

Summary

In one aspect, there is provided an AAV capsid protein having an amino acid sequence which:

- (i) has the sequence of SEQ ID NO: 3; or (ii) has at least 99% identity to the sequence of SEQ ID NO: 3 and is a hybrid capsid comprising regions of the AAV3B and AAV8 capsids. In one example, there is provided an AAV capsid protein having an amino acid sequence which has at least 98% identity to SEQ ID NO: 3 or at least 94% identity to SEQ ID NO: 4.

The inventors have surprisingly found that the novel capsid protein produces a capsid which provides a higher transduction rate than currently used AAV vectors. Further, it has been found that the prevalence of antibodies to the capsid in patients is lower than for some currently used AAV vectors.

In some examples, the amino acid sequence has at least 95% identity to the sequence of SEQ ID NO: 4. In particular examples, the amino acid sequence has at least 96% identity to the sequence of SEQ ID NO: 4. In further examples, the amino acid sequence has at least 97% identity to the sequence of SEQ ID NO: 4. In some examples, the amino acid sequence has at least 98% identity to the sequence of SEQ ID NO: 4. In other examples, the amino acid sequence has at least 99% identity to the sequence of SEQ ID NO: 4. In particular examples, the amino acid sequence has the sequence of SEQ ID NO: 4.

In some examples, the amino acid sequence has at least 98.5% identity to the sequence of SEQ ID NO: 3, preferably at least 99% identity to the sequence of SEQ ID NO: 3 and more preferably at least 99.5% identity to the sequence of SEQ ID NO: 3. In particular examples, the amino acid sequence has the sequence of SEQ ID NO: 3.

In some examples, the amino acid sequence has identity to the sequence of SEQ ID NO: 3. In other examples, the amino acid sequence has identity to the sequence of SEQ ID NO: 4. In preferred examples, the amino acid sequence has the sequence of SEQ ID NO: 4. More preferably, the amino acid sequence has the sequence of SEQ ID NO: 3.

The capsid protein having the amino acid sequence defined above is a functional capsid protein which can form a functional capsid, along with other necessary capsid proteins. A functional capsid is one which can enclose genetic material, enter a cell and transduce the cell with the genetic material. It would be well within the capabilities of a skilled person to determine whether a capsid is functional. For example, the experiments described below in

the detailed description of the invention can be used to determine whether a capsid can successfully transduce a cell.

SEQ ID NO: 3 is generated by cloning a 146 amino acid region encompassing the AAV8 VP1 region upstream of the VP2 and VP3 domains of the AAV3B capsid. SEQ ID NO: 4 is generated by cloning the AAV5 VP1 region upstream of the AAV3B VP2 and VP3 regions. Further details of AAV serotypes and capsid proteins is provided below.

A second aspect of the invention provides an AAV capsid comprising the AAV capsid protein described above. An AAV capsid is a protein shell made up of VP1, VP2 and VP3 proteins and which encapsidates (or encloses) the genetic material of the virus.

In a third aspect, there is provided a viral particle comprising the AAV capsid protein described above. The viral particle comprises the AAV capsid and the genetic material of the virus. In a fourth aspect, there is provided a nucleotide sequence which encodes for the capsid protein described above.

In one example, there is provided a recombinant AAV (rAAV) vector comprising a nucleotide sequence which encodes for the amino acid sequence described above. This means that the vector contains a nucleotide sequence encoding for a functional capsid protein.

Preferably the vector further comprises a promoter such that the nucleotide sequence is expressible. Preferably the vector comprises the AAV2 p40 viral promoter, which is constitutionally active in most mammalian cell types.

Accordingly, the present invention provides gene delivery vectors based on adeno-associated virus (AAV), which normally infects humans (e.g., serotypes 1, 2, 3A, 3B, 4, 5, and 6) or primates (e.g., serotypes 1 and 4). Further information on this virus is described in Kenneth I. Berns, "Parvoviridae: The Viruses and Their Replication," Chapter 69 in Fields Virology (3d Ed. 1996).

The genomic organization of all known AAV serotypes is very similar. The genome of AAV is a linear, single-stranded DNA molecule that is less than about 5,000 nucleotides (nt) in length. Inverted terminal repeats (ITRs) flank the unique coding nucleotide sequences for the

non-structural replication (Rep) proteins and the structural (VP) proteins. The VP proteins (VP1, VP2 and VP3) form the capsid. The terminal 145 nt are self-complementary and are organized so that an energetically stable intramolecular duplex forming a T-shaped hairpin may be formed. These hairpin structures function as an origin for viral DNA replication, serving as primers for the cellular DNA polymerase complex. Following wild type (wt) AAV infection in mammalian cells the Rep genes (i.e. encoding Rep78 and Rep52 proteins) are expressed from the P5 promoter and the P19 promoter, respectively, and both Rep proteins have a function in the replication of the viral genome. A splicing event in the Rep ORF results in the expression of actually four Rep proteins (i.e. Rep78, Rep68, Rep52 and Rep40). However, it has been shown that the unspliced mRNA, encoding Rep78 and Rep52 proteins, in mammalian cells are sufficient for AAV vector production. Also in insect cells the Rep78 and Rep52 proteins suffice for AAV vector production.

