(12) STANDARD PATENT (11) Application No. **AU 2019288291 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Recombinant adeno-associated virus products and methods for treating dystroglycanopathies and laminin-deficient muscular dystrophies

(51) International Patent Classification(s)
C07K 14/78 (2006.01) ***A61P 21/00*** (2006.01)
A61K 48/00 (2006.01) ***C07K 14/47*** (2006.01)

(21) Application No: **2019288291** (22) Date of Filing: **2019.06.18**

(87) WIPO No: **WO19/246125**

(30) Priority Data

(31) Number	(32) Date	(33) Country
62/686,522	**2018.06.18**	**US**

(43) Publication Date: **2019.12.26**
(44) Accepted Journal Date: **2024.09.26**

(71) Applicant(s)
Research Institute at Nationwide Children's Hospital

(72) Inventor(s)
MARTIN, Paul Taylor

(74) Agent / Attorney
WRAYS PTY LTD, L7 863 Hay St, Perth, WA, 6000, AU

(56) Related Art
REINHARD, J.R., et al, 2017, 'Linker proteins restore basement membrane and correct LAMA2 -related muscular dystrophy in mice', Science Translational Medicine, 9(396), pages 1-27
MCKEE, K.M., et al., 2017, 'Chimeric protein repair of laminin polymerization ameliorates muscular dystrophy phenotype"' Journal of Clinical Investigation, 127(3), pages 1075-1089

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 December 2019 (26.12.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/246125 A8

(51) International Patent Classification:
C07K 14/78 (2006.01) *A61K 48/00* (2006.01)
C07K 14/47 (2006.01) *A61P 21/00* (2006.01)

(21) International Application Number:
PCT/US2019/037769

(22) International Filing Date:
18 June 2019 (18.06.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/686,522 18 June 2018 (18.06.2018) US

(71) Applicant: **RESEARCH INSTITUTE AT NATIONWIDE CHILDREN'S HOSPITAL** [US/US]; 700 Children's Drive, Columbus, OH 43205 (US).

(72) Inventor: **MARTIN, Paul, Taylor**; 2520 Maryland Ave, Bexley, OH 43209 (US).

(74) Agent: **SINTICH, Sharon, M.**; Marshall, Gerstein & Borun LLP, 233 S. Wacker Drive, 6300 Wills Tower, Chicago, IL 60606-6357 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*
— *with information concerning refusal to authorize the rectification of an obvious mistake under Rule 91.1 (Rule 91.3(d))*

(48) Date of publication of this corrected version:
14 May 2020 (14.05.2020)

(15) Information about Correction:
see Notice of 14 May 2020 (14.05.2020)

WO 2019/246125 A8

(54) Title: RECOMBINANT ADENO-ASSOCIATED VIRUS PRODUCTS AND METHODS FOR TREATING DYSTROGLYCANOPATHIES AND LAMININ-DEFICIENT MUSCULAR DYSTROPHIES

(57) Abstract: Products and methods for treating dystroglycanopathies and laminin-deficient muscular dystrophies are provided. In the methods, a protein including a linker domain, such as the heparin-binding domain of Heparin-Binding Epidermal Growth Factor-Like Growth Factor (HBEGF), is delivered to patients.

RECOMBINANT ADENO-ASSOCIATED VIRUS PRODUCTS AND METHODS FOR TREATING DYSTROGLYCANOPATHIES AND LAMININ-DEFICIENT MUSCULAR DYSTROPHIES

[0001] This application claims priority benefit to U.S. Provisional Patent Application No. 62/686,522, filed June 18, 2018, which is incorporated by reference herein in its entirety.

Incorporation By Reference of Material Submitted Electronically

[0002] This application contains, as a separate part of the disclosure, a Sequence Listing in computer-readable form which is incorporated by reference in its entirety and identified as follows: Filename: 53147A_Seqlisting.txt; Size: 125,439 bytes, created; June 18, 2019.

Field of the Invention

[0003] Products and methods for treating dystroglycanopathies and laminin-deficient muscular dystrophies are provided. In the methods, a protein including a linker domain, such as the heparin-binding domain of Heparin-Binding Epidermal Growth Factor-Like Growth Factor (HBEGF), is delivered to patients. This linker protein assists in targeting a transgene to the extracellular matrix (ECM) of a muscle cell.

Background

[0004] Muscular dystrophies (MDs) are a group of genetic diseases. The group is characterized by progressive weakness and degeneration of the skeletal muscles that control movement. Some forms of MD develop in infancy or childhood, while others may not appear until middle age or later. The disorders differ in terms of the distribution and extent of muscle weakness (some forms of MD also affect cardiac muscle), the age of onset, the rate of progression, and the pattern of inheritance.

[0005] Congenital muscular dystrophy (CMD) describes a group of MDs in which the loss of muscle structural components results in neonatal hypotonia and progressive skeletal muscle weakness. These disorders are often associated with significant extramuscular complications, including brain and eye developmental defects, cognitive impairment, seizures, and respiratory and cardiac abnormalities, requiring regular medical management by a multidisciplinary team. The estimated incidence of CMD is 1 in 21,500 live births worldwide. Despite the gravity of these disorders, there are currently no approved and effective therapies. Dystroglycanopathies and merosin-deficient CMD Type 1A (MDC1A)

are two of the most common forms of CMD [Sframeli *et al., Neuromuscul Disord., 27*(9): 793-803 (2017)].

[0006] Dystroglycanopathies are caused by mutations in any of eighteen or more genes required for glycosylating α-dystroglycan. Proper glycosylation allows α-dystroglycan to bind components of the extracellular matrix (ECM). α-Dystroglycan, in turn, anchors to the sarcolemma by binding β-dystroglycan, a transmembrane protein. The number of susceptible genes makes impossible the development of a single gene-replacement therapy for dystroglycanopathy. Examples of dystroglycanopathies include the following. Walker-Warburg Syndrome (WWS) involves genetic mutations in *B3GLNT2, B4GAT1, DAG1, FKRP, FKTN, GMPPB, ISPD,* or *LARGE.* Muscle Eye Brain disease (MEB) involves genetic mutations in *B3GLNT2, B4GAT1, DAG1, FKRP, FKTN, GMPPB, ISPD,* or *LARGE.* Fukuyama CMD involves mutations in the *FKTN* gene. A group of congenital muscular dystrophies with cognitive impairment results from mutations in *FKRP, LARGE, POMT1, POMT2,* or *POMGNT1.* A group of CMDs without cognitive impairment are a result of genetic mutations in *FKRP* or *FKTN.* Limb Girdle Muscular Dystrophies LGMD 2I, 2K, 2M, 2N and 2O are associated with glycosylation abnormalities involving genetic mutations in *FKRP, FKTN, POMGNT1, POMT1,* or *POMT2.* Limb Girdle Muscular Dystrophies LGMD 2T and 2U are respectively a result of genetic mutations in *GMPPB* and *ISPD.* Other mutated genes in dystroglycanopathies include *DOLK, DPM1, DPM2, DPM3, GTDC2/AG061, TMEM5,* and *SK196.*

[0007] MDC1A is caused by mutations in the *LAMA2* gene, encoding the key ECM protein, laminin-α2, which binds glycosylated α-dystroglycan at the sarcolemma. The full *LAMA2* gene is over 9,000 base pairs in length.

[0008] A study by Reinhard and colleagues [Reinhard et al., Sci Transl Med., 9(396), (2017)] involved germline expression of fused domains from laminin-α4 and mini-agrin to incompletely ameliorate disease symptoms in the dyW/dyW mouse model of MDC1A.

[0009] Adeno-associated virus (AAV) is a replication-deficient parvovirus, the single-stranded DNA genome of which is about 4.7 kb in length including two 145 nucleotide inverted terminal repeat (ITRs). There are multiple serotypes of AAV. The nucleotide sequences of the genomes of AAV serotypes are known. For example, the complete genome of AAV-1 is provided in GenBank Accession No. NC_002077; the complete genome of AAV-2 is provided in GenBank Accession No. NC_001401 and Srivastava et al., J. Virol.,

45: 555-564 (1983); the complete genome of AAV-3 is provided in GenBank Accession No. NC_1829; the complete genome of AAV-4 is provided in GenBank Accession No. NC_001829; the AAV-5 genome is provided in GenBank Accession No. AF085716; the complete genome of AAV-6 is provided in GenBank Accession No. NC_00 1862; at least portions of AAV-7 and AAV-8 genomes are provided in GenBank Accession Nos. AX753246 and AX753249, respectively; the AAV-9 genome is provided in Gao et al., J. Virol., 78: 6381-6388 (2004); the AAV-10 genome is provided in *Mol. Ther., 13*(1): 67-76 (2006); the AAV-11 genome is provided in *Virology, 330*(2): 375-383 (2004); portions of the AAV-12 genome are provided in Genbank Accession No. DQ813647; portions of the AAV-13 genome are provided in Genbank Accession No. EU285562. The sequence of the AAV rh.74 genome is provided in see U.S. Patent 9,434,928, incorporated herein by reference. Cis-acting sequences directing viral DNA replication (rep), encapsidation/packaging and host cell chromosome integration are contained within the AAV ITRs. Three AAV promoters (named p5, p19, and p40 for their relative map locations) drive the expression of the two AAV internal open reading frames encoding rep and cap genes. The two rep promoters (p5 and p19), coupled with the differential splicing of the single AAV intron (at nucleotides 2107 and 2227), result in the production of four rep proteins (rep 78, rep 68, rep 52, and rep 40) from the rep gene. Rep proteins possess multiple enzymatic properties that are ultimately responsible for replicating the viral genome. The cap gene is expressed from the p40 promoter and it encodes the three capsid proteins VP1, VP2, and VP3. Alternative splicing and non-consensus translational start sites are responsible for the production of the three related capsid proteins. A single consensus polyadenylation site is located at map position 95 of the AAV genome. The life cycle and genetics of AAV are reviewed in Muzyczka, *Current Topics in Microbiology and Immunology, 158*: 97-129 (1992).

[0010] AAV possesses unique features that make it attractive as a vector for delivering foreign DNA to cells, for example, in gene therapy. AAV infection of cells in culture is noncytopathic, and natural infection of humans and other animals is silent and asymptomatic. Moreover, AAV infects many mammalian cells allowing the possibility of targeting many different tissues in vivo. Moreover, AAV transduces slowly dividing and non-dividing cells, and can persist essentially for the lifetime of those cells as a transcriptionally active nuclear episome (extrachromosomal element). The AAV proviral genome is inserted as cloned DNA in plasmids, which makes construction of recombinant genomes feasible. Furthermore, because the signals directing AAV replication and genome encapsidation are contained

within the ITRs of the AAV genome, some or all of the internal approximately 4.3 kb of the genome (encoding replication and structural capsid proteins, rep-cap) may be replaced with foreign DNA. To generate AAV vectors, the rep and cap proteins may be provided in trans. Another significant feature of AAV is that it is an extremely stable and hearty virus. It easily withstands the conditions used to inactivate adenovirus (56° to 65°C for several hours), making cold preservation of AAV less critical. AAV may even be lyophilized. Finally, AAV-infected cells are not resistant to superinfection.

[0011] There remains a need in the art for treatments for CMDs such as dystroglycanopathies and MDC1A.

[0011a] Any discussion of the prior art throughout the specification should in no way be considered as an admission that such prior art is widely known or forms part of common general knowledge in the field.

[0011b] Unless the context clearly requires otherwise, throughout the description and the claims, the words “comprise”, “comprising”, and the like are to be construed in an inclusive sense as opposed to an exclusive or exhaustive sense; that is to say, in the sense of “including, but not limited to”.

Summary

[0012] Provided herein are methods and products for treatment of CMDs such as dystroglycanopathies and laminin-deficient muscular dystrophies. The products include therapeutic proteins and rAAV encoding a disclosed therapeutic protein.

[0013] A polynucleotide is provided encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of Heparin-Binding Epidermal Growth Factor-Like Growth Factor (HBEGF), and a second domain comprising the G1-G5 domain of the human laminin alpha 2 (LAMA2) gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene;

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene,

e) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG 1 alpha or

f) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAGlalpha.

[0013a] According to a first aspect, the present invention provides a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of Heparin-Binding Epidermal Growth Factor-Like Growth Factor (HBEGF) encoded by the nucleotide sequence of SEQ ID NO: 13, and a second domain comprising the G1-G5 domain of the human laminin alpha 2 (LAMA2) encoded by the nucleotide sequence of SEQ ID NO: 15;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 14, and a second domain comprising the G1-G5 domain of the human LAMA2 encoded by the nucleotide sequence of SEQ ID NO: 15;

c) a first domain comprising the heparin-binding domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 13, and a second domain comprising the G3-G5 domain of the human LAMA2 gene encoded by the nucleotide sequence of SEQ ID NO: 16;

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 14, and a second domain comprising the G3-G5 domain of the human LAMA2 encoded by the nucleotide sequence of SEQ ID NO: 16;

e) a first domain comprising the heparin-binding domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 13, and a second domain comprising the processed native

alpha chain of the human dystroglycan gene (DAG1alpha) encoded by the nucleotide sequence of SEQ ID NO: 17; or

f) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 14, and a second domain comprising DAG1alpha encoded by the nucleotide sequence of SEQ ID NO: 17.

[0013b] According to a second aspect, the present invention provides a recombinant adeno-associate virus (rAAV), wherein the genome of the rAAV comprises the polynucleotide of the invention.

[0013c] According to a third aspect, the present invention provides a recombinant adeno-associate virus (rAAV), wherein the genome of the rAAV comprises

a) a polynucleotide sequence of SEQ ID NO: 1,

b) a polynucleotide sequence of SEQ ID NO: 3,

c) a polynucleotide sequence of SEQ ID NO: 5,

d) a polynucleotide sequence of SEQ ID NO: 7,

e) a polynucleotide sequence of SEQ ID NO: 9,

f) a polynucleotide sequence of SEQ ID NO: 11,

g) nucleotides 3590 to 8215 of SEQ ID NO: 2,

h) nucleotides 3590 to 8341 of SEQ ID NO: 4,

i) nucleotides 3609 to 6929 of SEQ ID NO: 6,

j) nucleotides 3590 to7036 of SEQ ID NO: 8,

k) nucleotides 3590 to 6340 of SEQ ID NO: 10,

l) nucleotides 3590 to 6049 of SEQ ID NO: 12,

m) the nucleotide sequence set out in Figure 13, or

n) the nucleotide sequence set out in Figure 14.

[0013d] According to a fourth aspect, the present invention provides an rAAV particle comprising an rAAV of the invention.

[0013e] According to a fifth aspect, the present invention provides a recombinant host cell comprising the polynucleotide of the invention optionally wherein the host cell is a Chinese hamster ovary (CHO) cell or HEK293 cell.

[0013f] According to a sixth aspect, the present invention provides a protein encoded by the polynucleotide of any one of claims 1 to 7.

[0013g] According to a seventh aspect, the present invention provides a composition comprising a polynucleotide of the invention, an rAAV of the invention, an rAAV particle of the invention or a protein of the invention.

[0013h] According to an eighth aspect, the present invention provides a method for treating a laminin-deficient muscular dystrophy comprising administering to a patient in need thereof a polynucleotide of the invention, an rAAV of the invention, an rAAV particle of the invention, a protein of the invention or a composition of the invention.

