAACCGGCA ACAACTTCCAGTTTACTTACACCTTCGAGGACGTGCCTTTCCACAGCAG CTACGCCACAGCCAGAGCTTGGACCGGCTGATGAATCCTCTGATTGA CCAGTACCTGTACTACTTGTCTCGGACTCAAACAACAGGAGGCACGGC AAATACGCAGACTCTGGGCTTCAGCCAAGGTGGGCCTAATACAATGGC CAATCAGGCAAAGAAGTGGCTGCCAGGACCCTGTTACCGCCAACAACG CGTCTCAACGACAACCGGGCAAAACAACAATAGCAACTTTGCCTGGAC TGCTGGGACCAAATACCATCTGAATGGAAGAAATTCATTGGTAACT TGGCATCGCTATGGCAACACACAAAGACGACGAGGAGCGTTTTTTTCC CAGTAACGGGATCCTGATTTTTTGGCAAACAATAATGCTGCCAGAGACAA TGCGGATTACAGCGATGTCATGCTCACCAGCGAGGAAGAAATCAAAAC CACTAACCTGTGGCTACAGAGGAATACGGTATCGTGGCAGATAACTT GCAGCAGCAAAACACGGCTCCTCAAATTGGAAGTGTCAACAGCCAGGG GGCCTTACCCGGTATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGG TCCCATCTGGGCCAAGATTCTCTCACACGGACGGCAACTTCCACCCGTCT CCGCTGATGGGCGGCTTTGGCCTGAAACATCCTCCGCCTCAGATCCTGA TCAAGAACACGCCTGTACCTGCGGATCCTCCGACCACCTTCAACCAGTC AAAGCTGAAGTCTTTCATCACGCAATACAGCACCGGACAGGTCAGCGT GGAAATTGAATGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACC CCGAGATCCAGTACACCTCCAATACTACAAATCTACAAGTGTGGACTT TGCTGTTAATACAGAAGGCGTGTACTCTGAACCCCGCCCCATTGGCACC CGTTACCTCACCCGTAATCTGTAA *B1 knob + CDRH3 is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:87 AAV8 VP2 Gly 455 B3.2 (B1 + GGGGSX2- GGGGSX2 linkers)	ATGACCGCTCCTGGAAAGAAGAGACCGGTAGAGCCATCACCCCAAGCGT TCTCCAGACTCCTCTACGGGCATCGGCAAGAAAGGCCAACAGCCCGCC AGAAAAAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTTCCA GACCCTCAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGA CCTAATACAATGGCTGCAGGCGGTGGCGCACCAATGGCAGACAATAAC GAAGGCGCCGACGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGAT TCCACATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGG GCCCTGCCCACCTACAACAACCACCTCTACAAGCAAATCTCCAACGGG ACATCGGGAGGAGCCACCAACGACAACACCTACTTCGGCTACAGCACC CCCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTTTCACCAC GTGACTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCCAAGA GACTCAGCTTCAAGCTCTTCAACATCCAGGTCAAGGAGGTCAACGCAGA ATGAAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCATCCAGG

