configuration, both vectors of the rAAV dual vector system require a promoter to generate two separate pre-mRNA molecules.

[11] For an optimal performance of mRNA-splicing and particularly mRNA *trans*-splicing, there is an unmet need for identification of optimized and experimentally validated strong ASS. In particular, there is a need for strong ASS regions, which can inter alia be used in the development of further AAV vectors, such as next-generation rAAV dual vector systems or AAVs delivering highly specific and efficient pre-mRNA *trans*-splicing molecules, which may be used in gene therapy.

SUMMARY

[12] The present disclosure complies with this need as described herein and as described in the Examples, Figures and Claims.

[13] In one aspect, the present disclosure provides a pre-mRNA *trans*-splicing molecule comprising, (i) an acceptor splice region, comprising (ia) a pyrimidine tract, wherein the pyrimidine tract comprises (iaa) 5 to 25 nucleotides; (iab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (ib) an acceptor splice site, (iba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and (ibb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; (ii) a nucleotide sequence of interest or a portion thereof; wherein the acceptor splice region is localized 3' or 5' to the nucleotide sequence of interest or the portion thereof; (iii) a binding domain targeting pre-mRNA, which is localized 3' or 5' to the nucleic acid sequence of interest or the portion thereof; and (iv) optionally a spacer sequence, wherein the spacer sequence is localized between the binding domain and the acceptor splice region. In one example the pre-mRNA *trans*-splicing molecule of claim 1 comprises (ii) the nucleotide sequence of interest or the portion thereof, wherein the acceptor splice region is localized 5' to the nucleotide sequence of interest or the portion thereof; (iii) a donor splice site, wherein the donor splice site is localized 3' to the nucleotide sequence of interest or the portion thereof; (iv) a first binding domain targeting pre-mRNA located 5' to the nucleotide sequence of interest or the portion thereof; (v) a second binding domain targeting pre-mRNA located 3' to the nucleotide sequence of interest or the portion thereof; (vi) optionally a first spacer sequence, wherein the first spacer is localized between the first binding domain and the acceptor splice region; and (vii) optionally a second spacer sequence, wherein the second spacer is localized between the second binding domain and the donor splice site.

[14] Preferably the acceptor splice region is localized 5' to the nucleotide sequence of interest or the portion thereof and the binding domain is localized 5' to acceptor splice region. The pre-mRNA *trans*-splicing molecule may further comprise a termination

sequence, preferably a polyA sequence. Preferably the pre-mRNA *trans*-splicing molecule according to the disclosure comprises a pyrimidine tract, wherein the 5 to 25 nucleotides of the pyrimidine tract comprise a sequence encoded by the sequence of TTTTTT or TCTTTT. Additionally or alternatively the sequence between the last pyrimidine of the pyrimidine tract and the acceptor splice site has less than 10, preferably less than 5, more preferably less than 3 bases. Additionally or alternatively the acceptor splice site has a sequence of CAGG.

[15] The acceptor splice region may comprise (a) 7 nucleotides 5' to the pyrimidine tract having at least 4 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotide is C; (b) 7 nucleotides 5' to the pyrimidine tract having at least 5 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotide is C; (c) 7 nucleotides 5' to the pyrimidine tract having at least 4 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotides are CAA; or (d) 7 nucleotides 5' to the pyrimidine tract having at least 5 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotides are CAAC. In one example the acceptor splice region is encoded by a nucleotide sequence which has the sequence of SEQ ID NO: 3 or 4 or which is at least 85 %, 90 %, 95 %, 97 %, 98 %, 99 % or 100 % identical to the sequence of SEQ ID NO: 3 or 4. Also provided is a DNA molecule, comprising a promoter and a sequence encoding the pre-mRNA *trans*-splicing molecule according to the disclosure, wherein the DNA molecule is preferably a vector or a plasmid. In one aspect, the present disclosure provides a DNA molecule comprising a promoter and a sequence encoding the pre-mRNA trans-splicing molecule as described herein.