[0013i] According to a ninth aspect, the present invention provides use of a polynucleotide of any one of the invention, an rAAV of the invention, an rAAV particle of the invention, a protein of the invention or a composition of the invention in the preparation of a medicament for the treatment of a laminin-deficient muscular dystrophy.

[0014] In one embodiment, the provided polynucleotides encode a protein, wherein the first domain of the protein is encoded by the nucleotide sequence of SEQ ID NO: 13 or SEQ ID NO: 14, and the second the second domain of the protein is encoded by the nucleotide sequence of SEQ ID NO: 15, SEQ ID NO: 16 or SEQ ID NO: 17.

[0015] For example, the provided polynucleotides comprises the nucleotide sequence of SEQ ID NO: 13 and SEQ ID NO: 15, or comprise the nucleotide sequence of SEQ ID NO: 13 and SEQ ID NO: 16, or comprise the nucleotide sequence of SEQ ID NO: 14 and SEQ ID NO: 15, or comprise the nucleotide sequence of SEQ ID NO: 14 and SEQ ID NO: 16, or comprise the nucleotide sequence of SEQ ID NO: 13 and SEQ ID NO: 17 or comprise the nucleotide sequence of SEQ ID NO: 14 and SEQ ID NO: 17.

[0016] In one embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 3, ii) a nucleotide sequence comprising nucleotides14 to 3235 set out in Figure 3, iii) the nucleotide sequence of SEQ ID NO: 1, or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 19.

[0017] In another embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 4, ii) a nucleotide sequence comprising nucleotides 14 to 3361 set forth in Figure 4, iii) the nucleotide sequence of SEQ ID NO: 3 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 20.

[0018] In a further embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 5, ii) a nucleotide sequence comprising nucleotides 14 to 1930 set forth in Figure 5, iii) the nucleotide sequence of SEQ ID NO: 5 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 21.

[0019] In another embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 6, ii) a nucleotide sequence comprising nucleotides 14 to 2056 set forth in Figure 6, iii) the nucleotide sequence of SEQ ID NO: 7 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 22.

[0020] In an embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 7, ii) a nucleotide sequence comprising nucleotides 14 to 1360 set forth in Figure 7, iii) the nucleotide sequence of SEQ ID NO: 9 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 23.

[0021] In another embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 8, ii) a nucleotide sequence

comprising nucleotides 14-1486 set forth in Figure 8, iii) the nucleotide sequence of SEQ ID NO: 11 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 24.

[0022] Therapeutic proteins encoded by any of the provided polynucleotides are also provided. For example, the provided proteins comprise the amino acid sequence of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23 or SEQ ID NO: 24.

[0023] In addition, the disclosure provides recombinant host cells comprising any of the polynucleotide described herein. In exemplary embodiments, the host cells, the polynucleotides are operatively linked to a transcriptional control element and these host cells express any of the polynucleotides disclosed herein. For example, the host cells are Chinese hamster ovary (CHO) cell or human HEK293 cell.

[0024] Further provided are recombinant adeno-associate virus (rAAV), wherein the genome of the rAAV comprises any of the polynucleotide described herein. For example, provided are rAAV, wherein the genome of the rAAV comprises a polynucleotide sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9 or SEQ ID NO: 11. In exemplary embodiments, the rAAV genome further comprises a muscle-specific transcriptional control element, such as a CMV promoter (SEQ ID NO: 18), MCK, NHCK, LAMA2 or tMCK. Any of the rAAV described herein comprise the AAV9, AAV10, AAVrh74, AAV8 or AAV6 capsid.

[0025] Also provided are rAAV, wherein the genome of the rAAV comprises nucleotides 3590 to 8215 of SEQ ID NO: 2, nucleotides 3590 to 8341 of SEQ ID NO: 4, nucleotides 3609 to 6929 of SEQ ID NO: 6, nucleotides 3590 to7036 of SEQ ID NO: 8, nucleotides 3590 to 6340 of SEQ ID NO: 10, nucleotides 3590 to 6049 of SEQ ID NO: 12, the nucleotide sequence set out in Figure 13, or the nucleotide sequence set out in Figure 14.

[0026] rAAV particles comprising any of the rAAV described herein are also provided. The disclosure also provides for compositions comprising any of the polynucleotides disclosed herein, any of the rAAV disclosed herein, any of the rAAV particles disclosed herein or any of the proteins disclosed herein.

[0027] Provided are methods for treating a laminin-deficient muscular dystrophy comprising administering to a patient in need thereof a rAAV, wherein the genome of the rAAV comprises a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene; or

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene.

[0028] For example, the methods of treating a laminin-deficient muscular dystrophy comprise administering to a patient in need thereof any of the following: any of the polynucleotides disclosed herein which encode a protein comprising LAMA2(G1-G5) or LAMA2(G3-G5) as the second domain, any of the rAAV or rAAV particles disclosed herein which comprise a polynucleotide encoding a protein comprising LAMA2(G1-G5) or LAMA2(G3-G5) as the second domain.

[0029] Further provided are methods for treating a laminin-deficient muscular dystrophy comprising administering to a patient in need thereof a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene; or

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene.

[0030] For example, the methods for treating a treating laminin-deficient muscular dystrophy comprise administering a protein to a patient in need thereof, wherein the protein

comprises the amino acid sequence of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21 or SEQ ID NO: 22.

[0031] Also provided are compositions for treating a laminin-deficient muscular dystrophy comprising a rAAV, wherein the genome of the rAAV comprises a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene; or

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene.

[0032] For example, the compositions for treating a laminin-deficient muscular dystrophy comprise any of the following: any of the polynucleotides disclosed herein which encode a protein comprising LAMA2(G1-G5) or LAMA2(G3-G5) as the second domain, any of the rAAV or rAAV particles disclosed herein which comprise a polynucleotide encoding a protein comprising LAMA2(G1-G5) or LAMA2(G3-G5) as the second domain.

[0033] Further provided are compositions for treating laminin-deficient muscular dystrophies comprising a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene; or

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene.

[0034] For example, the compositions for treating a treating laminin-deficient muscular dystrophy comprise a protein, wherein the protein comprises the amino acid sequence of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21 or SEQ ID NO: 22.

[0035] The disclosure also provides for a use of a rAAV for the preparation of a medicament for treating a laminin-deficient muscular dystrophy in a patient in need thereof, wherein the genome of the rAAV comprises a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene; or

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene.

[0036] For example, the disclosure also provides for use of any of the following: any of the polynucleotides disclosed herein which encode a protein comprising LAMA2(G1-G5) or LAMA2(G3-G5) as the second domain, any of the rAAV or rAAV particles disclosed herein which comprise a polynucleotide encoding a protein comprising LAMA2(G1-G5) or LAMA2(G3-G5) as the second domain for the preparation of a medicament for treating a laminin-deficient muscular dystrophy in a patient in need thereof.

[0037] Further provided are use of a protein for the preparation of a medicament for treating a laminin-deficient muscular dystrophy in a patient in need thereof, wherein the protein comprises:

a) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G1-G5 domain of the human LAMA2 gene;

c) a first domain comprising the heparin-binding domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene; or

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising the G3-G5 domain of the human LAMA2 gene.

[0038] For example, the disclosure also provides for use of a protein for the preparation of a medicament for treating a treating laminin-deficient muscular dystrophy in a patient in need thereof, wherein the protein comprises the amino acid sequence of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21 or SEQ ID NO: 22.

[0039] Also provided are methods for treating a dystroglycanopathy comprising administering to a patient in need thereof a rAAV, wherein the genome of the rAAV comprises a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG1alpha or

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAG1alpha.

[0040] For example, the methods of treating a dystroglycanopathy comprising administering to a patient in need thereof any of the following: any of the polynucleotides disclosed herein which encode a protein comprising DAG1alpha as the second domain, any of the rAAV or rAAV particles disclosed herein which comprise a polynucleotide encoding a protein comprising DAG1alpha as the second domain.

[0041] Still further provided are methods for treating a dystroglycanopathy comprising administering to a patient in need thereof a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG1alpha or

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAG1alpha.

[0042] For example, the methods for treating a dystroglycanopathy comprising administering a protein to a patient in need thereof, wherein the protein comprises the amino acid sequence of SEQ ID NO: 23 or SEQ ID NO: 24.

[0043] Also provided are compositions for treating a dystroglycanopathy comprising a rAAV, wherein the genome of the rAAV comprises a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG1alpha or

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAG1alpha.

[0044] For example, the disclosure provides compositions for treating a dystroglycanopathy comprising any of the following: any of the polynucleotides disclosed herein which encode a protein comprising DAG1alpha as the second domain, or any of the rAAV or rAAV particles disclosed herein which comprise a polynucleotide encoding a protein comprising DAG1alpha as the second domain.

[0045] Still further provided are compositions for treating a dystroglycanopathy comprising a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG1alpha or

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAG1alpha.

[0046] For example, the compositions for treating a dystroglycanopathy comprise a protein wherein the protein comprises the amino acid sequence of SEQ ID NO: 23 or SEQ ID NO: 24.

[0047] Also provided is use of a rAAV for the preparation of a medicament for treating a dystroglycanopathy in a patient in need thereof, wherein the genome of the rAAV comprises a polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG1alpha or

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAG1alpha.

[0048] For example, the disclosure provides for use of any of the following: any of the polynucleotides disclosed herein which encode a protein comprising DAG1alpha as the second domain or any of the rAAV or rAAV particles disclosed herein which comprise a polynucleotide encoding a protein comprising DAG1alpha as the second domain, for the preparation of a medicament for treating a dystroglycanopathy in a patient in need thereof.

[0049] Still further provided is a use of a protein for the preparation of a medicament for treating a dystroglycanopathy in a patient in need thereof, wherein the protein comprises:

a) a first domain comprising the heparin-binding domain of HBEGF and a second domain comprising DAG1alpha or

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF, and a second domain comprising DAG1alpha.

[0050] For example, the disclosure provides a use of a protein for the preparation of a medicament for treating a dystroglycanopathy in a patient in need thereof, wherein the protein comprises the amino acid sequence of SEQ ID NO: 23 or SEQ ID NO: 24.

[0051] In methods, uses or composition for treating laminin-deficient muscular dystrophy provided, the laminin-deficient muscular dystrophy may be, for example, MDC1A.

[0052] In any of the methods, uses or compositions for treating a dystroglycanopathy, the dystroglycanopathy may be, for example, Walker Warburg syndrome, Muscle Eye Brain disease, Fukuyama Congenital Muscular Dystrophy, MDC1C, MDC1D, LGMD2I, LGMD2K, LGMD2M, LGMD2N, LGMD2O, LGMD2P, LGMD2T or LGMD2U.

[0053] Examples of the provided proteins are described in Table 1.

Table 1:

Therapeutic protein	Linker/Laminin	Figure	Encoded by nucleotide in Figure or SEQ ID NO:	Protein SEQ ID NO:	Plasmid SEQ ID No:/ Figure
HB-LAMA2(G1-G5)	Ending after HB domain	3	14 to 3235 of Figure 3 SEQ ID NO: 1	19	SEQ ID NO: 2 Figure 19
HBEGF-LAMA2(G1-G5)	Complete soluble form HBEGF	4	14 to 3361 of Figure 4 SEQ ID NO: 3	20	SEQ ID NO: 4 Figure 20
HB-LAMA2(G3-G5)	Ending after HB domain	5	14 to 1930 of Figure 5 SEQ ID NO: 5	21	SEQ ID NO: 6 Figure 21

HBEGF-LAMA2(G3-G5)	Complete soluble form HBEGF	6	14 to 2056 of Figure 6 SEQ ID NO: 7	22	SEQ ID NO: 8 Figure 22
HB-DAG1	Ending after HB domain	7	14 to 1360 of Figure 7 SEQ ID NO: 9	23	SEQ ID NO: 10 Figure 23
HBEGF-DAG1	Complete soluble form HBEGF	8	14-1486 of Figure 8 SEQ ID NO: 11	24	SEQ ID NO: 12 Figure 24

Brief Description of the Drawings

[0054] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings and color photographs will be provided by the Office upon request and payment of the necessary fee.

[0055] Figure 1A depicts the dystrophin-associated glycoprotein (DAG) complex. Figure 1B shows α dystroglycan is not only abnormally glycosylated in dystroglycanopathies, which removes its normal laminin binding function, but α dystroglycan protein is reduced in diseased muscles. Figure 1C shows therapeutic proteins described herein will allow α dystroglycan to link to the muscle membrane by binding to β dystroglycan, which is present in normal amounts, and link to the ECM, even without its proper ECM-binding glycans, via binding of HBEGF to heparin sulfate proteoglycans of the ECM. This will reconstitute the lost linkages of α dystroglycan to the ECM and to the muscle membrane. Use of methods described herein providing these therapeutic proteins is indicated for treatment of all 18-plus genetic forms of dystroglycanopathies, making the methods powerful alternatives to gene replacement strategies in which each dystroglycanopathy would require development of a different gene therapy.

[0056] Figure 2A depicts MDC1A is caused by loss of function mutations in the *LAMA2* gene, which encodes laminin α2, an extracellular matrix (ECM) protein that surrounds each muscle cell in the body. *LAMA2* is required for muscle cell adherence to the ECM and for muscle membrane stability. Figure 2B shows therapeutic proteins described herein can anchor the LAMA2 G1-G5 domains to the ECM where the LAMA2 G1-G5 domains would normally be present, so the LAMA2 G1-G5 domains can function as they do in native laminin α2. Use of methods described herein providing these therapeutic proteins is thus indicated for MDC1A.

[0057] Figure 3 shows the polynucleotide sequence encoding a therapeutic protein HB-EGF (ending at heparin binding domain)-LAMA2 G1-G5. The therapeutic protein is encoded

by nucleotides of the invention comprises nucleotides 14 to 3235, which also correspond to SEQ ID NO: 1.

[0058] Figure 4 shows the polynucleotide sequence encoding a therapeutic protein HB-EGF (complete soluble form)-LAMA2 G1-G5. The therapeutic protein is encoded by nucleotides of the invention comprises nucleotides 14 to 3361, which also correspond to SEQ ID NO: 3.

[0059] Figure 5 shows the polynucleotide sequence encoding a therapeutic protein HB-EGF (ending at heparin binding domain)-LAMA2 G3-G5. The therapeutic protein is encoded by nucleotides of the invention comprises nucleotides 14 to 1930, which also correspond to SEQ ID NO: 5.

[0060] Figure 6 shows the polynucleotide sequence encoding a therapeutic protein HB-EGF (complete soluble form)-LAMA2 G3-G5. The therapeutic protein is encoded by nucleotides of the invention comprises nucleotides 14 to 2056, which also correspond to SEQ ID NO: 7.

[0061] Figure 7 shows the polynucleotide sequence encoding a therapeutic protein HB-EGF (ending at heparin binding domain)-DAG1 (native processed alpha DG gene). The therapeutic protein is encoded by nucleotides of the invention comprises nucleotides 14 to 1360, which also correspond to SEQ ID NO: 9.

[0062] Figure 8 shows the polynucleotide sequence encoding a therapeutic protein HB-EGF (complete soluble form)-DAG1 (native processed alpha DG gene). The therapeutic protein is encoded by nucleotides of the invention comprises nucleotides 14 to 1486, which also correspond to SEQ ID NO: 11.