Sequence identifier and name	Sequence
	TGTTTACGGACTCGGAGTACCAGCTGCCGTACGTTCTCGGCTCTGCCCCA CCAGGGCTGCCTGCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCCCAG TACGGCTACCTAACACTCAACAACGGTAGTCAGGCCGTGGGACGCTCC TCCTTCTACTGCCTGGAATACTTTCCTTCGCAGATGCTGAGAACCGGCA ACAACTTCCAGTTTACTTACACCTTCGAGGACGTGCCTTTCCACAGCAG CTACGCCCCACAGCCAGAGCTTGGACCGGTGATGAATCCTCTGATTGA CCAGTACCTGTACTACTTGTCTCGGACTCAAACAACAGGAagtggaggcggg agtggaggcgggagtGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACT GCTCCCCGGGATCGGAGTGCAAATGGGGATGCTACGCCCGTGGATG GACGGAGATACGGGGGATACGGCGCAGACTCGGGAGTGGGATCG gccccgtggaagtggcggaggtggaagtACGGCAAATCAGCAGACTCTGGGCTTCA GCCAAGGTGGGCCTAATACAATGGCCAATCAGGCAAAGAACTGGCTGC CAGGACCCTGTTACCGCCAACAACGCGTCTCAACGACAACCGGGCAAA ACAACAATAGCAACTTTGCCTGGACTGCTGGGACCAAATACCATCTGA ATGGAAGAAATTCATTGGCTAATCCTGGCATCGCTATGGCAACACACA AAGACGACGAGGAGCGTTTTTTTCCCAGTAACGGGATCCTGATTTTTTG CAAACAAAATGCTGCCAGAGACAATGCGGATTACAGCGATGTCATGCT CACCAGCGAGGAAGAAATCAAAACCACTAACCCCTGTGGCTACAGAGG AATACGGTATCGTGGCAGATAACTTGCAGCAGCAAAACACGGCTCCTC AAATTGGAACGTGCAACAGCCAGGGGGCCTTACCCGGTATGGTCTGGC AGAACCGGGACGTGTACCTGCAGGGTCCCATCTGGGCCAAGATTTCCTC ACACGGACGGCAACTTCCACCCGCTCTCCGCTGATGGGCGGCTTTGGCCT GAAACATCTCCGCTCAGATCCTGATCAAGAACACGCTGTACCTGC GGATCCTCCGACCACTTCAACCACTCAAAGCTGAAGTCTTTTCATCAGC CAATACAGCACCCGACAGGTCAGCGTGGAATGAATGGGAGCTGCAG AAGGAAAACAGCAAGCGCTGGAACCCCGAGATCCAGTACACCTCCAAC TACTACAAATCTACAAGTGTGGACTTTGCTGTTAATACAGAAGGCGTGT ACTCTGAACCCCGCCCCATTGGCACCCGTTACCTACCCGTAATCTGTA A *B1 knob is indicated in upper case, bold and single underlined. GGGGSX2 linkers are indicated in lower case, bold and single underlined.
SEQ ID NO:88 AAV1 Del Vp2	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTG AGGGCATTTCGCGAGTGGTGGGACTTGAAACCTGGAGCCCCGAAGCCCA AAGCCAACCAGCAAAAGCAGGACGACGGCCGGGTCTGGTGCTTCCTG GCTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGGGGGAGCCCC TCAACGCGGCGGACGCAGCGGCCCTCGAGCAGGACAAGGCCTACGACC AGCAGTCAAAGCGGGTGACAATCCGTACCTGCGGTATAACCAACGCCG ACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATAACGTCTTTTGGGGCA ACCTCGGGCGAGCAGTCTTCCAGGCCAAGAAGCGGGTTCTCGAACCTC TCGGTCTGGTTGAGGAAGGCGCTAAGAACCGCTCTGGAAAGAAACGTC CGGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCCTCGGGCATCGGCA AGACAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTG GCGACTCAGAGTCAGTCCCCGATCCACAACCTCTCGGAGAACCTCCAG CAACCCCCGCTGCTGTGGGACCTACTACAATGGCTTCAGGCGGTGGCG CACCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGTAATGCCT CAGGAAATTGGCATTGCGATTCCACATGGCTGGGCGACAGAGTCATCA CCACCAGCACCCGCACCTGGGCCTTGCCACCTACAATAACCACTCTA CAAGCAAATCTCCAGTGCTTCAACGGGGGCCAGCAACGACAACCACTA CTTCGGCTACAGCACCCCTGGGGGTATTTTGATTTCAACAGATTCCAC TGCCACTTTTACCACGTGACTGGCAGCGACTCATCAACAACAATGGG GATTCCGGGCCCAAGAGACTCAACTCAAACCTTCAACATCCAAGTCA AGGAGGTCACGACGAATGATGGCGTCAACAACCATCGCTAATAACCTTA