[16] In another aspect the disclosure relates to a method for producing a nucleic acid sequence, the method comprising (A) providing a first nucleic acid sequence comprising one or more donor splice site sequences; (B) providing a second nucleic acid sequence, wherein the second nucleic acid sequence comprises (i) an acceptor splice region comprising (ia) a pyrimidine tract, wherein the pyrimidine tract comprises (iaa) 5 to 25 nucleotides; (iab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (ib) an acceptor splice site, (iba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and (ibb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; (C) cleaving the first nucleic acid sequence at the one or more donor splice site sequence(s) and cleaving the second nucleic acid sequence in the acceptor splice site, (D) ligating the first cleaved nucleic acid sequence to the second cleaved nucleic acid sequence, thereby obtaining a nucleic acid sequence. The first nucleic acid sequence further comprises a nucleotide sequence of interest or a portion thereof, wherein at least a portion of the nucleotide sequence of interest is located 5' to the donor splice site and wherein the second nucleic acid sequence further comprises a nucleotide sequence of interest or a portion thereof, wherein at least a portion of the nucleotide sequence of interest is located 3' to the acceptor splice region. Preferably the first and the second nucleic acid sequence are introduced into a host cell, preferably wherein the first and the second nucleic acid sequence

are a recombinant nucleic acid sequence.

[17] In another aspect of an example of the method of the disclosure, the method comprises step (A) introducing into a host cell a first nucleic acid sequence comprising a pre-mRNA *trans*-splicing molecule sequence or a nucleic acid sequence encoding said pre-mRNA *trans*-splicing molecule, wherein the first pre-mRNA *trans*-splicing molecule comprises from 5' to 3', (a) a 5' portion of a nucleotide acid sequence of interest; (b) a donor splice site; (c) optionally a spacer sequence; (d) a first binding domain; and (e) optionally a termination sequence, preferably a polyA sequence; step (B) introducing into the host cell a second nucleic acid sequence comprising a pre-mRNA *trans*-splicing molecule sequence or a nucleic acid sequence encoding said pre-mRNA *trans*-splicing molecule, wherein the second pre-mRNA *trans*-splicing molecule comprises from 5' to 3'; (i) a second binding domain, which is complementary to the first target domain of the first nucleic acid sequence; (ii) an acceptor splice region sequence comprising; (iia) a pyrimidine tract, wherein the pyrimidine tract comprises; (iiaa) 5 to 25 nucleotides; (iiab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (iib) an acceptor splice site, (iiba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and (iibb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; (iii) a 3' portion of the nucleotide sequence of interest, and (iv) a termination sequence, preferably a polyA sequence; step (C) cleaving the first nucleic acid sequence at the donor splice site sequence and cleaving the second nucleic acid sequence in the acceptor splice site; and step (D) ligating the first cleaved nucleic acid sequence comprising the 5' portion of the nucleotide sequence of interest to the second cleaved nucleic acid sequence comprising the 3' portion of the nucleotide sequence of interest, thereby obtaining the nucleic acid sequence of interest.

[18] In yet another aspect, the disclosure relates to an adeno-associated virus (AAV) vector comprising at least two inverted terminal repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence comprises from 5' to 3': (i) a promoter; (ii) a binding domain; (iii) optionally a spacer sequence; (iv) an acceptor splice region sequence comprising (a) a pyrimidine tract, wherein the pyrimidine tract comprises (aa) 5 to 25 nucleotides; (ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (b)a acceptor splice site, wherein the acceptor splice site is located 3' to the pyrimidine tract; and wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; (v) a nucleotide sequence of interest or a portion thereof; and (vi) optionally a polyA sequence. In one aspect, the present disclosure provides an adeno-associated virus (AAV) vector system comprising (I) a first AAV vector comprising at least two inverted terminal repeats comprising a nucleic acid sequence between