[0063] Figure 9 shows expression of therapeutic proteins described herein by recombinant mammalian host cells.

[0064] Figure 10 shows that sHB-EGF can be secreted from muscles and stick to the extracellular matrix.

[0065] Figure 11 shows that sHB-EGF induces expression of therapeutic surrogate muscular dystrophy genes. Full length HBEGF does not induce therapeutic gene expression.

[0066] Figure 12 shows that sHB-EGF induces Akt tyrosine kinase cascade in skeletal muscle and can stimulate muscle growth and regeneration.

[0067] Figure 13 shows an exemplary rAAV genome encoding the therapeutic protein HB-EGF (complete soluble form)-LAMA2 G1-G5.

[0068] Figure 14 shows an exemplary rAAV genome encoding the therapeutic protein HB-EGF (complete soluble form)-DAG1 (native processed alpha DG gene).

[0069] Figure 15 provides immunohistochemical staining for HB-EGF and LG5 (denoted in figure as 4H8-2) after intramuscular injection of rAAV9.CMV vectors containing HBEGF, HBEGF.LAMA2(G1-G5), HBEGF.LAMA2(G3-G5), HB.LAMA2(G1-G5), HB.LAMA2(G3-G5), or LAMA2(G1-G5) in wild type mice. Mock injected mice (buffer alone) are shown as a negative control. 4H8-2 is an anti-laminin antibody to show muscle cells in the sections.

[0070] Figure 16 provides immunohistochemistry staining for HB-EGF and Laminin Globular Domain (LG5) in muscles injected IM with rAAV9.HBEGF-LAMA2(G1-G5), HB-LAMA2(G1-G5), or LAMA2(G1-G5). The lower panels below show staining for secondary antibody alone.

[0071] Figure 17 is a graph demonstrating that rAAV9.CMV.HB.LAMA(G1-G5) prevented loss of muscle strength in dy/dy mice. Mice were injected IV with $1x10^{12}$vg of rAAV9.CMV vectors containing HBEGF.LAMA2(G1-G5), HB.LAMA2(G1-G5), or HB.LAMA2(G3-G5). Mice were compared to mock-injected dy/dy disease controls and wild type normal controls at 2 months and 3 months post-injection. Mixed (50:50) female:male genders were used in all groups. Errors are SEM for n=12 (wild type and dy/dy mock), 6 (sHB-EGF.LAMA2G1-G5 and HB.LAMA2G1-G5) or 5 (HB.LAMA2G3-G5) animals per group, with five measures averaged per data point. *$p<0.05$, **$p<0.01$, ***$p<0.001$

[0072] Figure 18 provides immunohistochemistry staining for HB-EGF and LG5 to demonstrate expression of HB.LAMA2(G1-G5) in dy/dy muscle (triceps) at 4 months of age after IV injection at P1. Muscle sections from the triceps muscle of 4-month old wild type and dy/dy mice, either mock-injected or injected with $1x10^{12}$ vg rAAV9.CMV.HB-LAMA2(G1-G5) were stained with antibodies specific to HBEGF (green), to recognize transgenic protein, and to collagen IV (Col(IV), red), to recognize all muscle cells. DAPI is added in blue to stain nuclei. Merged tricolor images are shown.

[0073] Figure 19 provides the plasmid sequence of pAAV.CMV.HB.LAMA1(G1-G5) (SEQ ID NO: 2), the rAAV genome corresponds to nucleotides 3590 to 8215.

[0074] Figure 20 provides the plasmid sequence of pAAV.CMV.HBEGF LAMA2(G1-G5) (SEQ ID NO:4), the rAAV genome corresponds to nucleotides 3590-8341.

[0075] Figure 21 provides the plasmid sequence of pAAV.CMV. HB LAMA2 (G3-G5) (SEQ ID NO: 6), the rAAV genome corresponds to nucleotides 36909-6929.

[0076] Figure 22 provides the plasmid sequence of pAAV.CMV.HBEGF.LAMA2 (G3-G5) (SEQ ID NO: 8), the rAAV genome corresponds to nucleotides 3590-7036.

[0077] Figure 23 provides the plasmid sequence of pAAV.CMV. HB.DAG1 (alpha) (SEQ ID NO: 10), the rAAV genome corresponds to nucleotides, the rAAV genome corresponds to nucleotides 3590 to 6340.

[0078] Figure 24 provides the plasmid sequence of pAAV.CMV.HB.DAG1(alpha) (SEQ ID NO: 13), the rAAV genome corresponds to nucleotides 3590-6049.

Detailed Description

[0079] Methods and products are provided herein for treatment of dystroglycanopathies (including, but not limited to, Walker Warburg syndrome, Muscle Eye Brain disease, Fukuyama Congenital Muscular Dystrophy, MDC1C, MDC1D, LGMD2I, LGMD2K, LGMD2M, LGMD2N, LGMD2O, LGMD2P, LGMD2T and LGMD2U) and laminin-deficient muscular dystrophies (including, but not limited to, MDC1A) which utilize the lysine-rich heparin-binding domain of HBEGF. Heparin sulfate proteoglycans are abundant in the extracellular matrix (ECM) and, as shown herein, the overexpression of HBEGF in muscle leads to localization of HBEGF in the muscle ECM. In methods described herein, the membrane anchoring defects in dystroglycanopathies and laminin-deficient muscular dystrophies are treated using the heparin-binding domain of HBEGF as a "linker" domain in therapeutic proteins.

[0080] Here, the term "HBEGF" refers to the entire HBEGF sequence up to and including the bioactive EGF domain, but lacking the transmembrane domain, which thereby allows HBEGF secretion (Fig. 4). The HBEGF fragment contains four domains from the HBEGF gene, the signal peptide, which allows entry into the secretory pathway, the prepro-peptide, which allows folding and stabilization of the protein, the heparin binding domain, which allows for increased interaction with the extracellular matrix, and the bioactive EGF domain, which allows for HBEGF signaling. In the proteins disclosed herein, the coding sequence for these domains are then then linked to laminin alpha 2 or dystroglycan coding sequences. A second "HB" fragment is also used (Fig. 3). The HB fragment only contains 3 of the four domains found in HBEGF: the signal peptide, the pre-pro-peptide, and the heparin binding domain (and so HB lacks the bioactive EGF domain). When linked to laminin alpha 2 or

dystroglycan protein fragments, the HB domain allows for increased association with the ECM but without increasing EGF or HBEGF signaling.

[0081] The HBEGF or HB linker domain targets a protein to the extracellular matrix of a cell and acts to anchor this protein to the extracellular domain of a cell, such as a muscle cell. Polynucleotides encoding therapeutic proteins are delivered to a patient (for example, delivery by a recombinant AAV encoding the therapeutic proteins), or the therapeutic proteins are delivered to a patient.

[0082] For example, for all of the dystroglycanopathies, a coding sequence for HBEGF heparin-binding domain is fused to a coding sequence for α-dystroglycan, creating a polynucleotide encoding the therapeutic protein HBEGF-DAG1(α). In addition to α-dystroglycan hypoglycosylation, α-dystroglycan protein levels are reduced in dystroglycanopathies. In methods described herein, the HBEGF domain of HBEGF-DAG1(α) binds the ECM heparin sulfate proteoglycans, while the α-dystroglycan domain binds β-dystroglycan, linking the sarcolemma to the ECM despite hypoglycosylation in the dystroglycanopathies. Four examples of such HBEGF-DAG1(α) therapeutic proteins are:
HB-EGF (ending at heparin binding domain)-LAMA2 G1-G5 (encoded by the polynucleotide of Figure 3),
HB-EGF (complete soluble form)-LAMA2 G1-G5 (encoded by the polynucleotide of Figure 4),
HB-EGF (ending at heparin binding domain)-LAMA2 G3-G5 (encoded by the polynucleotide of Figure 5), and
HB-EGF (complete soluble form)-LAMA2 G3-G5 (encoded by the polynucleotide of Figure 6).

[0083] The term "complete soluble form" herein indicates the therapeutic protein comprises the HBEGF heparin-binding and EGF-like domains, but not the transmembrane portion of HBEGF. The combination of the HBEGF heparin-binding and EGF-like domains of HBEGF corresponds to the cleaved, active, soluble isoform of HBEGF. This term is referred to as 'HBEGF" herein.

[0084] The term "ending at heparin binding domain" herein indicates that that therapeutic protein comprises only HBEGF heparin-binding domain and does not comprise the EGF-like domain or the transmembrane portion of HBEGF. This term is abbreviated as 'HB" herein.

[0085] For example, for laminin-deficient muscular dystrophies such as MDC1A, a coding sequence for HBEGF heparin-binding domain is fused to a coding sequence for the globular

(G) domains 1-5 of laminin-α2, creating a polynucleotide encoding the therapeutic protein HBEGF-LAMA2(G1-5). The laminin-α2 G domains bind glycosylated α-dystroglycan and also integrins at the sarcolemma, and are encoded by a part of the LAMA2 gene. In methods described herein, the HBEGF domain of HBEGF-LAMA2(G1-5) binds the ECM heparin sulfate proteoglycans, while the G domains bind α-dystroglycan, linking the sarcolemma to the ECM despite the absence of full-length laminin-α2 in MDC1A. Two examples of such HBEGF-LAMA2(G1-5) therapeutic proteins are:
HB-EGF (ending at heparin binding domain)-DAG1 (native processed alpha DG gene) (encoded by the polynucleotide of Figure 7) and
HB-EGF (complete soluble form)-DAG1 (native processed alpha DG gene) (encoded by the polynucleotide of Figure 8).

[0086] Furthermore, both dystroglycanopathies and laminin-deficient muscular dystrophies (such as MDC1A) are associated with reduced muscle regeneration and, in embodiments of methods described herein wherein the therapeutic proteins comprise a HBEGF heparin-binding domain and HBEGF EGF-like domain, patients also benefit from trophic signaling of the HBEGF EGF-like domain of the therapeutic proteins which results in the alteration of expression of genes including *Pax7*, *MyoD*, *Myogenin* and *Myh3* increasing myogenesis and muscle regeneration.

[0087] Thus, polynucleotides are provided encoding the therapeutic proteins. Embodiments include a polynucleotide comprising the polynucleotide sequence set forth in Figure 3, 4, 5, 6, 7 or 8. Other embodiments include a polynucleotide encoding the same amino acid sequence as the polynucleotide sequence set forth in Figure 3, 4, 5, 6, 7 or 8. Still other embodiments include a polynucleotide comprising a polynucleotide consisting of the polynucleotide sequence set forth in Figure 3, 4, 5, 6, 7 or 8.

[0088] In one embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 3, ii) a nucleotide sequence comprising nucleotides14 to 3235 set out in Figure 3, iii) the nucleotide sequence of SEQ ID NO: 1, or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 19.

[0089] In another embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 4, ii) a nucleotide sequence comprising nucleotides 14 to 3361 set forth in Figure 4, iii) the nucleotide sequence of SEQ ID NO: 3 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 20.

[0090] In a further embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 5, ii) a nucleotide sequence comprising nucleotides 14 to 1930 set forth in Figure 5, iii) the nucleotide sequence of SEQ ID NO: 5 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 21.

[0091] In another embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 6, ii) a nucleotide sequence comprising nucleotides 14 to 2056 set forth in Figure 6, iii) the nucleotide sequence of SEQ ID NO: 7 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 22.

[0092] In an embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 7, ii) a nucleotide sequence comprising nucleotides 14 to 1360 set forth in Figure 7, iii) the nucleotide sequence of SEQ ID NO: 9 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 23.

[0093] In another embodiment, the provided polynucleotides comprise one of the following: i) the nucleotide sequence set forth in Figure 8, ii) a nucleotide sequence comprising nucleotides 14-1486 set forth in Figure 8, iii) the nucleotide sequence of SEQ ID NO: 11 or iv) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 24.

[0094] Other polynucleotides provided include, but are not limited to, a polynucleotide that encodes an amino acid variant of a therapeutic polypeptide which retains the binding activity of the therapeutic protein, which polynucleotide has a nucleotide sequence at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the protein-coding nucleotide sequence set out in Figure 3, 4, 5, 6, 7 or 8 or the nucleotide sequence of any of the provided polynucleotides.

[0095] Also provided herein are polynucleotides that encode an amino acid variant of a therapeutic polypeptide which retains the binding activity of the therapeutic protein, which polynucleotide hybridizes under stringent conditions to the protein-coding nucleotide sequence set out in Figure 3, 4, 5, 6, 7 or 8, or the complement thereof or the nucleotide sequence of any of the provided polynucleotides. The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Hybridization stringency is principally determined by temperature, ionic strength, and the concentration of denaturing agents such as formamide. Examples of stringent conditions for hybridization and washing

are 0.015 M sodium chloride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015M sodium citrate, and 50% formamide at 42°C. See Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, (Cold Spring Harbor, N.Y. 1989).

[0096] "Retains the binding activity" is contemplated herein to mean that the amino acid variant of the therapeutic protein encoded by a polynucleotide competes for binding with a therapeutic protein encoded by the nucleotide sequence set out in Figure 3, 4, 5, or 6 to heparin sulfate proteoglycans and β-dystroglycan; or for binding with a therapeutic protein encoded by the nucleotide sequence set out in Figure 7 or 8 to heparin sulfate proteoglycans and α-dystroglycan; or a therapeutic protein comprising the amino acid sequence of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21SEQ ID NO: 22, SEQ ID NO: 23 OR SEQ ID NO: 24 to heparin sulfate proteoglycans and α-dystroglycan.

[0097] Recombinant expression vectors comprising one or more of the polynucleotides described herein are also provided. Recombinant AAV genomes comprising a polynucleotide described herein are also provided.

[0098] In expression vectors or recombinant AAV genomes described herein, the polynucleotide encoding the therapeutic protein is operatively linked to transcriptional control elements (including, but not limited to, promoters, enhancers and/or introns), specifically transcriptional control elements functional in target cells of interest. For example, suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, retroviruses, hepatitis-B virus, and Simian Virus 40 (SV40). Other suitable mammalian promoters include heterologous mammalian promoters, for example, heat-shock promoters and the actin promoter. Also for example, AAV delivery methods may comprise transducing muscle or liver cells using muscle-specific transcriptional control elements, including, but not limited to, those derived from the actin and myosin gene families, such as from the myoD gene family [See Weintraub *et al., Science, 251*: 761-766 (1991)], the myocyte-specific enhancer binding factor MEF-2 [Cserjesi and Olson, *Mol Cell Biol, 11*: 4854-4862 (1991)], control elements derived from the human skeletal actin gene [Muscat *et al., Mol Cell Biol, 7*: 4089-4099 (1987)], muscle creatine kinase sequence elements [See Johnson *et al., Mol Cell Biol, 9*:3393-3399 (1989)] and the murine creatine kinase enhancer (mCK) element, control elements derived from the skeletal fast-twitch

troponin C gene, the slow-twitch cardiac troponin C gene and the slow-twitch troponin I gene: hypoxia-inducible nuclear factors [Semenza *et al., Proc Natl Acad Sci USA, 88*: 5680-5684 (1991)], steroid-inducible elements and promoters including the glucocorticoid response element (GRE) [Mader and White, *Proc. Natl. Acad. Sci. USA, 90*: 5603-5607 (1993)], the tMCK promoter [see Wang *et al., Gene Therapy, 15*: 1489-1499 (2008)], hybrid α-myosin heavy chain enhancer-/MCK enhancer-promoter (MHCK7) promoter [Salva *et al. Mol Ther,* 15: 320-329 (2007), the CK6 promoter [see Wang *et al., supra*] and other control elements. Thus, one example of a muscle-specific transcriptional control element is the tMCK promoter. An example of a liver-specific promoter is LSP [Wang and Verma, *Proc. Natl. Acad. Sci. USA, 96,* 3906-3910 (1999)]. As another promoter example, for production of the therapeutic proteins in recombinant host cells, the promoter may be a constitutive promoter such as the cytomegalous virus (CMV) promoter. Another example is LAMA2.