Sequence identifier and name	Sequence
	CCAGCACGGTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGT CCTCGGCTCTGCGCACCAGGGCTGCCTCCCTCCGTTCCCGGCGGACGTG TTCATGATTCCGCAATACGGCTACCTGACGCTCAACAATGGCAGCCAA GCCGTGGGACGTTTCATCCTTTTACTGCCTGGAATATTTCCCTTCTCAGAT GCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAGGAAGT GCCTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGAT GAATCCTCTCATCGACCAATACCTGTATTACCTGAACAGAACTCAAAAT CAGTCCGGAAGTGCCCAAAACAAGGACTTGCTGTTTAGCCGTGGGTCT CCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACCTGGACCTGTT ATCGGCAGCAGCGCTTTCTAAAAACAAAACAGACAACAACAACAGC AATTTTACCTGGACTGGTGCTTCAAAATATAACCTCAATGGGCGTGAAT CCATCATCAACCCTGGCACTGCTATGGCCTCACACAAAGACGACGAAG ACAAGTTCTTTCCCATGAGCGGTGTCATGATTTTTGGAAAAGAGAGCGC CGGAGCTTCAAACACTGCATTGGACAATGTCATGATTACAGACGAAGA GGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGT GGCAGTCAATTTCCAGAGCAGCAGCACAGACCCTGCGACCGGAGATGT GCATGCTATGGGAGCATTACCTGGCATGGTGTGGCAAGATAGAGACGT GTACCTGCAGGGTCCCATTGTTGGGCCAAAATTCCTCACACAGATGGACA CTTTCACCCGTCTCCTCTTATGGGCGGCTTTGGACTCAAGAACCCGCCT CCTCAGATCCTCATCAAAAACACGCCTGTTCTGCGAATCCTCCGGCGG AGTTTTCAGCTACAAAGTTTGCTTCATTCACCCAATACTCCACAGG ACAAGTGAGTGTGGAAATTGAATGGGAGCTGCAGAAAGAAAACAGCA AGCGCTGGAATCCCGAAGTGCAGTACACATCCAATTATGCAAAATCTG CCAACGTTGATTTTACTGTGGACAACAATGGACTTTATACTGAGCCTCG CCCCATTGGCACCCGTTACCTTACCCGTCCCCTGTAA *VP2 start codon is in upper case, bold and single underlined, and G was mutated to C in the VP2 start codon.
SEQ ID NO:89 AAV1 VP2 N-term B3.1	ATGGGTTGTCCCCCTGGCTACAAGAGCGGCGTGGACTGCTCCCCG GGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGA <u>TACGGGGGATACGGCGCAGACTCGGGAGTGGGATCGgtggagcgggg</u> <u>gt</u>ACCGCTCCTGGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGC CAGACTCCTCCTCGGGCATCGGCAAGACAGGCCAGCAGCCCGCTAAAA AGAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTCCCCGATC CACAACCTCTCGGAGAACCTCCAGCAACCCCCGCTGCTGTGGGACCTA CTACAATGGCTTCAGGCGGTGGCGCACC AATGGCAGACAATAACGAAG GCGCCGACGGAGTGGGTAAATGCCTCAGGAAATTGGCATTGCGATTCCA CATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGACCTGGGCCCT TGCCACCTACAATAACCACCTCTACAAGCAAATCTCCAGTGCTTCAAC GGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGG GTATTTTGATTTCAACAGATTCCACTGCCACTTTTACCACGTGACTGG CAGCGACTCATCAACAACAATTGGGGATTCCGGCCCAAGAGACTCAAC TTCAAACCTTTCAACATCCAAGTCAAGGAGGTCACGACGAATGATGGC GTCACAACCATCGCTAATAACCTTACCAGCACGGTTCAAGTCTTCTCGG ACTCGGAGTACCAGTTCCTGACGTCTCGGCTCTGCGCACCAGGGCTG CCTCCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCGCAATACGGCTAC CTGACGCTCAACAATGGCAGCCAAGCCGTGGGACGTTTCATCCTTTTACT GCCTGGAATATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTAC CTTCAGCTACACCTTTGAGGAAGTGCTTTCCACAGCAGCTACGCGCAC AGCCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATACCTG TATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAG GACTTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCA

Sequence identifier and name	Sequence
	AAAACTGGCTACCTGGACCCTGTTATCGGCAGCAGCGCTTTCTAAAA CAAAAACAGACAACAACAACAGCAATTTTACCTGGACTGGTGCTTCAA AATATAACCTCAATGGGCGTGAATCCATCATCAACCCTGGCACTGCTAT GGCCTCACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGT CATGATTTTTTGAAAAAGAGAGCGCCGGAGCTTCAAACACTGCATTGGA CAATGTCATGATTACAGACGAAGAGGAAATTAAGCCACTAACCTGT GGCCACCGAAAGATTTGGGACCGTGGCAGTCAATTTCCAGAGCAGCAG CACAGACCCTGCGACCGGAGATGTGCATGCTATGGGAGCATTACCTGG CATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGGC CAAAATTCCCTCACAGATGGACACTTTTACCCGCTCTCTCTTATGGGC GGCTTTGGACTCAAGAACCCTCCTCAGATCCTCATCAAAAAACAG CCTGTTCTGCGAATCCTCCGGCGGAGTTTTAGCTACAAAGTTTGCTT CATTATCACCCAATACTCCACAGGACAAGTGAGTGTGGAAATTGAAT GGGAGCTGCAGAAAGAAAAACAGCAAGCGCTGGAATCCCGAAGTGCAG TACACATCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACA ACAATGGACTTTATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTAC CCGTCCCCTGTAA *B1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:90 AAVTT Del Vp2	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTG AAGGAATAAGACAGTGGTGGAAGCTCAAACCTGGCCCACCACCACCAA AGCCCGCAGAGCGGCATAAGGACGACAGCAGGGGTCTTGTGCTTCTCTG GGTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGGGAGAGCCGG TCAACGAGGCAGACGCGCGGCCCTCGAGCAGCACAAGCCTACGACC GGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACCACGCCG ACGCGGAGTTTCAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGGCA ACCTCGGACGAGCAGTCTTCCAGGCGAAAAAGAGGATCCTTGAACCTC TGGGCCTGGTTGAGGAACCTGTAAAG<u>ACCG</u>CTCCGGGAAAAAAGAGGC CGGTAGAGCACTCTCTGCCGAGCCAGACTCCTCCTCGGGAACCGGAA AGAGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTG GAGACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAG CAGCCCCCTCTGGTCTGGGAATAATACGATGGCTAGCGGCAGTGGCG CACCAATGGCAGACAATAACGAGGGCGCCGACGGAGTGGGTAAATTCCT CGGGAAATTGGCATTGCGATTCCACATGGATGGGCGACAGAGTCATCA CCACCAGCACCCGAACCTGGGCCCTGCCCACCTACAACAACCACCTCT ACAAAACAATTTCCAGCCAATCAGGAGCCTCGAACGACAATCACTACT TTGGCTACAGCACCCCTTGGGGGTATTTGACTTCAACAGATTCCACTG CCACTTTTACCACGTGACTGGCAAAGACTCATCAACAACAACCTGGGG ATTCCGACCCAAGAGACTCAGCTTCAAGCTCTTTAACATTCAAGTCAAA GAGGTCACGCAGAATGACGGTACGACGACGATTGCCAATAACCTTACC AGCACGGTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCC TCGGCTCGGCGCATCAAGGATGCCTCCCGCCGTTCCAGCAGACGTCTT CATGGTGCCACAGTATGGATACCTCACCTGAACAACGGGAGTCAGGC AGTAGGACGCTCTTCATTTTACTGCCTGGAGTACTTTCTTCTCAGATG CTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAGGACGTTC CTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAA TCCTCTCATCGACCAGTACCTGTATTACTTGAGCAGAACAAACACTCCA AGTGGAACCACCAGATGTCAAGGCTTCAGTTTTCTCAGGCCGGAGCG AGTGACATTGGGACCAGTCTAGGAACTGGCTTCTGGACCCGTGTTACC GCCAGCAGCGAGTATCAAAGACAGCCGCGGATAACAACAACAGTGAC TACTCGTGAGCTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCT CTGGTGAATCCGGGCCCGGCCATGGCAAGCCACAAGGACGATGAAGA