these two inverted terminal repeats, wherein said nucleic acid sequence between these two inverted terminal repeats comprises from 5' to 3': (a) a promoter, (b) a nucleotide sequence encoding an N-terminal portion of a polypeptide of interest; (c) a donor splice site; (d) optionally a spacer sequence; (e) a first binding domain; and (f) optionally a termination sequence, preferably a polyA sequence; (II) a second AAV vector comprising a nucleic acid sequence comprising at least two inverted terminal repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence between these two inverted terminal repeats comprises from 5' to 3' (i) a promoter; (ii) a second binding domain, which is complementary to the first binding domain of the first AAV vector; (iii) optionally a spacer sequence (iv) an acceptor splice region sequence comprising (a) a pyrimidine tract, wherein the pyrimidine tract comprises (aa) 5 to 25 nucleotides; (ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (b) an acceptor splice site, (ba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and (bb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; (v) a nucleotide sequence encoding a C-terminal portion of the polypeptide of interest; (va) wherein the C-terminal portion of the polypeptide of interest and the N-terminal portion of the polypeptide of interest reconstitute the polypeptide of interest; and (vi) a termination sequence, preferably a polyA sequence. The AAV vector may also be part of an AAV vector system comprising (I) a first AAV vector comprising at least two inverted terminal

repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence between these two inverted terminal repeats comprises from 5' to 3': (a) a promoter, (b) a nucleotide sequence encoding an N-terminal portion of a polypeptide of interest; (c) a donor splice site; (d) optionally a spacer sequence; (e) a first binding domain; and (f) optionally a termination sequence, preferably a polyA sequence; (II) a second AAV vector comprising a nucleic acid sequence comprising at least two inverted terminal repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence between these two inverted terminal repeats comprises from 5' to 3' (i) a promoter; (ii) a second binding domain, which is complementary to the first binding domain of the first AAV vector; (ii) optionally a spacer sequence (iii) an acceptor splice region sequence comprising (a) a pyrimidine tract, wherein the pyrimidine tract comprises (aa) 5 to 25 nucleotides; (ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (b) an acceptor splice site, wherein the acceptor splice site is located 3' to the pyrimidine tract; and wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; (iv) a nucleotide sequence encoding a C-terminal portion of the polypeptide of interest; (iva) wherein the C-terminal portion of the polypeptide of interest and the N-terminal portion of the polypeptide of interest reconstitute the polypeptide of interest; and (v) a termination sequence, preferably a polyA sequence. In one example the polypeptide is a full-length polypeptide and the first AAV vector comprises a N-terminal portion of a full-length polypeptide of interest and the second AAV vector comprises a C-terminal portion of the full-length polypeptide of interest.

[19] Also provided is a nucleic acid sequence comprising an acceptor splice region sequence comprising (ia) a pyrimidine tract, wherein the pyrimidine tract comprises (aa) 5 to 25 nucleotides; (ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (ib) a acceptor splice site, ba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and bb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G; and (ii) a nucleotide sequence of interest, wherein the nucleotide sequence of interest (iia) is located 3' or 5' to the acceptor splice region. In one aspect, the present disclosure provides a nucleic acid sequence comprising (i) an acceptor splice region sequence comprising (ia) a pyrimidine tract, wherein the pyrimidine tract comprises (aa) 5 to 25 nucleotides; (ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U); (ib) a acceptor splice site, ba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and bb) wherein the acceptor splice site comprises from 5' to 3' a sequence of CAGG; and (ii) a nucleotide sequence of interest, wherein the nucleotide sequence of interest (iia) is located 3' or 5' to the

acceptor splice region.

[20] In certain examples according to the disclosure the acceptor splice region described herein comprises a pyrimidine tract, wherein the 5 to 25 nucleotides of the pyrimidine tract comprise a sequence encoded by the sequence of TTTTTT or TCTTTT. The sequence between the last pyrimidine of the pyrimidine tract and the acceptor splice site preferably has less than 10, more preferably less than 5, even more preferably less than 3 bases. Moreover, the acceptor splice site preferably has the sequence CAGG. The acceptor splice region may further comprise 7 nucleotides 5' to the pyrimidine tract having at least 4

nucleotides of the sequence CAACGAG, wherein the first 5' nucleotide is C. In a preferred example the splice acceptor region comprises or consists of the sequence of SEQ ID NO: 3 or 4 or is encoded by a nucleic acid sequence comprising or consisting of the sequence of SEQ ID NO: 3 or 4.