[0099] For the expression of therapeutic proteins described herein, provided are expression systems and constructs in the form of plasmids, expression vectors, transcription or expression cassettes that comprise at least one polynucleotide as described herein are also provided, as well host cells comprising such expression systems or constructs. As used herein, "vector" means any molecule or entity (*e.g.*, polynucleotide, plasmid, bacteriophage or virus) suitable for use to transfer protein coding information into a host cell. Examples of vectors include, but are not limited to, plasmids, viral vectors, non-episomal mammalian vectors and expression vectors, for example, recombinant expression vectors. Expression vectors, such as recombinant expression vectors, are useful for transformation of a host cell.

[00100] Host cells are provided into which an expression vector, such as a recombinant expression vector, has been introduced. A host cell can be any prokaryotic cell (for example, E. coli) or eukaryotic cell (for example, yeast, insect, or mammalian cells (e.g., CHO cells)). Expression vectors can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. For stable transfection of mammalian cells, a gene that encodes a selectable marker (e.g., for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced polynucleotide can be identified by drug selection, among other methods. Methods for introduction of heterologous polynucleotides into mammalian cells are well known in the art and include, but are not limited to, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast

fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, mixing nucleic acid with positively-charged lipids, and direct microinjection of the DNA into nuclei.

[00101] The method selected will in part be a function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan, and are set forth, for example, in Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001).

[00102] A host cell, when cultured under appropriate conditions, synthesizes protein that can be subsequently collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not soluble). The selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity (such as glycosylation or phosphorylation) and ease of folding into a biologically active molecule. As one example, Chinese hamster ovary cells overexpressing LARGE (CHO-LARGE cells) [Yoon et al., A Method to Produce and Purify Full-Length Recombinant Alpha Dystroglycan: Analysis of N- and O-Linked Monosaccharide Composition in CHO Cells with or without LARGE Overexpression, PLoS Curr. (2013 January 2)] are contemplated for use in producing glycosylated therapeutic proteins described herein.

[00103] Mammalian cell lines available as hosts for expression are well known in the art and include, but are not limited to, immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to Chinese hamster ovary (CHO) cells, CHO-LARGE cells, HEK293 cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and other cell lines standard in the art.

[00104] The rAAV genomes provided herein lack AAV rep and cap DNA. Recombinant AAV genomes provided herein comprise a polynucleotide encoding a therapeutic protein as described above and one or more AAV ITRs flanking the polynucleotide. AAV DNA in the rAAV genomes may be from any AAV serotype for which a recombinant virus can be derived including, but not limited to, AAV serotypes AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, AAV-7, AAV-8, AAV-9, AAV-10, AAV-11, AAV-12, AAV-13 and AAV rh.74. Other types of rAAV variants, for example rAAV with capsid mutations, are also contemplated. See, for example, Marsic *et al., Molecular Therapy, 22*(11): 1900-1909 (2014). As noted in the Background section above, the nucleotide sequences of the genomes of

various AAV serotypes are known in the art. To promote skeletal muscle specific expression, AAV1, AAV5, AAV6, AAV8 or AAV9 may be used.

[00105] DNA plasmids are provided that comprise rAAV genomes. The DNA plasmids are transferred to cells permissible for infection with a helper virus of AAV (including, but not limited to, adenovirus, E1-deleted adenovirus or herpesvirus) for assembly of the rAAV genome into infectious viral particles. Techniques to produce rAAV particles, in which an AAV genome to be packaged, rep and cap genes, and helper virus functions are provided to a cell are standard in the art. Production of rAAV requires that the following components are present within a single cell (denoted herein as a packaging cell): a rAAV genome, AAV rep and cap genes separate from (i.e., not in) the rAAV genome, and helper virus functions. The AAV ITRs and rep and cap genes may be from any AAV serotype for which recombinant virus can be derived and may be from a different AAV serotype than the rAAV genome ITRs, including, but not limited to, AAV serotypes AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, AAV-7, AAV-8, AAV-9, AAV-10, AAV-11, AAV-12, AAV-13 and AAV rh.74. Production of pseudotyped rAAV is disclosed in, for example, WO 01/83692 which is incorporated by reference herein in its entirety.

[00106] A method of generating a packaging cell is to create a cell line that stably expresses all the necessary components for AAV particle production. For example, a plasmid (or multiple plasmids) comprising a rAAV genome lacking AAV rep and cap genes, AAV rep and cap genes separate from the rAAV genome, and a selectable marker, such as a neomycin resistance gene, are integrated into the genome of a cell. AAV genomes have been introduced into bacterial plasmids by procedures such as GC tailing (Samulski et al., 1982, Proc. Natl. Acad. S6. USA, 79:2077-2081), addition of synthetic linkers containing restriction endonuclease cleavage sites (Laughlin et al., 1983, Gene, 23:65-73) or by direct, blunt-end ligation (Senapathy & Carter, 1984, J. Biol. Chem., 259:4661-4666). The packaging cell line is then infected with a helper virus such as adenovirus. The advantages of this method are that the cells are selectable and are suitable for large-scale production of rAAV. Other examples of suitable methods employ adenovirus or baculovirus rather than plasmids to introduce rAAV genomes and/or rep and cap genes into packaging cells.

[00107] General principles of rAAV production are reviewed in, for example, Carter, *Current Opinions in Biotechnology*, 1533-1539 (1992); and Muzyczka, *Curr. Topics in Microbial. And Immunol., 158*:97-129 (1992). Various approaches are described in Ratschin *et al., Mol. Cell. Biol., 4*:2072 (1984); Hermonat *et al., Proc. Natl. Acad. Sci. USA, 81*:6466

(1984); Tratschin *et al., Mol. Cell. Biol., 5*:3251 (1985); McLaughlin *et al., J. Virol., 62*:1963 (1988); Lebkowski *et al., Mol. Cell. Biol., 7*:349 (1988); Samulski *et al., J. Virol., 63*:3822-3828 (1989); U.S. Patent No. 5,173,414; WO 95/13365 and corresponding U.S. Patent No. 5,658.776 ; WO 95/13392; WO 96/17947; PCT/US98/18600; WO 97/09441 (PCT/US96/14423); WO 97/08298 (PCT/US96/13872); WO 97/21825 (PCT/US96/20777); WO 97/06243 (PCT/FR96/01064); WO 99/11764; Perrin *et al., Vaccine, 13*:1244-1250 (1995); Paul *et al., Human Gene Therapy, 4*:609-615 (1993); Clark *et al., Gene Therapy 3*:1124-1132 (1996); U.S. Patent. No. 5,786,211; U.S. Patent No. 5,871,982; U.S. Patent. No. 6,258,595; and McCarty, *Mol. Ther., 16*(10): 1648-1656 (2008). The foregoing documents are hereby incorporated by reference in their entirety herein, with particular emphasis on those sections of the documents relating to rAAV production.

[00108] The invention thus provides packaging cells that produce infectious rAAV. In one embodiment packaging cells may be stably transformed cancer cells such as HeLa cells, 293 cells and PerC.6 cells (a cognate 293 line). In another embodiment, packaging cells are cells that are not transformed cancer cells, such as low passage 293 cells (human fetal kidney cells transformed with E1 of adenovirus), MRC-5 cells (human fetal fibroblasts), WI-38 cells (human fetal fibroblasts), Vero cells (monkey kidney cells) and FRhL-2 cells (rhesus fetal lung cells).

[00109] Recombinant AAV, which herein are replication-deficient, infectious, encapsidated viral particles (rAAV), comprise a rAAV genome. A rAAV encodes a therapeutic protein described herein. The rAAV genomes lack AAV rep and cap DNA, that is, there is no AAV rep or cap DNA between the ITRs of the rAAV genomes.

[00110] The rAAV may be purified by methods standard in the art such as by column chromatography or cesium chloride gradients. Methods for purifying rAAV vectors from helper virus are known in the art and include methods disclosed in, for example, Clark *et al., Hum. Gene Ther., 10*(6): 1031-1039 (1999); Schenpp and Clark, *Methods Mol. Med., 69*: 427-443 (2002); U.S. Patent No. 6,566,118; and WO 98/09657.

[00111] In another embodiment, the invention contemplates compositions comprising rAAV or therapeutic protein described herein. Compositions of the invention comprise rAAV or therapeutic protein in a pharmaceutically acceptable carrier. The compositions may also comprise other ingredients such as diluents and adjuvants. Acceptable carriers, diluents and adjuvants are nontoxic to recipients and are preferably inert at the dosages and

concentrations employed, and include buffers such as phosphate, citrate, or other organic acids; antioxidants such as ascorbic acid; low molecular weight polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as Tween, pluronics or polyethylene glycol (PEG).

[00112] Titers of rAAV to be administered in methods described herein will vary depending, for example, on the particular rAAV, the mode of administration, the treatment goal, the individual, and the cell type(s) being targeted, and may be determined by methods standard in the art. Titers of rAAV may range from about $1x10^{10}$, about $1x10^{11}$, about $1x10^{12}$, about $1x10^{13}$ to about $1x10^{14}$ or more DNase resistant particles (DRP) per ml. Dosages may also be expressed in units of viral genomes (vg) as understood in the art.

[00113] Methods of transducing a target cell with rAAV, *in vivo* or *in vitro*, are contemplated by the invention. The *in vivo* methods comprise the step of administering an effective dose, or effective multiple doses, of a composition comprising a rAAV described herein to an animal (including a human patient) in need thereof.

[00114] Dosages and the frequency of administration of therapeutic proteins described herein may vary according to such factors as the route of administration, the particular therapeutic protein administered, and the size and general condition of the patient. Appropriate dosages can be determined by procedures known in the pertinent art, e.g., in clinical trials that may involve dose escalation studies. In view of these factors, a typical dose for a therapeutic protein described herein may range from about 0.1 pg/kg to up to about 30 mg/kg or more. Further, a dose may range from 0.1 pg/kg up to about 30 mg/kg, from 1 pg/kg up to about 30 mg/kg, from 10 pg/kg up to about 10 mg/kg, from about 0.1 mg/kg to 5 mg/kg, or from about 0.3 mg/kg to 3 mg/kg.

[00115] Methods of treating a patient with a therapeutic protein described herein are thus also provided. The methods comprise the step of administering an effective dose, or effective multiple doses, of a composition comprising a therapeutic protein described herein to an animal (including a human patient) in need thereof.

[00116] If the dose is administered prior to development of a disorder/disease, the administration is prophylactic. If the dose is administered after the development of a disorder/disease, the administration is therapeutic. An "effective dose" is a dose that alleviates (eliminates or reduces) at least one symptom associated with the disorder/disease state being treated, that slows or prevents progression to a disorder/disease state, that slows or prevents progression of a disorder/disease state, that diminishes the extent of disease, that results in remission (partial or total) of disease, and/or that prolongs survival. Methods described herein result in one or more of improved ambulation time, limb grip strength, decreased muscle pathology, and decreased neural pathology in a treated patient. Other endpoints achieved by methods described herein are one or more of increased muscle fiber size, decreased number of small oxidative fibers, correction of muscle atrophy, increased muscular force, and increased muscle regeneration in the treated patient. Dystroglycanopathies and laminin-deficient muscular dystrophies are contemplated for prevention or treatment according to methods of the invention.

[00117] Combination therapies are also contemplated by the invention. Combination therapies as used herein includes both simultaneous treatment, or sequential treatments. Combinations of methods described herein with standard medical treatments are specifically contemplated, as are combinations with novel therapies.

[00118] Administration of an effective dose of the compositions of rAAV or therapeutic protein may be by routes standard in the art including, but not limited to, intramuscular, intraparenteral, intravenous, intrathecal, oral, buccal, nasal, pulmonary, intracranial, intraosseous, intraocular, rectal, or vaginal. Route(s) of administration and serotype(s) of AAV components of the rAAV (in particular, the AAV ITRs and capsid protein) of the invention may be chosen and/or matched by those skilled in the art taking into account the infection and/or disease state being treated and the target cells/tissue(s) that are to express the therapeutic proteins.

[00119] In particular, actual administration of rAAV of the present invention may be accomplished by using any physical method that will transport the rAAV recombinant vector into the target tissue of an animal. Administration according to the invention includes, but is not limited to, injection into muscle, the bloodstream and/or directly into the liver. Simply resuspending a rAAV in phosphate buffered saline or lactated Ringer's solution has been demonstrated to be sufficient to provide a vehicle useful for muscle tissue expression, and there are no known restrictions on the carriers or other components that can be co-

administered with the rAAV (although compositions that degrade DNA should be avoided in the normal manner with rAAV). Capsid proteins of a rAAV may be modified so that the rAAV is targeted to a particular target tissue of interest such as muscle. See, for example, WO 02/053703, the disclosure of which is incorporated by reference herein. Pharmaceutical compositions can be prepared as injectable formulations or as topical formulations to be delivered to the muscles by transdermal transport. Numerous formulations for both intramuscular injection and transdermal transport have been previously developed and can be used in the practice of the invention. The rAAV can be used with any pharmaceutically acceptable carrier for ease of administration and handling.

[00120] For purposes of intramuscular injection, solutions of rAAV or therapeutic protein in an adjuvant such as sesame or peanut oil or in aqueous propylene glycol can be employed, as well as sterile aqueous solutions. Such aqueous solutions can be buffered, if desired, and the liquid diluent first rendered isotonic with saline or glucose. Solutions of rAAV as a free acid (DNA contains acidic phosphate groups) or a pharmacologically acceptable salt can be prepared in water suitably mixed with a surfactant such as hydroxpropylcellulose. A dispersion of rAAV can also be prepared in glycerol, liquid polyethylene glycols and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms. In this connection, the sterile aqueous media employed are all readily obtainable by standard techniques well-known to those skilled in the art.

[00121] The pharmaceutical forms of rAAV or therapeutic protein suitable for systemic (*e.g.*, intravenous) injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating actions of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of a dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal and the like. In many

cases it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by use of agents delaying absorption, for example, aluminum monostearate and gelatin.

[00122] Sterile injectable solutions are prepared by incorporating rAAV in the required amount in the appropriate solvent with various other ingredients enumerated above, as required, followed by filter sterilization. Generally, dispersions are prepared by incorporating the sterilized active ingredient into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze-drying technique that yield a powder of the active ingredient plus any additional desired ingredient from the previously sterile-filtered solution thereof.

[00123] Transduction with rAAV may also be carried out *in vitro*. In one embodiment, desired target muscle cells are removed from the subject, transduced with rAAV and reintroduced into the subject. Alternatively, syngeneic or xenogeneic muscle cells can be used where those cells will not generate an inappropriate immune response in the subject.