Sequence identifier and name	Sequence
	AAAGTACTTTCCTCAGAGCGGGGTTCTCATCTTTGGGAAGCAAGACTCA GGCAAAACAAATGTGGACATTGAAAAGGTCATGATTACAGACGAAGA GGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGT ATCTACCAACCTCCAGAGCGGCAACACCCAAGCAGTACCAGCGATGT CAACACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGT GTACCTTCAGGGGCCCATCTGGGCAAAGATTCCACACACGGACGGACA TTTTACCCCTCTCCCTCATGGGTGGATTTCGGACTTAAACACCCCTCCTC CACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCAC CTTCAGTGCGGCAAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAAACAGCAA ACGCTGGAATCCCGAAATTCAGTACACTTCCAACATAACAAGTCTGTT AATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCAGAGCCTCGCC CCATTGGCACCAGATACCTGACTCGTAATCTGTAA *VP2 start codon is in upper case, bold and single underlined, and G was mutated to C in the VP2 start codon.
SEQ ID NO:91 AAVTT VP2 N-term B3.1	ATGGGGTTGTCCCCCTGGCTACAAGAGCGGGCGTGGACTGCTCCCCG GGATCGGAGTGCAAATGGGGATGCTACGCCGTGGATGGACGGAGA TACGGGGGATACGGCGCAGACTCGGGAGTGGGATCGgtggaggcggtg gtACCGCTCCGGGAAAAAAGAGGCCGCTAGAGCACTCTCCTGCCGAGCC AGACTCCTCCTCGGGAACCGGAAAGAGCGGCCAGCAGCCTGCAAGAA AAAGATTGAATTTTGGTCAGACTGGAGACGCAGACTCAGTACCTGACC CCCAGCCTCTCGGACAGCCACCAGCAGCCCCCTCTGGTCTGGGAACATA ATACGATGGCTAGCGGCGAGTGGCGCACCAATGGCAGACAATAACGAG GGCGCCGACGGAGTGGGTAATTCCTCGGGAAATTGGCATTGCGATTCC ACATGGATGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCCACCTACAACAACCACTCTACAAACAAATTTCCAGCCAATCA GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGG TATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGC AAAGACTCATCAACAACAACCTGGGGATTCCGACCCAAGAGACTCAGCT TCAAGCTCTTTAACATTCAAGTCAAAGAGGTACGCAGAAATGACGGTA CGACGACGATTGCCAATAACCTTACCAGCACGGTTCAGGTGTTTACTGA CTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATG CCTCCCGCCGTTCCAGCAGACGTCTTCATGGTGCCACAGTATGGATAC CTACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACT GCCTGGAGTACTTTCTTCTCAGATGCTGCGTACCGGAAACAACCTTAC CTTCAGCTACACTTTTGAGGACGTTCTTTCCACAGCAGCTACGCTCAC AGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTACCTGT ATTACTTGAGCAGAACAACAACCTCCAAGTGGAACCAACACGATGTCAA GGCTTCAGTTTTTCTCAGGCCGGAGCGAGTGACATTTCGGGACCAGTCTA GGAACCTGGCTTCTGGACCCTGTTACCGCCAGCAGCGAGTATCAAAGA CAGCCGCGGATAACAACAACAGTGACTACTCGTGGACTGGAGCTACCA AGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCCGGCCA TGGCAAGCCACAAGGACGATGAAGAAAAGTACTTTCCTCAGAGCGGGG TTCTCATCTTTGGGAAGCAAGACTCAGGCAAAACAAATGTGGACATTG AAAAGGTCATGATTACAGACGAAGAGGAAATCAGGACAACCAATCCC GTGGCTACGGAGCAGTATGGTTCTGTATCTACCAACCTCCAGAGCGGC AACACCCAAGCAGCTACCAGCGATGTCAACACACAAGGCGTTCTTCCA GGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGG GCAAAGATTCCACACACGGACGGACATTTTCACCCCTCTCCCTCATGG GTGGATTTCGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACAC CCCGGTACCTGCGAATCCTTCGACCACCTTCAGTGCGGCAAAGTTTGCT