BRIEF DESCRIPTION OF THE FIGURES

[21] **Fig. 1** mRNA splicing of *RHO* minigenes in HEK293 cells and in transduced photoreceptors.

A, Schematic true to scale view of the *RHO* gene. The boxes represent the exons and the thick lines in between represent the introns, the respective start and stop codon is indicated. The asterisk symbolizes the c.620T>G mutation in exon 3. **B** and **C**, Rhodopsin minigenes encompassing the coding parts of the exons and the flanking introns were driven either by a CMV promoter for expression in HEK293 cells (B), or by a human rhodopsin (hRHO) promoter for expression in photoreceptors (C). To enable packaging of the *RHO* minigene into AAV vectors, intron 1 was shortened as indicated. For protein visualization and detection, *RHO* was C-terminally fused either to citrine (in B), or to a myc-tag (in C). The primers shown as arrows in B and C were used for specific detection of the wild type (WT) or mutant splicing products derived from HEK293 cells (D) or from transduced photoreceptors (E). **D**, RT-PCR analysis from HEK293 cells transiently transfected with WT and mutant *RHO* minigenes. **E**, Left, schematic representation of the subretinal *RHO* minigene delivery to murine retinas at postnatal day 14 (P14). Right, RT-PCR from injected murine retinas containing the respective WT or mutant *RHO* minigenes. RT-PCR was performed four weeks post injection. All experiments were repeated once.

[22] **Fig. 2** Identification of the most efficient ASS_620 sequence.

A, Composition of the single elements of a functional ASS. The branch point (BP) is underlined representing the consensus sequence for human branch points, with the branch point nucleoside adenosine highlighted in bold. The polypyrimidine tract (Poly-C/T) is represented as an empty box and the ASS spanning the intron-exon boundary is represented by a grey filled box. **B**, Exon-intron organization of the human *RHO* gene and zoomed view in the DNA sequence of exon 3 of the human *RHO* gene for wild type (WT, upper panel) and the c.620T>G mutant (lower panel). Intronic sequences are shown in lower case, exonic sequences are shown in upper case (bold) letters. The single ASS elements are highlighted according to the scheme shown in A, the BP sequence is underlined (thick line), the Poly-C/T sequence is marked by an empty box, the ASS sequence(s) is/are marked by a filled grey box. The ATG sequence in the WT converted to the AGG sequence in the c.620T>G mutant is marked by a thin underline. Note that the c.620T>G mutation generates a novel canonical ASS sequence. The other two elements required for a

CLAIMS

1. A pre-mRNA *trans*-splicing molecule comprising,
   (i) an acceptor splice region, comprising
      (ia) a pyrimidine tract, wherein the pyrimidine tract comprises
         (iaa) 5 to 25 nucleotides;
         (iab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U);
      (ib) an acceptor splice site,
         (iba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and
         (ibb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G;
   (ii) a nucleotide sequence of interest or a portion thereof; wherein the acceptor splice region is localized 3' or 5' to the nucleotide sequence of interest or the portion thereof;
   (iii) a binding domain targeting pre-mRNA, which is localized 3' or 5' to the nucleic acid sequence of interest or the portion thereof; and
   (iv) optionally a spacer sequence, wherein the spacer sequence is localized between the binding domain and the acceptor splice region.

2. The pre-mRNA *trans*-splicing molecule of claim 1 comprising,
   (i) the nucleotide sequence of interest or the portion thereof, wherein the acceptor splice region is localized 5' to the nucleotide sequence of interest or the portion thereof;
   (ii) a binding domain targeting pre-mRNA located 5' to the nucleotide sequence of interest or the portion thereof; and
   (iii) a spacer sequence, wherein the spacer is localized between the first binding domain and the acceptor splice region.

3. The pre-mRNA *trans*-splicing molecule of claim 1 or 2, wherein the acceptor splice region is localized 5' to the nucleotide sequence of interest or the portion thereof and the binding domain is localized 5' to acceptor splice region.

4. The pre-mRNA *trans*-splicing molecule of any one of claims 1 to 3 further comprising a termination sequence, preferably a polyA sequence.