[00124] Suitable methods for the transduction and reintroduction of transduced cells into a subject are known in the art. In one embodiment, cells can be transduced *in vitro* by combining rAAV with muscle cells, *e.g.*, in appropriate media, and screening for those cells harboring the DNA of interest using conventional techniques such as Southern blots and/or PCR, or by using selectable markers. Transduced cells can then be formulated into pharmaceutical compositions, and the composition introduced into the subject by various techniques, such as by intramuscular, intravenous, subcutaneous and intraperitoneal injection, or by injection into smooth and cardiac muscle, using *e.g.*, a catheter.

[00125] Transduction of cells with rAAV of the invention results in sustained expression of a therapeutic protein described herein. The present invention thus provides methods of administering rAAV which express a therapeutic protein described herein to a patient, preferably a human being. These methods include transducing tissues (including, but not limited to, tissues such as muscle, organs such as liver and brain, and glands such as salivary glands) with one or more rAAV of the present invention.

[00126] Muscle tissue is an attractive target for *in vivo* DNA delivery, because it is not a vital organ and is easy to access. The invention contemplates sustained expression of therapeutic protein described herein from transduced muscle cells.

[00127] By "muscle cell" or "muscle tissue" is meant a cell or group of cells derived from muscle of any kind [for example, skeletal muscle and smooth muscle (*e.g.*, from the digestive tract, urinary bladder, blood vessels or cardiac tissue)]. Such muscle cells may be differentiated or undifferentiated, such as myoblasts, myocytes, myotubes, cardiomyocytes and cardiomyoblasts.

[00128] The term "transduction" is used to refer to the administration/delivery of therapeutic protein to a recipient cell either *in vivo* or *in vitro*, via a rAAV of the invention resulting in expression of therapeutic protein by the recipient cell.

[00129] Thus, methods are provided herein of administering an effective dose (or doses administered essentially simultaneously or doses given at intervals) of rAAV that encode a therapeutic protein described herein to a patient in need thereof.

[00130] Methods are also provided herein of administering an effective dose (or doses administered essentially simultaneously or doses given at intervals) of a therapeutic protein described herein to a patient in need thereof.

Examples

[00131] Aspects and embodiments of the invention are illustrated by the following examples. Example 1 describes constructs encoding therapeutic proteins of the disclosure. Example 2 describes recombinant expression of the therapeutic proteins in cultured host cells. Example 3 describes experiments demonstrating heparin-binding domain targets LAMA2(g1-G50 to the muscle of wild type mice. Example 4 describes experiments to demonstrate efficacy of AAV-mediated HBEGF-LAMA2(G1-5) expression in reducing symptoms and pathology in the dy^W/dy^W mouse model of MDC1A. Example 5 describes experiments to demonstrate efficacy of AAV-mediated HBEGF-DAG1(α) expression in reducing symptoms and pathology in the mouse models of dystroglycanopathy. Example 6 describes properties of (domains of) sHBEGF contemplated herein as useful for its application as a linker domain in therapeutic proteins described herein and as a trophic factor in various methods described herein.

Example 1

Constructs encoding therapeutic proteins

[00132] Six exemplary DNA constructs encoding therapeutic proteins including an HBEGF EGF domain were generated as follows:

HB-EGF (ending at heparin binding domain)-LAMA2 G1-G5 (encoded by the polynucleotide of Figure 3),

HB-EGF (complete soluble form)-LAMA2 G1-G5 (encoded by the polynucleotide of Figure 4),

HB-EGF (ending at heparin binding domain)-LAMA2 G3-G5 (encoded by the polynucleotide of Figure 5),

HB-EGF (complete soluble form)-LAMA2 G3-G5 (encoded by the polynucleotide of Figure 6),

HB-EGF (ending at heparin binding domain)-DAG1 (native processed alpha DG gene) (encoded by the polynucleotide of Figure 7), and

HB-EGF (complete soluble form)-DAG1 (native processed alpha DG gene) (encoded by the polynucleotide of Figure 8).

[00133] The constructs were expressed from plasmids in CHO cells. CHO cells were transfected with plasmids containing one of the constructs or mock-transfected (-).

[00134] The transfected CHO cells were stained with antibodies against HB-EGF, dystroglycan, or laminin-α2 G5 domain. Results are shown in Figure 9A. Also, culture media was collected from each plate 48 hours post-transfection and cell lysis was performed. Heparin-agarose pull-down was performed on both cell lysate and culture media and loaded in Western blot along with whole cell lysate. Results are shown in Figure 9B.

Example 2

Recombinant expression of therapeutic proteins in cultured host cells

[00135] The constructs of Example 1 were also subcloned into a plasmid to produce AAV9 vectors encoding the therapeutic proteins.

[00136] AAV vectors carrying one of the therapeutic genes of Example 1 under the transcriptional control of the cytomegalovirus (CMV) promoter were produced.

[00137] rAAV vectors were produced by a modified cross-packaging approach whereby the AAV type 2 vector genome can be packaged into multiple AAV capsid serotypes [Rabinowitz *et al., J Virol.* 76 (2):791-801 (2002)]. Production was accomplished using a

standard three plasmid DNA/CaPO4 precipitation method using HEK293 cells. HEK293 cells were maintained in DMEM supplemented with 10% fetal bovine serum (FBS) and penicillin and streptomycin. The production plasmids were: (i) plasmids encoding the therapeutic proteins, (ii) rep2-capX modified AAV helper plasmids encoding cap serotype 9 isolate, and (iii) an adenovirus type 5 helper plasmid (pAdhelper) expressing adenovirus E2A, E4 ORF6, and VA I/II RNA genes. A quantitative PCR-based titration method was used to determine an encapsidated vector genome (vg) titer utilizing a Prism 7500 Taqman detector system (PE Applied Biosystems). [Clark *et al., Hum Gene Ther.* 10 (6): 1031-1039 (1999)]. A final titer (vg ml^{-1}) was determined by quantitative reverse transcriptase PCR using the specific primers and probes utilizing a Prism 7500 Real-time detector system (PE Applied Biosystems, Grand Island, NY, USA). Aliquoted viruses were kept at −80 °C until use.

[00138] The rAAV set out in Table 2 were used in the experiments described herein.

Table 2

AAV construct	Membrane linker (first domain)	Second Domain	Denoted herein as	SEQ ID NO:	rAAV genome nt's
rAAV.CMV.HB.LAMA2(G1-G5)	Heparin binding domain only	Laminin alpha 2 G1-G5	HB-LAMA2(G1-G5)	2	Nt. 3590-8215
rAAV9.CMV.HBEGF.LAMA2(G1-G5)	Complete soluble form	Laminin alpha 2 G1-G5	HBEGF-LAMA2(G1-G5)	4	Nt. 3590-8341
rAAV.CMV.HB.LAMA2(G3-G5)	Heparin binding domain only	Laminin alpha 2 G3-G5	HB-LAMA2(G3-G5)	6	Nt. 3609-6929
rAAV.CMV.HBEGF.LAMA2(G3-G5)	Complete soluble form	Laminin alpha 2 G3-G5	HBEGF-LAMA2(G3-G5)	8	Nt. 3590 - 7036
rAAV.CMV.HB.DAG1	Heparin binding domain only	DAG1	HB-DAG1	10	Nt. 3590 to 6340
rAAV.CMV.HBEGF.DAG1	Complete soluble form	DAG1	HBEGF.DAG1	12	Nt. 3590 to 6049

Example 3

Heparin-binding Domain Targets LAMA2(G1-G5) to the muscle in Wild Type Mice

[00139] Wild type mice were injected intramuscularly in the gastrocnemius muscle with $5x10^{11}$vg pf rAAV9.CMV vectors containing HBEGF, HBEGF.LAMA2(G1-G5),

HBEGF.LAMA2(G3-G5), HB.LAMA2(G1-G5), HB.LAMA2(G3-G5), or LAMA2(G1-G5) (see Table 2). Cells were stained with antibody specific to human HB-EGF or recombinant G5 domain of LAMA2. Mock injected mice (buffer alone) are shown as a negative control. 4H8-2 is an anti-laminin antibody to show muscle cells in the sections.

[00140] As shown in Figure 15, HBEGF and HBEGF.LAMA2(G1-G5) reduced muscle growth and/or induced mild muscle atrophy, much as we had previously shown for overexpression of HBEGF, while IM injection of rAAV9.CMV.HB.LAMA2(G1-G5) lead to secretion and localization of LAMA2(G1-G5) protein in the extracellular matrix. In addition, HB. LAMA2(G1-G5)-expressing muscles appeared larger than normal wild type muscles. HB.LAMA2(G3-G5) showed lower overall protein staining than HB.LAMA2(G1-G5). The ECM targeting function of the HB domain of HBEGF allows for secretion and targeting of LAMA2(G1-G5) protein to the muscle extracellular matrix. By contrast, expression of LAMA2(G1-G5) without the HB domain led to very poor protein production and no detectable ECM localization. Thus, the constructs comprising the HB domain alone, rather than the full HB-EGF domain serves the purpose of targeting LAMA2(G1-G5) to the muscle ECM, and successfully does so without having the negative consequences of EGF signaling, as this construct has the EGF domain from HBEGF deleted.

[00141] The pre-pro peptide portion of HBEGF, which is still present in the HB construct, may also improve protein folding and/or expression for LAMA2(G1-G5) relative to LAMA2(G1-G5) alone, which only contains the signal peptide from HBEGF. Last, HB.LAMA2(G1-G5), when localized appropriately, may improve muscle growth, even in normal muscles.

[00142] Figure 16 shows staining of HBEGF.LAMA2(G1-G5), HB.LAMA2(G1-G5) and LAMA2(G1-G5) using an antibody to human HBEGF protein and an antibody to the G5 domain of human LAMA2. This data confirmed that expression of LAMA2(G1-G5) alone leads to very poor protein production in muscle, while inclusion of the HB domain improves expression for LAMA2(G1-G5), which was visualized with an anti-laminin antibody as well as an HBEGF antibody.

Example 4

Efficacy of AAV9-mediated HBEGF-LAMA2(G1-5) expression in reducing symptoms and pathology in the dy^W/dy^W mouse model of MDC1A

[00143] dy^W/dy^W mice [Nonaka, *Lab Anim Sci.*, 48(1):8-17 (1998)] have a loss-of-function mutation in *Lama2*, resulting in impaired laminin-α2 production, similar to MDC1A pathogenesis. Dy^W/dy^W mice have decreased size, grip strength, and lifespan compared to wild-type mice. They display muscle atrophy, dystrophic muscle pathology, and severely impaired ambulation by three months of age. As such, these mice are an appropriate and robust model for testing MDC1A therapy.

[00144] To demonstrate the therapeutic efficacy of the LAMA2 expressing rAAV genomes provided in Example 2, 8 dy^W/dy^W pups were injected intravenously through the facial vein at postnatal day 1 with either a low dose, 10^{11} viral genomes (vg), or a high dose, 10^{12} vg, rAAV9.CMV.HBEGF.LAMA2(G1-G5), or rAAV9.CMV.HB.LAMA2(G1-G5) or rAAV.CMV.HB.LAMA2(G3-G5). Mock injections of AAV buffer in control dy^W/dy^W pups were also performed.

[00145] At 2 and at 3 months post-injection, grip strength in the forelimb muscles was analyzed (Figure 17). At 4 months of age, the mice were euthanized and limb muscles were dissected and analyzed for expression of recombinant protein. As shown in Fig. 17, both HB.LAMA2(G3-G5) and HB.LAMA2(G1-G5) prevented loss of grip strength in dy/dy mice, showing a significant change from mock-treated dy/dy mice and bringing strength values to within the range seen in untreated wild type mice of the same age. Thus, both HB-LAMA2(G3-G5) and HB-LAMA2(G1-G5) show a therapeutic effect in the dy/dy model for MDC1A. In this experiment, HB.LAMA2(G3-G5) did not reach significance at 2 months relative to mock-treated dy/dy mice, while HB.LAMA2(G1-G5) did.

[00146] The role of transgene expression in preventing muscle pathology by comparing the percentage of myofibers with central nuclei, myofiber diameter and area, and variance in myofiber diameter in treated dy/dy mice was also analyzed. Muscle pathology intransgene-expressing myofibers was compared to the same pathology measures in non-expressing myofibers using the triceps muscle. This experiment demonstrated that expression of HB.LAMA2(G1-G5) increased muscle size. An example of staining showing such changes is shown in Figure 18. When quantified across all muscles, the average cross-sectional muscle area was $2328mm^2$ in expressing myofibers versus $1082mm^2$ in nonexpressing myofibers (n=4 muscles each with 400 myofibers analyzed per muscle), which was a two-fold average increase in muscle size with treatment. The variance in myofiber diameter index (Diameter SD/Mean X1000) was reduced from 620 in non-expressing muscles to 431 in expressing muscles (250 or lower is considered normal), and the percentage of myofibers

with central nuclei, an indicator of a cycle of muscle degeneration and regeneration, was reduced from 28% in non-expressing myofibers to 14% in expressing ones (n=2 each).

[00147] While not reduced to zero pathology, it is important to remember that AAV requires 3-4 weeks to achieve maximal gene expression, so some pathology will develop in these animals prior to when therapeutic transgene expression occurs. In all such experiments, the average level of muscle transduction was 26±1% (n=4 triceps muscles analyzed, 400 fibers each). The take home from these pathological measures is that HB-LAMA2(G1-G5) not only appears to prevent, at least in part, muscle damage in dy/dy muscles, but it also increases muscle growth back to, and perhaps beyond, wild type levels.

Example 5

Efficacy of AAV-mediated HBEGF-DAG1(α) expression in reducing symptoms and pathology in mouse models of dystroglycanopathy

[00148] Mice lacking dystroglycan or the α-dystroglycan-glycosylating enzyme, fukutin, encoded by the *Fktn* gene, are embryonic lethal, and cannot be used to study dystroglycanopathy therapy. Myf5Cre-Fktn^{loxP} mice [Kanagawa *et al., Hum Mol Genet.*, 22(15):3003-3015 (2013)], in which *Fktn* deletion is restricted to skeletal muscle are used to demonstrate efficacy. Myf5Cre-Fktn^{loxP} mice have decreased body weight, grip strength, and lifespan compared to wild-type mice. They also display dystrophic muscle pathology.

[00149] To demonstrate the therapeutic efficacy of rAAV9.CMV.HBEGF-DAG1(α) in the Myf5Cre-Fktn^{loxP} mouse model, the same injection protocol and assessments are performed as described in Example 3.

[00150] Another mouse model of dystroglycanopathy the Large-vls mouse mutant (Lee et al., Mol. Cell. Neurosci. 30: 160-172, 2005). Several Large vls mice were IM injected with 1x10^{12}vg rAAV9.CMV.HB-DAG1 IV via the facial vein at P1. The tests of grip strength of these mice suggest potential improvement. Of the 4-7 animals analyzed per group at 2 months, forelimb grip strength is reduced from 4.7±.2 g/g in wild type to 3.8±0.1g/g in untreated Large vls mice (p=0.0005), while pAAV9.CMV.HB-DAG1(α) treatment of Large vls mice (IV with 1x10^{12}vg at P1) showed improvement in grip strength, to 4.4±.3g/g (p=0.06 versus mock-treated Large vls). This data is very close to significance, and a significant difference may be achieved with additional measures.