In an AAV suitable for use as a gene therapy vector, the vector genome typically comprises a nucleic acid of interest to be packaged for delivery to a target cell. According to this particular example, the nucleic acid is located between the viral ITRs at either end of the vector genome. It is possible for an AAV genome to function with only one ITR. Thus, in a gene therapy vector of the invention based on AAV, the vector genome is flanked by at least one ITR, but, more typically, by two AAV ITRs (generally with one either side of the vector genome, i.e. one at the 5' end and one at the 3' end). There may be intervening sequences between the nucleic acid in the vector genome and one or more of the ITRs.

In the context of the invention "at least one ITR" is understood to mean a palindromic sequence, comprising mostly complementary, symmetrically arranged sequences also referred to as "A," "B," and "C" regions. The ITR functions as an origin of replication, a site having a "cis" role in replication, i.e., being a recognition site for trans-acting replication proteins such as e.g. Rep 78 (or Rep68) which recognize the palindrome and specific sequences internal to the palindrome. One exception to the symmetry of the ITR sequence is the "D" region of the ITR. It is unique (not having a complement within one ITR). Nicking of single-stranded DNA occurs at the junction between the A and D regions. It is the region where new DNA synthesis initiates. The D region normally sits to one side of the palindrome and provides directionality to the nucleic acid replication step. An AAV replicating in a mammalian cell typically has two ITR sequences. It is, however, possible to engineer an ITR so that binding

sites are on both strands of the A regions and D regions are located symmetrically, one on each side of the palindrome. On a double-stranded circular DNA template (e.g., a plasmid), the Rep78- or Rep68-assisted nucleic acid replication then proceeds in both directions and a single ITR suffices for parvoviral replication of a circular vector. Thus, one ITR nucleotide
5 sequence can be used in the context of the present invention. Preferably, however, two or another even number of regular ITRs are used. Most preferably, two ITR sequences are used. For safety reasons it may be desirable to construct an AAV vector that is unable to further propagate after initial introduction into a cell. Such a safety mechanism for limiting undesirable vector propagation in a recipient may be provided by using AAV with a chimeric
10 ITR as described in US 2003148506.

Those skilled in the art will appreciate that the viral Rep protein(s) used for producing an AAV vector of the invention may be selected with consideration for the source of the viral ITRs. For example, the AAV5 ITR typically interacts more efficiently with the AAV5 Rep
15 protein, although it is not necessary that the serotype of ITR and Rep protein(s) are matched.

The ITR(s) used in the invention are typically functional, i.e. they may be fully resolvable and are AAV sequences, with serotypes 1, 2, 3, 4, 5 or 6 being preferred. Resolvable AAV ITRs according to the present invention need not have a wild-type ITR sequence (e. g., a wild-type
20 sequence may be altered by insertion, deletion, truncation or missense mutations), as long as the ITR mediates the desired functions, e.g., virus packaging, integration, and/or provirus rescue, and the like.

In further preferred examples, the AAV cap genes and AAV rep genes are deleted from the
25 template genome (and thus from the virion DNA produced therefrom). This configuration maximizes the size of the nucleic acid sequence(s) that can be carried by the AAV capsid.

AAV sequences that may be used in the present invention for the production of AAV gene therapy vectors can be derived from the genome of any AAV serotype. Generally, the AAV
30 serotypes have genomic sequences of significant homology at the amino acid and the nucleic acid levels, provide an identical set of genetic functions, produce virions which are essentially physically and functionally equivalent, and replicate and assemble by practically identical mechanisms. For the genomic sequence of the various AAV serotypes and an overview of the

The vector may comprise other elements to allow the functional therapeutic protein to be expressed. Such elements are well known to a person skilled in the art.

Preferably, the nucleic acids and amino acid sequences described above are isolated.

5

It would be well with the capabilities of a skilled person to produce the nucleic acid molecules and amino acid sequences described above. This could be done, for example, using chemical synthesis of a given sequence.

10 The invention also relates to a method of transferring a nucleic acid of interest into a mammal comprising introducing a recombinant AAV vector into a mammal, the recombinant AAV vector comprising a gene of interest which is encapsidated into a capsid comprising the capsid protein described above.

15 In one aspect, the invention also provides a host cell comprising the capsid protein described above or the nucleotide sequence described above.

As used herein, the term "host" refers to organisms and/or cells which harbour a protein or a vector of the invention, as well as organisms and/or cells that are suitable for use in
20 expressing a recombinant gene or protein. It is not intended that the present invention be limited to any particular type of cell or organism. Indeed, it is contemplated that any suitable organism and/or cell will find use in the present invention as a host. A host cell may be in the form of a single cell, a population of similar or different cells, for example in the form of a culture (such as a liquid culture or a culture on a solid substrate), an organism or part thereof.

25

A host cell according to the invention may permit the expression of a therapeutic nucleic acid molecule. Thus, the host cell may be, for example, a bacterial, a yeast, an insect or a mammalian cell.