5. The pre-mRNA trans-splicing molecule of any one of claims 1 to 4, wherein
(a) the 5 to 25 nucleotides of the pyrimidine tract comprise a sequence encoded by the sequence of TTTTTT or TCTTTT;
(b) the sequence between the last pyrimidine of the pyrimidine tract and the acceptor splice site has less than 10, preferably less than 5, more preferably less than 3 bases; and/or
(c) the acceptor splice site has a sequence of CAGG.

6. The pre-mRNA trans-splicing molecule of any one of claims 1 to 5, wherein the acceptor splice region further comprises
(a) 7 nucleotides 5' to the pyrimidine tract having at least 4 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotide is C;
(b) 7 nucleotides 5' to the pyrimidine tract having at least 5 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotide is C;
(c) 7 nucleotides 5' to the pyrimidine tract having at least 4 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotides are CAA; or
(d) 7 nucleotides 5' to the pyrimidine tract having at least 5 nucleotides of the sequence CAACGAG, wherein the first 5' nucleotides are CAAC.

7. A DNA molecule comprising a promoter and a sequence encoding the pre-mRNA *trans*-splicing molecule according to any one of claims 1 to 6.

8. A method for producing a nucleic acid sequence, the method comprising
(A) providing a first nucleic acid sequence comprising one or more donor splice site sequences;
(B) providing a second nucleic acid sequence, wherein the second nucleic acid sequence comprises
(i) an acceptor splice region sequence comprising
(ia) a pyrimidine tract, wherein the pyrimidine tract comprises
(iaa) 5 to 25 nucleotides;
(iab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U);
(ib) an acceptor splice site,
(iba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and
(ibb) wherein the acceptor splice site comprises from 5' to 3'

a sequence of NAGG, wherein N is a A, C, T/U or G;

(C) cleaving the first nucleic acid sequence at the one or more donor splice site sequence(s) and cleaving the second nucleic acid sequence in the acceptor splice site,

(D) ligating the first cleaved nucleic acid sequence to the second cleaved nucleic acid sequence,

thereby obtaining a nucleic acid sequence.

9. The method of claim 8, wherein the first nucleic acid sequence further comprises a nucleotide sequence of interest or a portion thereof, wherein at least a portion of the nucleotide sequence of interest is located 5' to the donor splice site and wherein the second nucleic acid sequence further comprises a nucleotide sequence of interest or a portion thereof, wherein at least a portion of the nucleotide sequence of interest is located 3' to the splice acceptor splice region.

10. The method of claim 8 or 9, wherein the first and the second nucleic acid sequence are introduced into a host cell, preferably wherein the first and the second nucleic acid sequence are a recombinant nucleic acid sequence.

11. The method of any one of claims 8 to 10, comprising

(A) introducing into a host cell a first nucleic acid sequence comprising a pre-mRNA trans-splicing molecule sequence or a nucleic acid sequence encoding said pre-mRNA *trans*-splicing molecule, wherein the first pre-mRNA *trans*-splicing molecule comprises from 5' to 3'

(a) a 5' portion of a nucleotide acid sequence of interest;

(b) a donor splice site;

(c) optionally a spacer sequence;

(d) a first binding domain; and

(e) optionally a termination sequence, preferably a polyA sequence, and

(B) introducing into the host cell a second nucleic acid sequence comprising a pre-mRNA *trans*-splicing molecule sequence or a nucleic acid sequence encoding said pre-mRNA *trans*-splicing molecule, wherein the second pre-mRNA *trans*-splicing molecule comprises from 5' to 3';

(i) a second binding domain, which is complementary to the first target domain of the first nucleic acid sequence;

(ii) an acceptor splice region sequence comprising;

(iia) a pyrimidine tract, wherein the pyrimidine tract comprises

(iiaa) 5 to 25 nucleotides;

(iiab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U);

(iib) a acceptor splice site,

(iiba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and

(iibb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G;

(iii) a 3' portion of the nucleotide sequence of interest, and

(iv) a termination sequence, preferably a polyA sequence,

(C) cleaving the first nucleic acid sequence at the donor splice site sequence and cleaving the second nucleic acid sequence in the acceptor splice site; and

(D) ligating the first cleaved nucleic acid sequence comprising the 5' portion of the nucleotide sequence of interest to the second cleaved nucleic acid sequence comprising the 3' portion of the nucleotide sequence of interest, thereby obtaining the nucleic acid sequence of interest.