Example 6

sHBEGF as a Linker Protein and Trophic Factor

[00151] Using the heparin-binding domain and EGF-like domain of soluble HBEGF (sHBEGF) in various exemplary therapeutic proteins described herein provides a dual benefit to patients. Including both domains adds increased muscle membrane stability from the inclusion of the heparin-binding domain along with LAMA2(G1-5) or DAG1(α), and then inclusion of the EGF-like domain additionally provides a stimulus for muscle regeneration. Figures 10, 11 and 12 show sHBEGF activates an Akt kinase pathway in muscle and increases the expression of muscle regeneration markers. Expression of activated Akt kinase in muscle has previously been shown to stimulate profound muscle growth, akin to what is seen with myostatin inhibitors. The presence of the EGF-like domain of HBEGF in various therapeutic proteins described herein, therefore, adds an additional therapeutic effect for treatment of the diseases described herein.

[00152] The gastrocnemius muscle on the left side of 5-week-old male C57BL/6J mice was injected with $5x10^{10}$ vector genomes of r(ds)AAV9.CMV.HB-EGF or r(ds)AAV9.CMV.sHB-EGF in a volume of 50 µL sterile PBS using a 0.3 mL insulin syringe near the midpoint of the muscle. Muscles on the contralateral (right) side of the mouse were mock-injected with an identical volume of sterile PBS. At 4- or 12-weeks post-injection, mice were sacrificed and dissected. Gastrocnemius muscles were embedded in O.C.T. Compound (Fisher Scientific, Pittsburgh, PA) and snap-frozen in liquid nitrogen-cooled isopentane.

[00153] sHB-EGF expression was visualized using an antibody that recognizes sHB-EGF and co-stained with either an antibody to Galgt2 protein or the CT glycan. sHB-EGF was expressed along the sarcolemmal membrane of skeletal myofibers in muscles analyzed at 4 weeks post-injection with r(ds)AAV9.CMV.sHB-EGF. Figure 9 shows that sHB-EGF can be secreted from muscles and stick to the extracellular matrix, supporting its use as a linker protein.

[00154] Figure 10 shows that sHB-EGF induces expression of therapeutic surrogate muscular dystrophy genes.

[00155] Figure 11 shows that sHB-EGF induces Akt tyrosine kinase cascade in skeletal muscle and can stimulate muscle growth and regeneration.

[00156] While the present invention has been described in terms of specific embodiments, it is understood that variations and modifications will occur to those skilled in the art. Accordingly, only such limitations as appear in the claims should be placed on the invention.

[00157] All documents referred to in this application are hereby incorporated by reference in their entirety.

CLAIMS

1. A polynucleotide encoding a protein comprising:

a) a first domain comprising the heparin-binding domain of Heparin-Binding Epidermal Growth Factor-Like Growth Factor (HBEGF) encoded by the nucleotide sequence of SEQ ID NO: 13, and a second domain comprising the G1-G5 domain of the human laminin alpha 2 (LAMA2) encoded by the nucleotide sequence of SEQ ID NO: 15;

b) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 14, and a second domain comprising the G1-G5 domain of the human LAMA2 encoded by the nucleotide sequence of SEQ ID NO: 15;

c) a first domain comprising the heparin-binding domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 13, and a second domain comprising the G3-G5 domain of the human LAMA2 gene encoded by the nucleotide sequence of SEQ ID NO: 16;

d) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 14, and a second domain comprising the G3-G5 domain of the human LAMA2 encoded by the nucleotide sequence of SEQ ID NO: 16;

e) a first domain comprising the heparin-binding domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 13, and a second domain comprising the processed native alpha chain of the human dystroglycan gene (DAG1alpha) encoded by the nucleotide sequence of SEQ ID NO: 17; or

f) a first domain comprising the heparin-binding domain of HBEGF and the EGF-like domain of HBEGF encoded by the nucleotide sequence of SEQ ID NO: 14, and a second domain comprising DAG1alpha encoded by the nucleotide sequence of SEQ ID NO: 17.

2. The polynucleotide of claim 1 comprising i) nucleotides 14 to 3235 set forth in Figure 3, ii) the nucleotide sequence of SEQ ID NO: 1, or iii) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 19.

3. The polynucleotide of claim 1 comprising i) nucleotides 14 to 3361 set forth in Figure 4, ii) the nucleotide sequence of SEQ ID NO: 3 or iii) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 20.

4. The polynucleotide of claim 1 comprising i) nucleotides 14 to 1930 set forth in Figure 5, ii) the nucleotide sequence of SEQ ID NO: 5 or iii) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 21.

5. The polynucleotide of claim 1 comprising i) nucleotides 14 to 2056 set forth in Figure 6, the nucleotide sequence of SEQ ID NO: 7 or iii) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 22.

6. The polynucleotide of claim 1 comprising i) nucleotides 14 to 1360 set forth in Figure 7, ii) the nucleotide sequence of SEQ ID NO: 9 or iii) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 23.

7. The polynucleotide of claim 1 comprising i) nucleotides 14-1486 set forth in Figure 8, ii) the nucleotide sequence of SEQ ID NO: 11 or iii) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 24.

8. A recombinant adeno-associate virus (rAAV), wherein the genome of the rAAV comprises the polynucleotide of any one of claims 1-7.

9. A recombinant adeno-associate virus (rAAV), wherein the genome of the rAAV comprises

a) a polynucleotide sequence of SEQ ID NO: 1,

b) a polynucleotide sequence of SEQ ID NO: 3,

c) a polynucleotide sequence of SEQ ID NO: 5,

d) a polynucleotide sequence of SEQ ID NO: 7,

e) a polynucleotide sequence of SEQ ID NO: 9,

f) a polynucleotide sequence of SEQ ID NO: 11,

g) nucleotides 3590 to 8215 of SEQ ID NO: 2,

h) nucleotides 3590 to 8341 of SEQ ID NO: 4,

i) nucleotides 3609 to 6929 of SEQ ID NO: 6,

j) nucleotides 3590 to7036 of SEQ ID NO: 8,

k) nucleotides 3590 to 6340 of SEQ ID NO: 10,

l) nucleotides 3590 to 6049 of SEQ ID NO: 12,

m) the nucleotide sequence set out in Figure 13, or

n) the nucleotide sequence set out in Figure 14.

10. An rAAV of claim 8 or claim 9, wherein the genome of the rAAV further comprises a muscle-specific transcriptional control element.

11. An rAAV of claim 10, wherein the muscle-specific transcriptional control elements is the CMV promoter.

12. The rAAV of any one of claims 8 to 11, comprising AAV9, AAV10, AAVrh74, AAV8 or AAV6 capsid.

13. An rAAV particle comprising an rAAV of any one of claims 8 to 12.

14. A recombinant host cell comprising the polynucleotide of any one of claims 1-7 optionally wherein the host cell is a Chinese hamster ovary (CHO) cell or HEK293 cell.

15. A protein encoded by the polynucleotide of any one of claims 1 to 7.

16. A composition comprising a polynucleotide of any one of claims 1 to 7, an rAAV of any one of claims 8 to 12, an rAAV particle of claim 13 or a protein of claim 15.

17. A method for treating a laminin-deficient muscular dystrophy comprising administering to a patient in need thereof a polynucleotide of any one of claims 1 to 7, an rAAV of any one of claims 8 to 12, an rAAV particle of claim 13, a protein of claim 15 or a composition of claim 16.

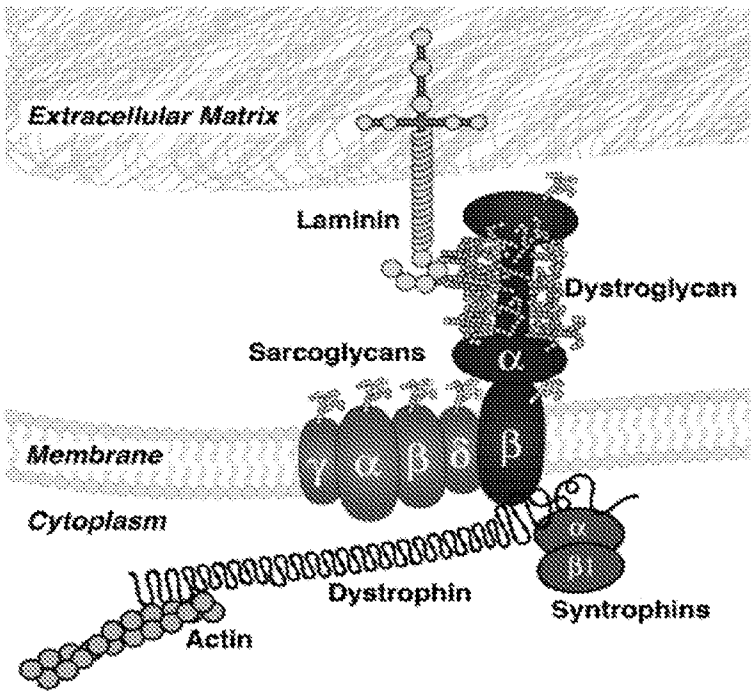
18. Use of a polynucleotide of any one of claims 1 to 7, an rAAV of any one of claims 8 to 12, an rAAV particle of claim 13, a protein of claim 15 or a composition of claim 16 in the preparation of a medicament for the treatment of a laminin-deficient muscular dystrophy.

19. The method of claim 17 or use of 18 wherein the laminin-deficient muscular dystrophy is MDC1A.

20. The method of claim 17 or the use of claim 18 wherein the laminin-deficient muscular dystrophy is Walker Warburg syndrome, Muscle Eye Brain disease, Fukuyama Congenital Muscular Dystrophy, MDC1C, MDC1D, LGMD2I, LGMD2K, LGMD2M, LGMD2N, LGMD2O, LGMD2P, LGMD2T or LGMD2U.

Figure 1

A


Extracellular Matrix
Laminin
Dystroglycan
Sarcoglycans
α
γ α β δ β
Membrane
Cytoplasm
α
β1
Dystrophin
Syntrophins
Actin

B

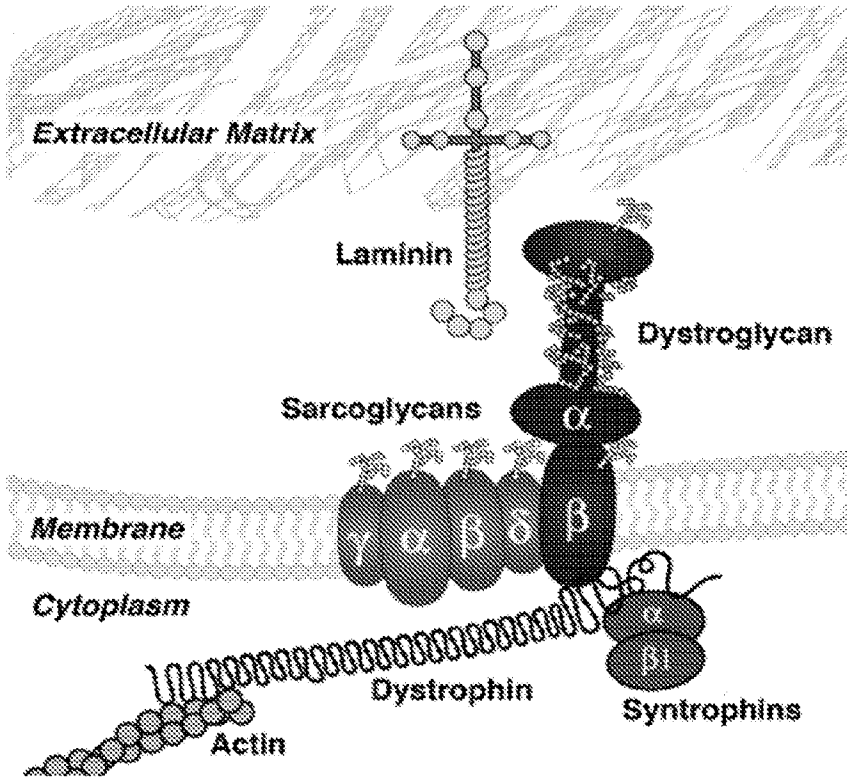

Extracellular Matrix
Laminin
Dystroglycan
Sarcoglycans
α
γ α β δ β
Membrane
Cytoplasm
α
β1
Dystrophin
Syntrophins
Actin

Figure 1 (Continued)

C

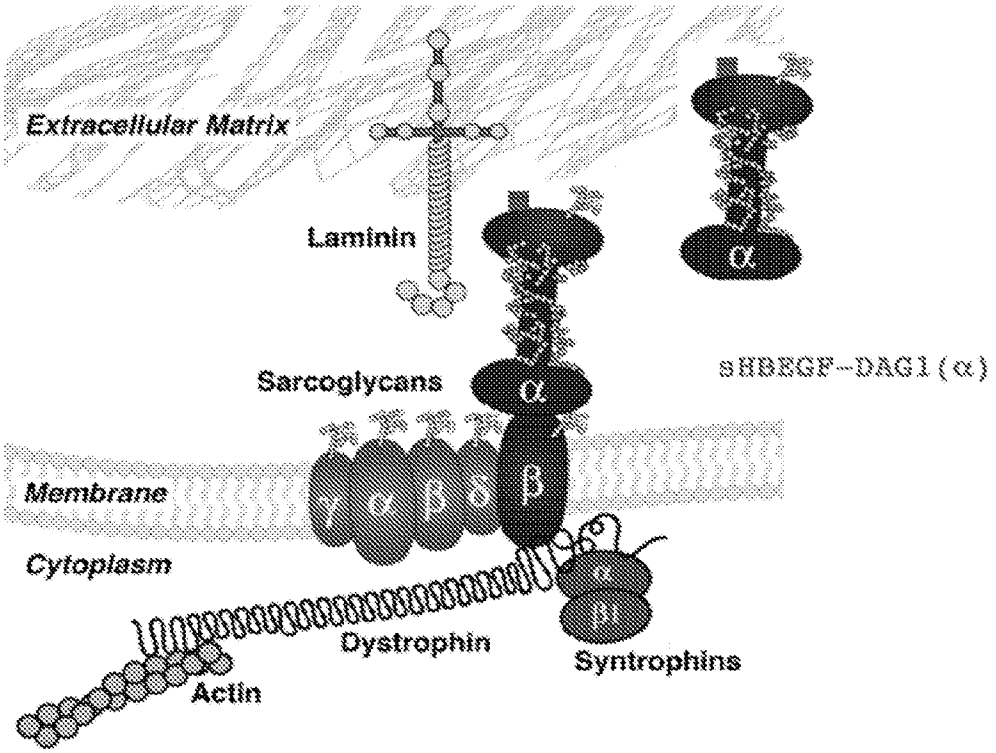


Figure 2

A

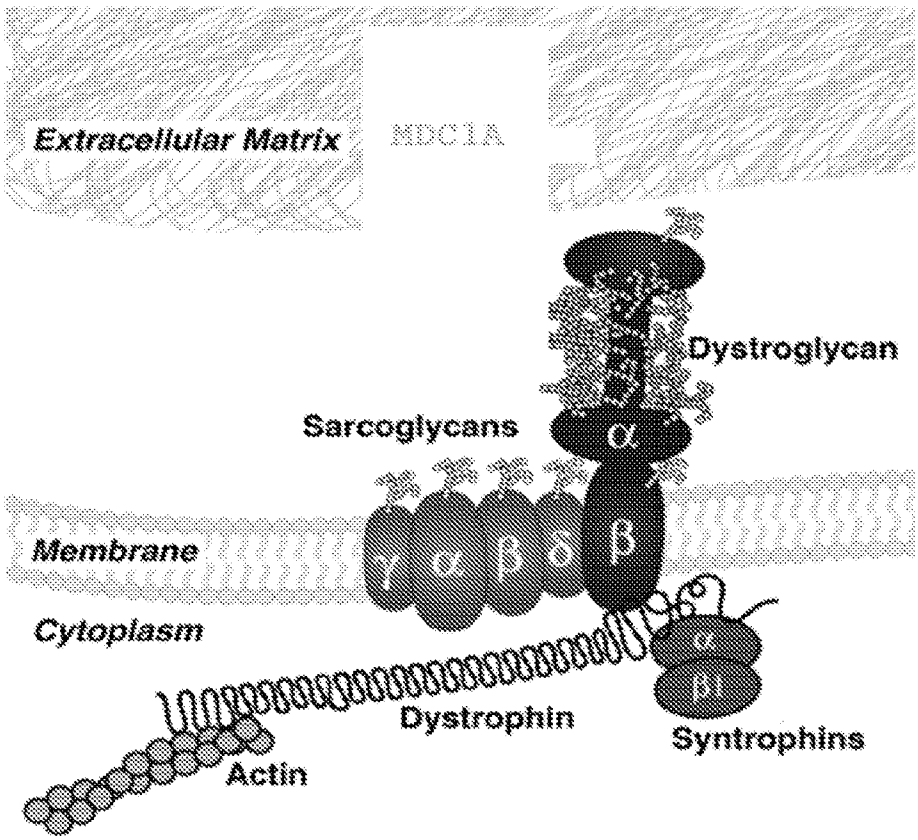


Figure 2 cont.