30 In another aspect, the invention provides a transgenic animal comprising cells comprising the capsid protein having described above or the nucleotide sequence described above. Preferably the animal is a non-human mammal, especially a primate. Alternatively, the animal may be a rodent, especially a mouse; or may be canine, feline, ovine or porcine.

In one aspect, the invention provides a pharmaceutical composition comprising the capsid protein of the present invention or a nucleotide sequence of the invention and one or more pharmaceutically acceptable excipients. The one or more excipients include carriers, diluents and/or other medicinal agents, pharmaceutical agents or adjuvants, etc.

5

The higher levels of gene transfer achieved with the sequences of the present invention represents an advance beyond the current state of the art. An improvement in potency has been observed which will permit gene transfer using lower vector doses, thus improving safety (especially reducing liver toxicity which is dose dependent) and at a lower cost.

10 Accordingly the sequences and vector of the present invention may have a clinical applicability to disorders such as congenital FVII deficiency, Gaucher's disease, OTC deficiency, Fabry's disease, glycogen storage diseases, α -1-antitrypsin deficiency, progressive familial intrahepatic cholestasis, Wilson's disease, Crigler Najjar syndrome and hepatocellular carcinoma amongst others.

15

Accordingly, the invention may provide a method of treating one or more of the diseases listed above comprising administering a therapeutically effective amount of a vector as described above to a patient suffering from said disease. Preferably, the patient is human.

20 When said disease is "treated" in the above method, this means that one or more symptoms of said disease are ameliorated. It does not mean that the symptoms of said disease are completely remedied so that they are no longer present in the patient, although in some methods, this may be the case. The method of treating results in one or more of the symptoms of said disease being less severe than before treatment.

25

A "therapeutically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired therapeutic result.

Further, the invention may provide the capsid protein as described above, or a vector as
30 described above for use in therapy, preferably for one of the afore-mentioned diseases.

In addition, the invention may provide for the use of the capsid protein as described above or a vector as described above in the manufacture of a medicament for treating one of the diseases as mentioned above.

5 In the description above, the term “identity” is used to refer to the similarity of two sequences. For the purpose of this invention, it is defined here that in order to determine the percent identity of two amino acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a amino acid for optimal alignment with a second amino acid sequence). The amino acid residues are then compared. When a
10 position in the first sequence is occupied by the same amino acid or amino acid residue as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = number of identical positions/total number of positions (i.e. overlapping positions) x 100). Preferably, the two
15 sequences are the same length.

A sequence comparison may be carried out over the entire lengths of the two sequences being compared or over fragment of the two sequences. Typically, the comparison will be carried out over the full length of the two sequences being compared. However, sequence identity
20 may be carried out over a region of, for example, about twenty, about fifty, about one hundred, about two hundred, about five hundred, about 1000 or about 2000 or more contiguous amino acid residues.

The skilled person will be aware of the fact that several different computer programs are
25 available to determine the homology or identity between two sequences. In preferred examples, the identity between two sequences is analysed using the software package Clone Manager Professional version 9 (preferably, version 9.4). This analysis tool is produced by Sci-Ed Software (Scientific & Educational Software, 11010 Lake Grove Blvd, Ste 100, PMB 122, Morrisville, NC 27560, USA - <http://www.scied.com/index.htm>). The settings used to
30 compare the sequences are preferably as follows: alignment: Global DNA alignment; parameters: both strands; scoring matrix: linear (mismatch 2, OpenGap 4, ExtGap 1). Alternatively methods such as Fast Scan –MaxScore and Fast Scan MaxQual can also be used with the same software and local settings.

CLAIMS

1. An AAV capsid protein having an amino acid sequence which:
 - (i) has the sequence of SEQ ID NO: 3; or
 - (ii) has at least 99% identity to the sequence of SEQ ID NO: 3 and is a hybrid
- 5 capsid comprising regions of the AAV3B and AAV8 capsids.
2. The AAV capsid protein of claim 1, wherein the amino acid sequence has at least 99.5% identity to the sequence of SEQ ID NO: 3.
- 10 3. The AAV capsid protein of claim 1 or claim 2, wherein the amino acid sequence has the sequence of SEQ ID NO. 3.
4. An AAV capsid comprising the capsid protein of any one of claims 1 to 3.
- 15 5. A viral particle comprising the capsid protein of any one of claims 1 to 3.
6. A nucleotide sequence which encodes for the capsid protein of any one of claims 1 to 3.
- 20 7. A host cell comprising the capsid protein of any one of claims 1 to 4 or the nucleotide sequence of claim 6.
8. A non-human transgenic animal comprising cells comprising the capsid protein of any one of claims 1 to 3 or the nucleotide sequence of claim 6.
- 25 9. A pharmaceutical composition comprising the capsid protein of any one of claims 1 to 3 or the nucleotide sequence of claim 6, and one or more pharmaceutically acceptable excipients.
- 30 10. The capsid protein of any one of claims 1 to 3 or the nucleotide sequence of claim 6 for use in therapy.