12. An adeno-associated virus (AAV) vector comprising at least two inverted terminal repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence comprises from 5' to 3 '

(i) a promoter;

(ii) a binding domain;

(iii) optionally a spacer sequence;

(iv) an acceptor splice region sequence comprising

(a) a pyrimidine tract, wherein the pyrimidine tract comprises

(aa) 5 to 25 nucleotides;

(ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U);

(b) a acceptor splice site,

ba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and

bb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G;

(v) a nucleotide sequence of interest or a portion thereof; and

(vi) optionally a poly A sequence.

13. An adeno-associated virus (AAV) vector system comprising

(I) a first AAV vector comprising at least two inverted terminal repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence between these two inverted terminal repeats comprises from 5' to 3'

(a) a promoter,

(b) a nucleotide sequence encoding an N-terminal portion of a polypeptide of interest;

(c) a donor splice site;

(d) optionally a spacer sequence;

(e) a first binding domain; and

(f) optionally a termination sequence, preferably a polyA sequence;

(II) a second AAV vector comprising a nucleic acid sequence comprising at least two inverted terminal repeats comprising a nucleic acid sequence between these two inverted terminal repeats, wherein said nucleic acid sequence between these two inverted terminal repeats comprises from 5' to 3'

(i) a promoter;

(ii) a second binding domain, which is complementary to the first binding domain of the first AAV vector;

(iii) optionally a spacer sequence

(iv) an acceptor splice region sequence comprising

(a) a pyrimidine tract, wherein the pyrimidine tract comprises

(aa) 5 to 25 nucleotides;

(ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U);

(b) an acceptor acceptor splice site,

(ba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and

(bb) wherein the acceptor splice site comprises from 5' to 3' a sequence of NAGG, wherein N is a A, C, T/U or G;

(v) a nucleotide sequence encoding a C-terminal portion of the polypeptide of interest;

(va) wherein the C-terminal portion of the polypeptide of interest and the N-terminal portion of the polypeptide of interest reconstitute the polypeptide of interest; and

(vi) a termination sequence, preferably a polyA sequence.

14. The AAV vector system of claim 13, wherein the polypeptide is a full-length polypeptide

and the first AAV vector comprises an N-terminal portion of a full-length polypeptide of interest and the second AAV vector comprises a C-terminal portion of the full-length polypeptide of interest.

15. A nucleic acid sequence comprising
    (i) an acceptor splice region sequence comprising
        (ia) a pyrimidine tract, wherein the pyrimidine tract comprises
            (aa) 5 to 25 nucleotides;
            (ab) wherein at least 60 % of the nucleotides within these 5 to 25 nucleotides are pyrimidine bases such as cytosine (C), thymine (T) and/or uracil (U);
        (ib) a acceptor splice site,
            ba) wherein the acceptor splice site is located 3' to the pyrimidine tract; and
            bb) wherein the acceptor splice site comprises from 5' to 3' a sequence of CAGG; and
    (ii) a nucleotide sequence of interest, wherein the nucleotide sequence of interest
        (iia) is located 3' or 5' to the acceptor splice region.

16. The method of any one of claims 8 to 11 the AAV vector of claim 12, the AAV vector system of claim 13 or 14 and the nucleic acid of claim 15, wherein
    (a) the 5 to 25 nucleotides of the pyrimidine tract comprise the sequence TTTTTT or TCTTTT;
    (b) the acceptor splice site has a sequence of CAGG; and/or
    (c) wherein the splice acceptor region comprises the sequence of SEQ ID NO: 3 or 4.

17. The pre-mRNA trans-splicing molecule of any one of claims 1 to 6, the adeno-associated virus vector of claim 12 or the adeno-associated virus vector system of claim 13 or 14 for use in therapy.