Sequence identifier and name	Sequence
	TCCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAGATCGAG TGGGAGCTGCAGAAGGAAAACAGCAAACGCTGGAATCCCGAAATTCA GTACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGAC ACTAATGGCGTGTATTTCAGAGCCTCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA *B1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:92 AAV1 VP2 N-term F3.1	<u>ATGCGATGCCCGGACGGAAGCATATGGAGGTGCGGGGATGTGGAC</u> <u>GGAGCGACGTGTTCAATTGCGACCATGCGCCTGTTGGGTTTCAAG</u> <u>TCATCAGGTGGCTGGTACTGTAAGCCACCGGCGACT</u><u>ggtggaggcggggt</u> ACCGCTCCTGGAAAGAAACGTCCTGGTAGAGCAGTCGCCACAAGAGCCA GACTCCTCCTCGGGCATCGGCAAGACAGGCCAGCAGCCCGCTAAAAAG AGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTCCCCGATCCAC AACCTCTCGGAGAACCTCCAGCAACCCCCGCTGCTGTGGGACCTACTA CAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCG CCGACGGAGTGGGTAATGCCTCAGGAAATTGGCATTGCGATTCCACAT GGCTGGGCGACAGAGTCATCACCACCAGCACCCGCACCTGGGCCTTGC CCACCTACAATAACCACCTCTACAAGCAAATCTCCAGTGCTTCAACGG GGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGGGT ATTTTGATTTCAACAGATTCCACTGCCACTTTTACCACGTGACTGGCA GCGACTCATCAACAACAATTGGGGATTCCGGCCCCAAGAGACTCAACTT CAAACCTTCAACATCCAAGTCAAGGAGGTCACGACGAATGATGGCGT CACAACCATCGCTAATAACCTTACCAGCACGGTTCAAGTCTTCTCGGAC TCGGAGTACCAGCTTCCGTACGTCTCGGCTCTGCGCACCAAGGGCTGCC TCCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCGCAATACGGCTACCT GACGCTCAACAATGGCAGCCAAGCCGTGGGACGTTTCATCCTTTTACTGCG CTGGAATATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTTACCT TCAGCTACACCTTTGAGGAAGTGCCTTTCCACAGCAGCTACGCGCACA GCCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATACCTGT ATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCCAAAACAAGG ACTTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCAA AAACTGGCTACCTGGACCCTGTTATCGGCAGCAGCGCGTTTCTAAAAC AAAAACAGACAACAACAACAGCAATTTTACCTGGACTGGTGCTTCAAA ATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACTGCTATG GCCTCACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTC ATGATTTTTGGAAAAGAGAGCGCCGGAGCTTCAAACACTGCATTGGAC AATGTCATGATTACAGACGAAGAGGAAATTAAGCCACTAACCCGTGTG GCCACCGAAAGATTTGGGACCGTGGCAGTCAATTTCCAGAGCAGCAGC ACAGACCCTGCGACCGGAGATGTGCATGCTATGGGAGCATTACCTGGC ATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTGTTGGGCC AAAATTCCTCACACAGATGGACACTTTACCCGTCTCCTCTTATGGGCG GCTTTGGACTCAAGAACCCGCCTCCTCAGATCCTCATCAAAAACACGCC TGTTCTGCGAATCCTCCGGCGGAGTTTTTCAGCTACAAAGTTTGCTTCA TTCATCACCCAATACTCCACAGGACAAGTGAGTGTGGAAATTGAATGG GAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGACAGTA CACATCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAC AATGGACTTTTATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTACCC GTCCCCTGTAA *F1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.