B

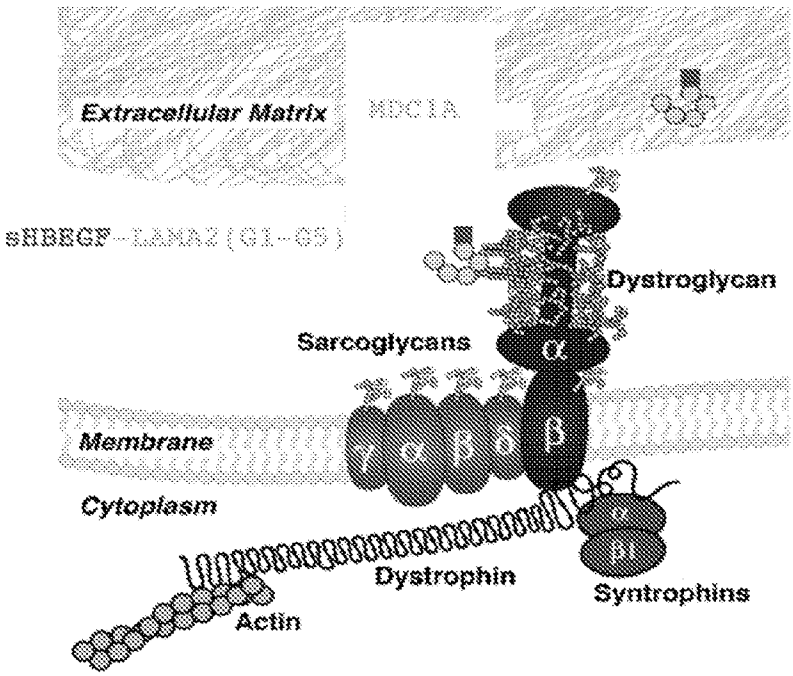


Figure 3

HB-EGF (ending at heparin binding domain)-LAMA2 G1-G5

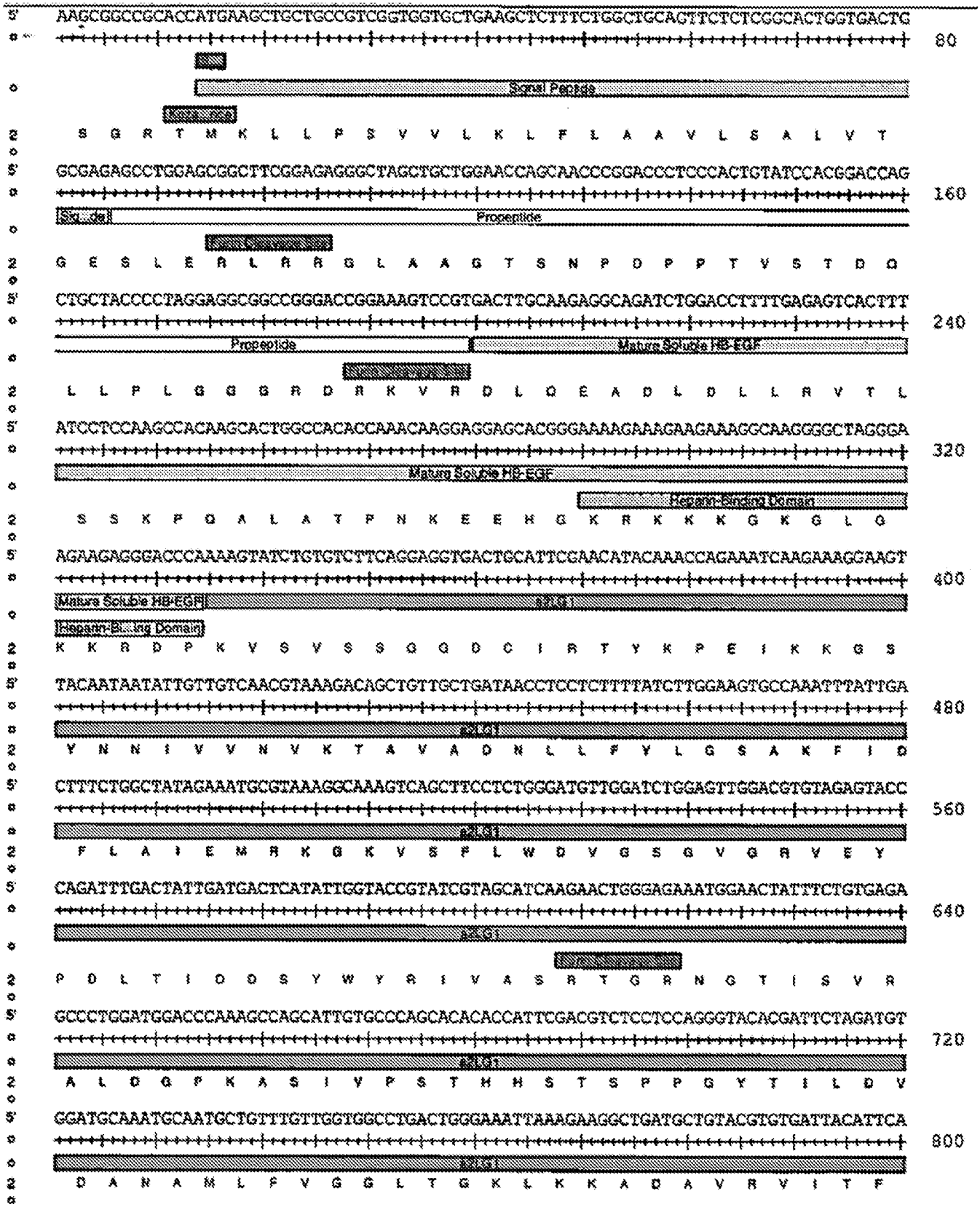


Figure 3 (Continued)

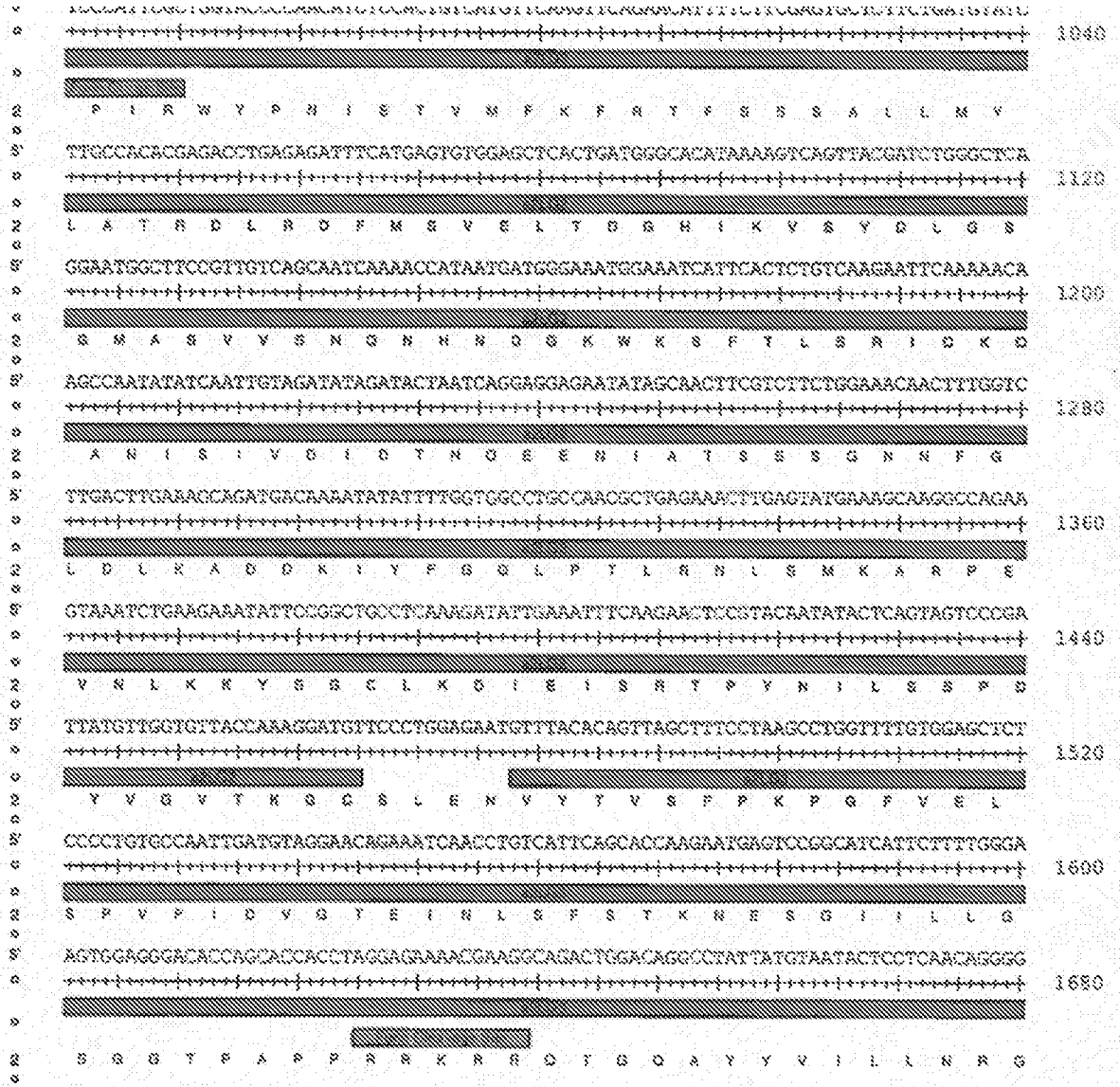
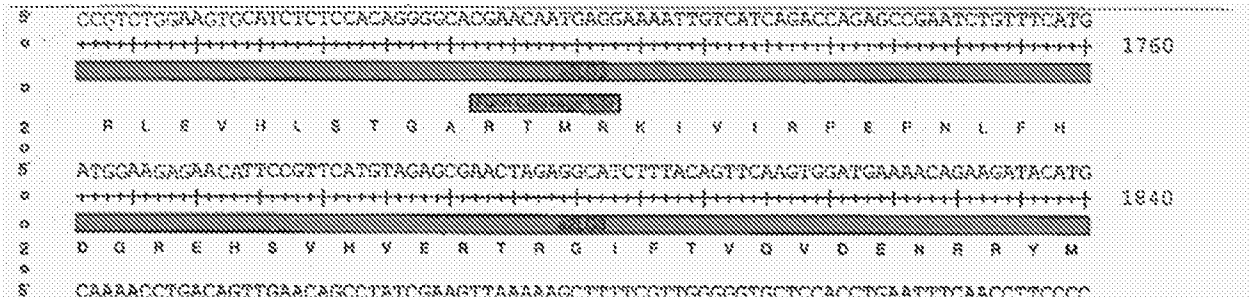


Figure 3 (Continued)



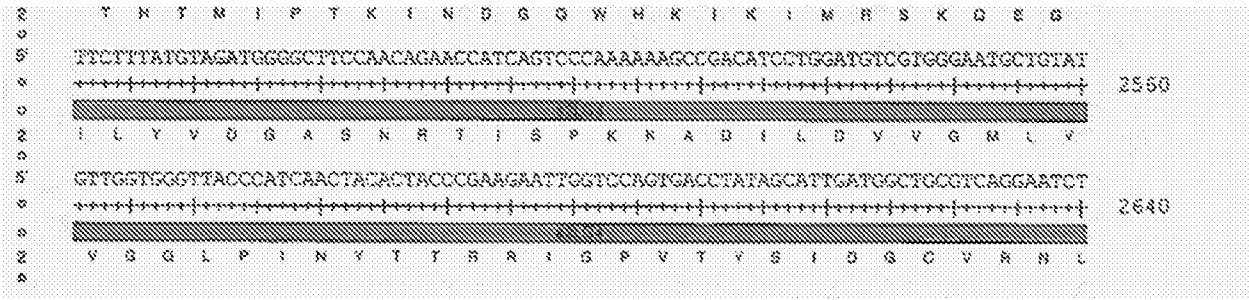


Figure 3 (Continued)