Sequence identifier and name	Sequence
SEQ ID NO:93 AAVTT VP2 N-term G3.1	<u>ATGTCGTGTCCGGATGGATACACATACCGCTTCGGTAGTGT</u><u>TTGTA</u> <u>CAGCACGCTGGCGTAATGACTGCTTTCGCGATGCGGATTATGATTA</u> <u>TCAGCGAAGATGCTACACCACTTCCCTTCCACAGACG</u><u>ggtggaggcggg</u> <u>gt</u>ACCGCTCCGGGAAAAAAGAGGCCGGTAGAGCACTCTCCTGCCGAGCC AGACTCCTCCTCGGGAACCGGAAAGAGCGGCCAGCAGCCTGCAAGAA AAAGATTGAATTTTGGTCAGACTGGAGACGCAGACTCAGTACCTGACC CCCAGCCTCTCGGACAGCCACCAGCAGCCCCCTCTGGTCTGGGAACTA ATACGATGGCTAGCGGCAGTGGCGCACCAATGGCAGACAATAACGAG GGCGCCGACGGAGTGGGTAATTCCTCGGGAATTGGCATTGCGATTCC ACATGGATGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCACCTACAACAACCACCTCTACAACAAATTTCAGCCCAATCA GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGG TATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGC AAAGACTCATCAACAACAACCTGGGGATTCCGACCCAAGAGACTCAGCT TCAAGCTCTTTAACATTCAAGTCAAAGAGGTACGCAGAATGACGGTA CGACGACGATTGCCAATAACCTTACCAGCACGGTTCAGGTGTTTACTGA CTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATG CCTCCCGCCGTTCCAGCAGACGTCTTCATGGTGCCACAGTATGGATAC CTCACCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACT GCCTGGAGTACTTTCTTCTCAGATGCTGCGTACCGGAAACAACCTTTAC CTTCAGCTACACTTTTGAGGACGTTCTTTCCACAGCAGCTACGCTCAC AGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTACCTGT ATTACTTGAGCAGAACAACACTCCAAGTGGAACCACCACGATGTCAA GGCTTCAGTTTTCTCAGGCCGGAGCGAGTGACATTCTGGGACCACTCA GGAAC TGGCTTCTG GACCTGTTACCGCCAGCAGCGAGTATCAAAGA CAGCCGCGGATAACAACAACAGTGACTACTCGTGGACTGGAGCTACCA AGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCGGCCA TGGCAAGCCACAAGGACGATGAAGAAAAGTACTTTCTCAGAGCGGGG TTCTCATCTTTGGGAAGCAAGACTCAGGCAAAACAAATGTGGACATTG AAAAGGTCATGATTACAGACGAAGAGGAAATCAGGACAACCAATCCC GTGGCTACGGAGCAGTATGGTTCTGTATCTACCAACCTCCAGAGCGGC AACACCCAAGCAGCTACCAGCGATGTCAACACACAAGGCGTTCTTCCA GGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGGCCCATCTGG GCAAAGATTCCACACACGGACGGACATTTTCACCCCTCTCCCCTCATGG GTGGATTCTGGACTTAAACACCCTCCTCCACAGATTCTCATCAAGAACAC CCCGGTACCTGCGAATCCTTCGACCACCTTCAGTGCGGCAAAGTTTGCT TCCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAGATCGAG TGGGAGCTGCAGAAGGAAAACAGCAAACGCTGGAATCCCGAAATTCA GTACACTTCCAATAACAAGTCTGTTAATGTGGACTTTACTGTGGAC ACTAATGGCGTGTATTCAGAGCCTCGCCCCATTGGCACCAGATACCTGA CTCGTAATCTGTAA *G1 knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.
SEQ ID NO:94 AAV8 VP2 N-term E3.1	<u>ATGACTTGCCCGGATGGGTACACACGGACCAACTACTACTGT</u><u>CGGA</u> <u>GGGACGGCTGCGGCTCATGGTGCAACGGTGCGGAGCGTCAGCAGC</u> <u>CGTGTATAAGAGGACCCTGTTGTTGTGACCTTACT</u><u>ggtggaggcgggggtA</u> CCGCTCCTGGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTCTC CAGACTCCTCTACGGGCATCGGCAAGAAAGGCCAACAGCCCCGCCAGAA AAAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTTCCAGACC CTCAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGACCTA ATACAATGGCTGCAGGCGGTGGCGCACCAATGGCAGACAATAACGAA

Sequence identifier and name	Sequence
	GGCGCCGACGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCC ACATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCC CTGCCCACCTACAACAACCACCTCTACAAGCAAATCTCCAACGGGACA TCGGGAGGAGCCACCAACGACAACACCTACTTTCGGCTACAGCACCCCC TGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTTTCACCACGTG ACTGGCAGCGACTCATCAACAACAACCTGGGGATTCCGGCCCCAAGAGAC TCAGCTTCAAGCTCTTCAACATCCAGGTCAAGGAGGTCACGCAGAATG AAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCATCCAGGTGT TTACGGACTCGGAGTACCAGCTGCCGTACGTTCTCGGCTCTGCCCCACCA GGGCTGCCTGCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCCCAGTAC GGCTACCTAACACTCAACAACGGTAGTCAGGCCGTGGGACGCTCCTCC TTCTACTGCCTGGAATACTTTCCTTCGCAGATGCTGAGAACC GGCAACA ACTTCCAGTTTACTTACACCTTCGAGGACGTGCCTTTCCACAGCAGCTA CGCCCACAGCCAGAGCTTGGACCGGCTGATGAATCCTCTGATTGACCA GTACCTGTACTACTTGTCTCGGACTCAAACAACAGGAGGCACGGCAAA TACGCAGACTCTGGGCTTCAGCCAAGGTGGGCCTAATAAATGGCCAA TCAGGCAAAGAACTGGCTGCCAGGACCCTGTTACCGCCAACAACGCGT CTCAACGACAACCGGGCAAAACAACAATAGCAACTTTGCCTGGACTGC TGGGACCAAATACCATCTGAATGGAAGAAATTCATTGGCTAATCCTGG CATCGCTATGGCAACACACAAAGACGACGAGGAGCGTTTTTTTCCCAG TAACGGGATCCTGATTTTTTGGCAAACAAAATGCTGCCAGAGACAATGC GGATTACAGCGATGTCATGCTCACCAGCGAGGAAGAAATCAAACCAC TAACCCTGTGGCTACAGAGGAATACGGTATCGTGGCAGATAACTTGCA GCAGCAAAACACGGCTCCTCAAATTGGAACTGTCAACAGCCAGGGGGC CTTACCCGGTATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTCC CATCTGGGCCAAGATTCCCTCACACGGACGGCAACTTCCACCCGTCTCCG CTGATGGGCGGCTTTGGCCTGAAACATCCTCCGCCTCAGATCCTGATCA AGAACACGCCTGTACCTGCGGATCCTCCGACCACCTTCAACCAGTCAA AGCTGAACTCTTTCATCACGCAATACAGCACCGGACAGGTCAGCGTGG AAATTGAATGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCC GAGATCCAGTACACCTCCAATACTACAAATCTACAAGTGTGGACTTTG CTGTTAATACAGAAGGCGTGTACTCTGAACCCCGCCCCATTGGCACCCG TTACCTCACCCGTAATCTGTAA *El knob is indicated in upper case, bold and single underlined. GGGGG linker is indicated in lower case, bold and single underlined.

CLAIMS

1. An engineered capsid protein comprising an AAV capsid protein and an antibody fragment, wherein the antibody fragment comprises a bovine ultralong CDR-H3, or a portion thereof.
2. An engineered capsid protein comprising an AAV capsid protein and an antibody fragment, optionally according to claim 1, wherein the antibody fragment comprises a knob domain of an ultralong CDR-H3, or a portion thereof.
3. The engineered capsid protein according to claim 2, wherein the antibody fragment does not comprise a stalk of an ultralong CDR-H3.
4. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment binds an antigen.
5. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is 5 amino acids in length or more, 10 amino acids in length or more, 15 amino acids in length or more, 20 amino acids in length or more, 25 amino acids in length or more, 30 amino acids in length or more, 35 amino acids in length or more, 40 amino acids in length or more, 45 amino acids in length or more, 50 amino acids in length or more, 55 amino acids in length or more, or 60 amino acids in length or more.
6. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is up to 69 amino acids in length.
7. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is between 5 and 55, or between 15 and 50, or between 20 and 45 or between 25 and 40 amino acids in length.
8. The capsid protein according to any one of the preceding claims, wherein the antibody fragment comprises a (Z₁) X₁ C X₂ motif at its N-terminal extremity, wherein:

Z₁ is present or absent, and when Z₁ is present, Z₁ represents 1 amino acid or 2, 3, 4, or 5 independently selected amino acids;

X₁ is any amino acid residue, preferably selected from the list consisting of Serine, Threonine, Asparagine, Alanine, Glycine, Proline, Histidine, Lysine, Valine, Arginine, Isoleucine, Leucine, Phenylalanine and Aspartic acid; and,

C is cysteine; and,

X₂ is an amino acid selected from the list consisting of Proline, Arginine, Histidine, Lysine, Glycine and Serine.

9. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment comprises a sequence which is a variant of a naturally occurring sequence.
10. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment further comprises at least one bridging moiety between two amino acids, optionally wherein the bridging moiety is a disulphide bond.
11. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is fully bovine.
12. The engineered capsid protein according to any one of claims 1 to 10, wherein the antibody fragment is chimeric.
13. The engineered capsid protein according to any one of the preceding claims, wherein the AAV capsid protein comprises a naturally occurring, or a variant or an artificial AAV sequence or a combination thereof.
14. The engineered capsid protein according to any one of the preceding claims, wherein the AAV capsid protein comprises a sequence of an AAV selected from the group consisting of AAV1, AAV2, AAV true type (AAV-TT), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh10, AAV11, AAV12, or AAV13, or a combination thereof.
15. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is inserted within the AAV capsid protein, optionally via a linker.
16. The engineered capsid protein according to claim 15, wherein the antibody fragment is inserted within the AAV capsid protein via one linker, wherein optionally the linker is genetically fused to the antibody fragment, optionally at its C-terminal end.
17. The engineered capsid protein according to claim 15, wherein the antibody fragment is inserted within the AAV capsid protein, via at least two linkers.
18. The engineered capsid protein according to claim 17, wherein at least one linker is fused, optionally genetically, to the N-terminal end of the antibody fragment, and at least one linker is fused, optionally genetically, to the C-terminal end of the antibody fragment.

19. The engineered capsid protein according to any one of the preceding claims, wherein the AAV capsid protein is a VP1, a VP2, or a VP3.
20. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is inserted within the common VP3 region of the AAV capsid protein, optionally, within the GH loop of the common VP3 region.
21. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is inserted within the variable region VR-IV of the AAV capsid protein.
22. The engineered capsid protein according to any one of the preceding claims, wherein the antibody fragment is inserted within the AAV capsid protein after amino acid residue Gly455 of AAV9 with reference to SEQ ID NO: 40 (or a corresponding amino acid residue of another AAV).
23. The engineered capsid protein according to any one of claims 1-20, wherein the antibody fragment is inserted within the variable region VR-V of the AAV capsid protein.
24. The engineered capsid protein according to claim 23, wherein the antibody fragment is inserted within the AAV capsid protein after amino acid residue Asn498 of AAV9 with reference to SEQ ID NO:40, or a corresponding amino acid residue of another AAV.
25. The engineered capsid protein according to any one of claims 1-20, wherein the antibody fragment is inserted within the variable region VR-VIII of the AAV capsid protein.
26. The engineered capsid protein according to claim 25, wherein the antibody fragment is inserted within the AAV capsid protein after amino acid residue Gln588 of AAV9 with reference to SEQ ID NO:40, or a corresponding amino acid residue of another AAV.
27. The engineered capsid protein according to any one of claims 1-19, wherein the AAV capsid protein is VP1 or VP2 and wherein the antibody fragment is inserted within the VP1/VP2 common region of the capsid protein.
28. The engineered capsid protein according to any one of claims 1-19 or 27, wherein the antibody fragment is inserted within the N-terminus of the VP2 capsid protein.
29. The engineered capsid protein according to claim 28, wherein the antibody fragment is inserted within the N-terminus of the VP2 capsid protein of AAV9 before amino

acid residue Thr1 with reference to SEQ ID NO:41, or a corresponding amino acid in another AAV.

30. An AAV capsid comprising an engineered capsid protein according to any one of the preceding claims.
31. An AAV capsid comprising two engineered capsid proteins according to any one of the preceding claims, wherein the first and second engineered capsid proteins are respectively VP1 and VP2, or VP1 and VP3, or VP2 and VP3.
32. An AAV capsid comprising three engineered capsid proteins according to any one of the preceding claims, wherein the first, second and third engineered capsid proteins are respectively VP1, VP2 and VP3.
33. A nucleic acid encoding an engineered capsid protein according to any one of claims 1 to 29, or a capsid according to any one of claims 30 to 32.
34. A vector comprising the nucleic acid according to claim 33.
35. A recombinant adeno-associated virus (rAAV) particle comprising the engineered capsid protein according to any one of claims 1 to 29, or the capsid according to any one of claims 30 to 32, or the nucleic acid according to claim 33 or the vector according to claim 34, and a transgene.
36. The rAAV particle according to claim 35, wherein the transgene encodes a peptide, a polypeptide or a nucleic acid molecule, optionally wherein the nucleic acid molecule is a small interfering RNA (siRNA), small or short hairpin RNA (shRNA), micro RNA (miRNA).
37. A host cell comprising the nucleic acid according to claim 33 or a vector according to claim 34 and/or which produces a viral particle according to claim 35 or 36.
38. A method for producing an AAV particle comprising an engineered capsid protein according to any one of claims 1 to 29, and a transgene, said method comprising:
 - a) providing a first vector comprising a first nucleotide sequence encoding an AAV capsid protein, and a second nucleotide sequence encoding the antibody fragment; wherein the first nucleotide sequence and the second nucleotide sequence are genetically fused, optionally via a nucleotide sequence coding for a linker;
 - b) providing a second vector comprising the transgene;
 - c) providing a third, Helper vector;
 - d) transfecting a host cell with the first, second and third vector;
 - e) recovering the AAV particle from the host cell.

39. A vector according to claim 34, or an AAV particle according to claim 35 or 36 or obtained according to claim 38, for use as a medicament.