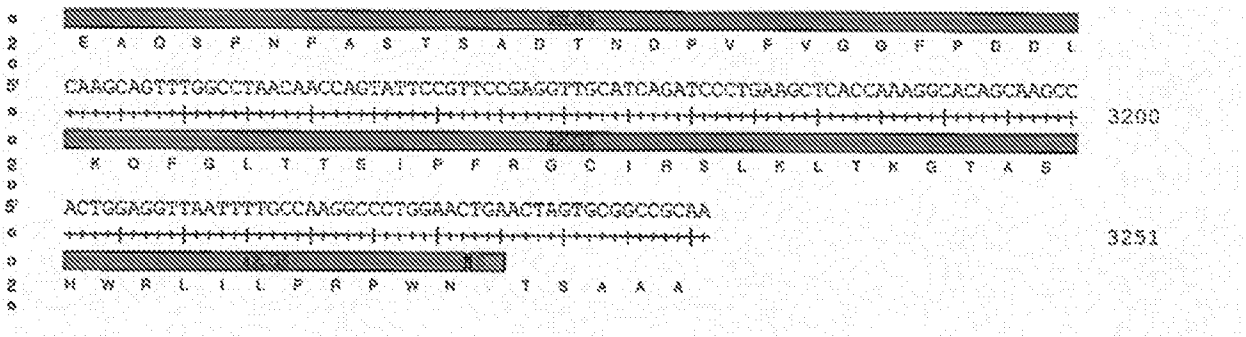


Figure 4

HB-EGF (complete soluble form)-LAMA2 G1-G5

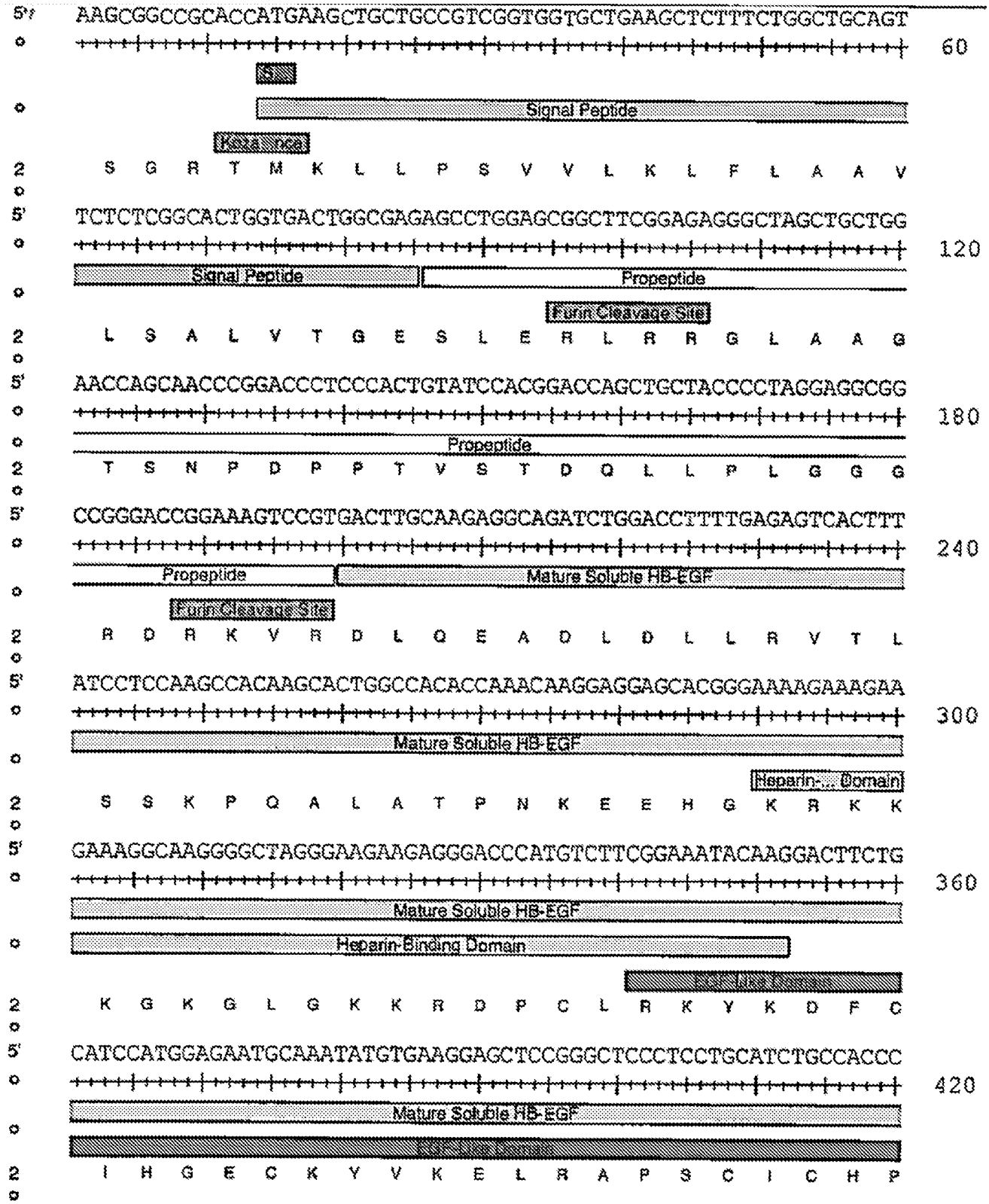


Figure 4 (Continued)

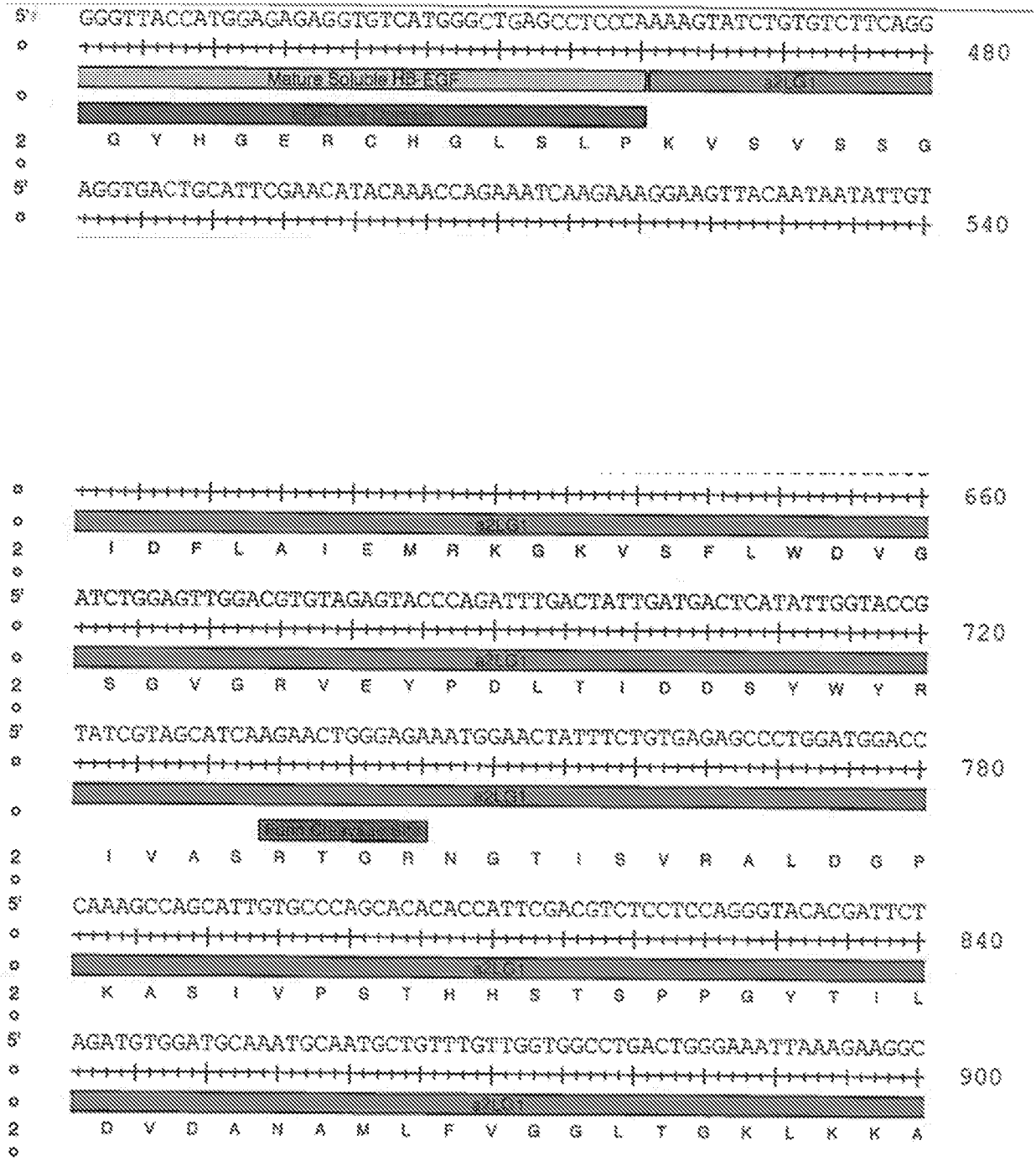


Figure 4 (Continued)

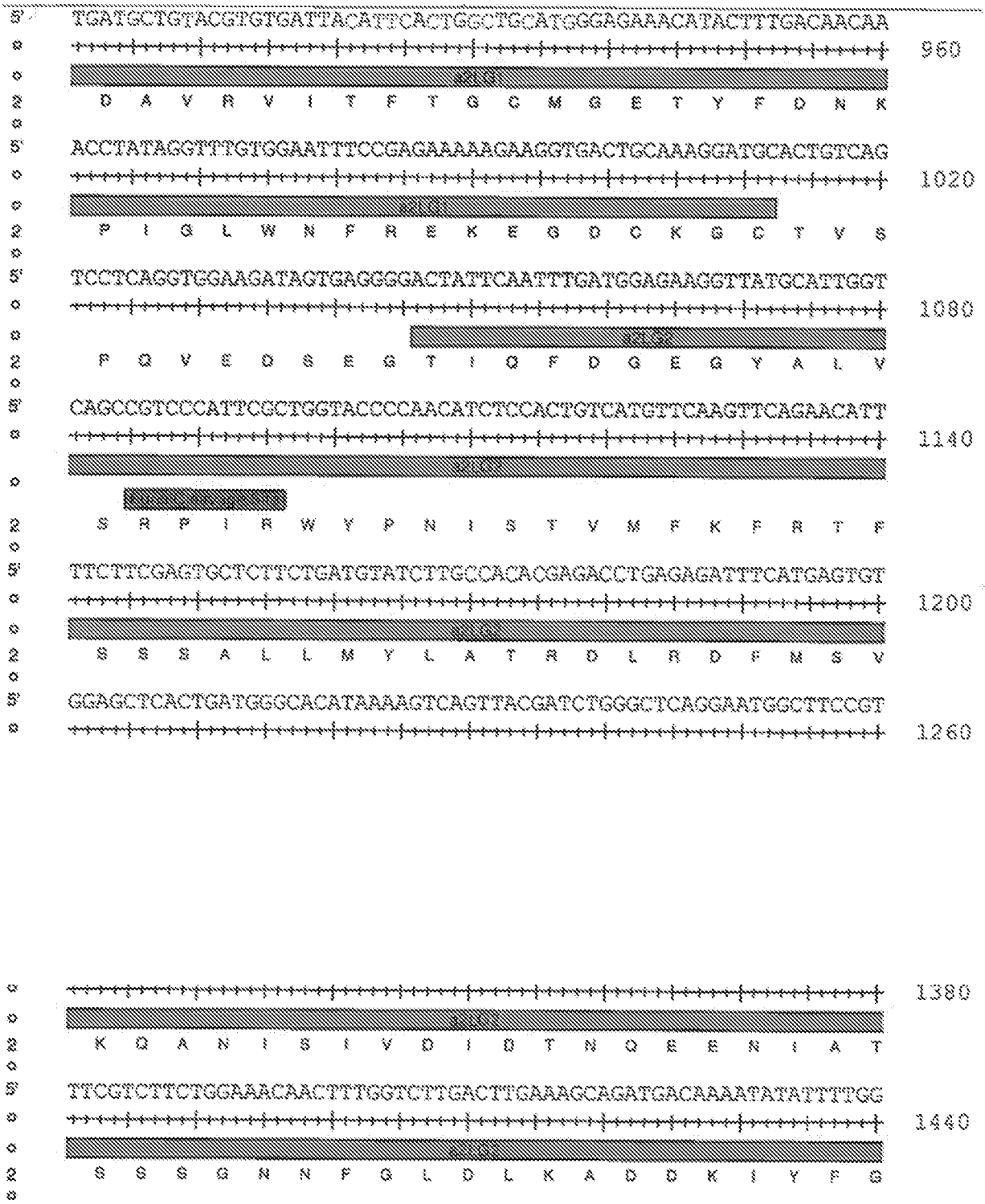


Figure 4 (Continued)

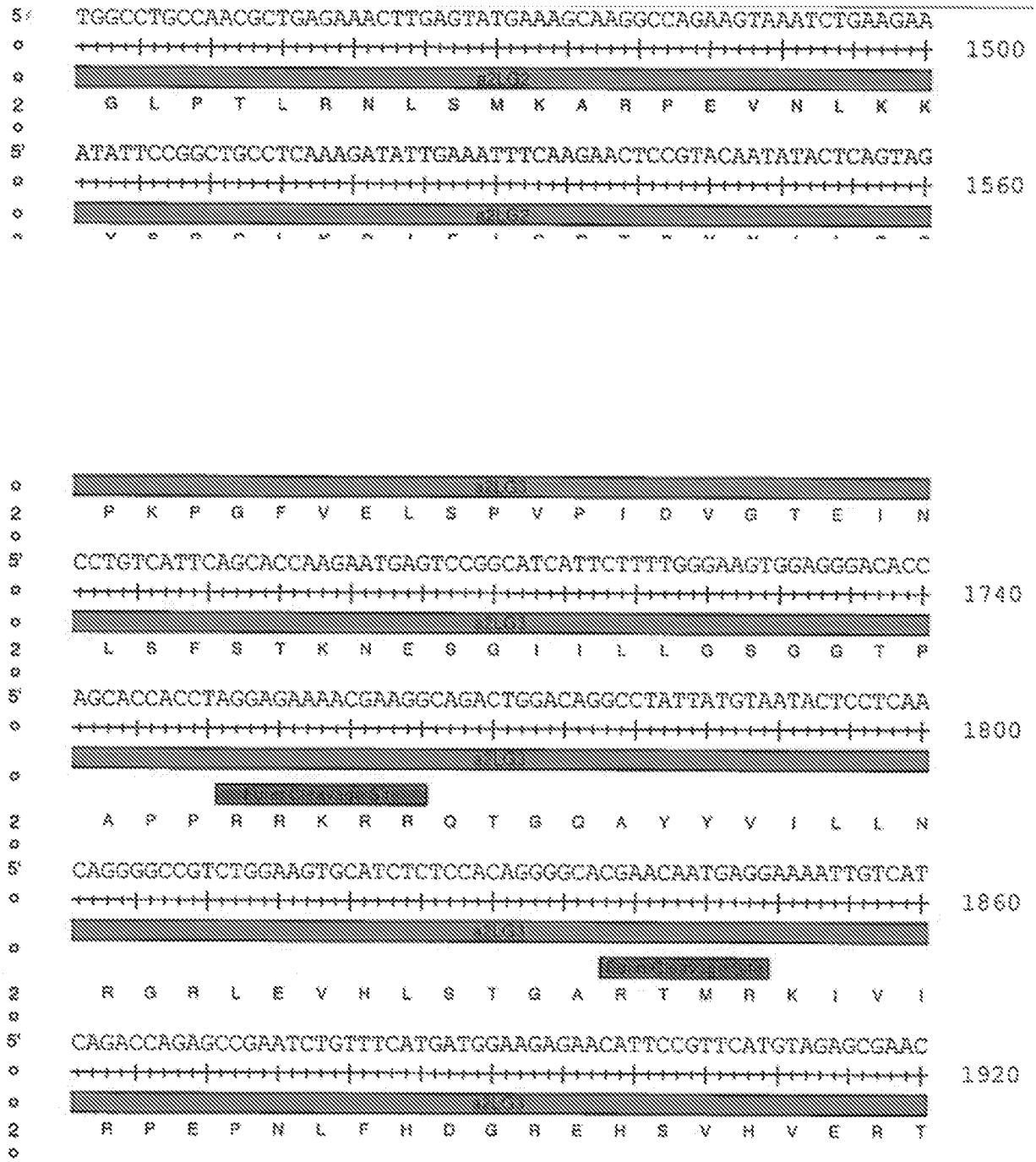


Figure 4 (Continued)

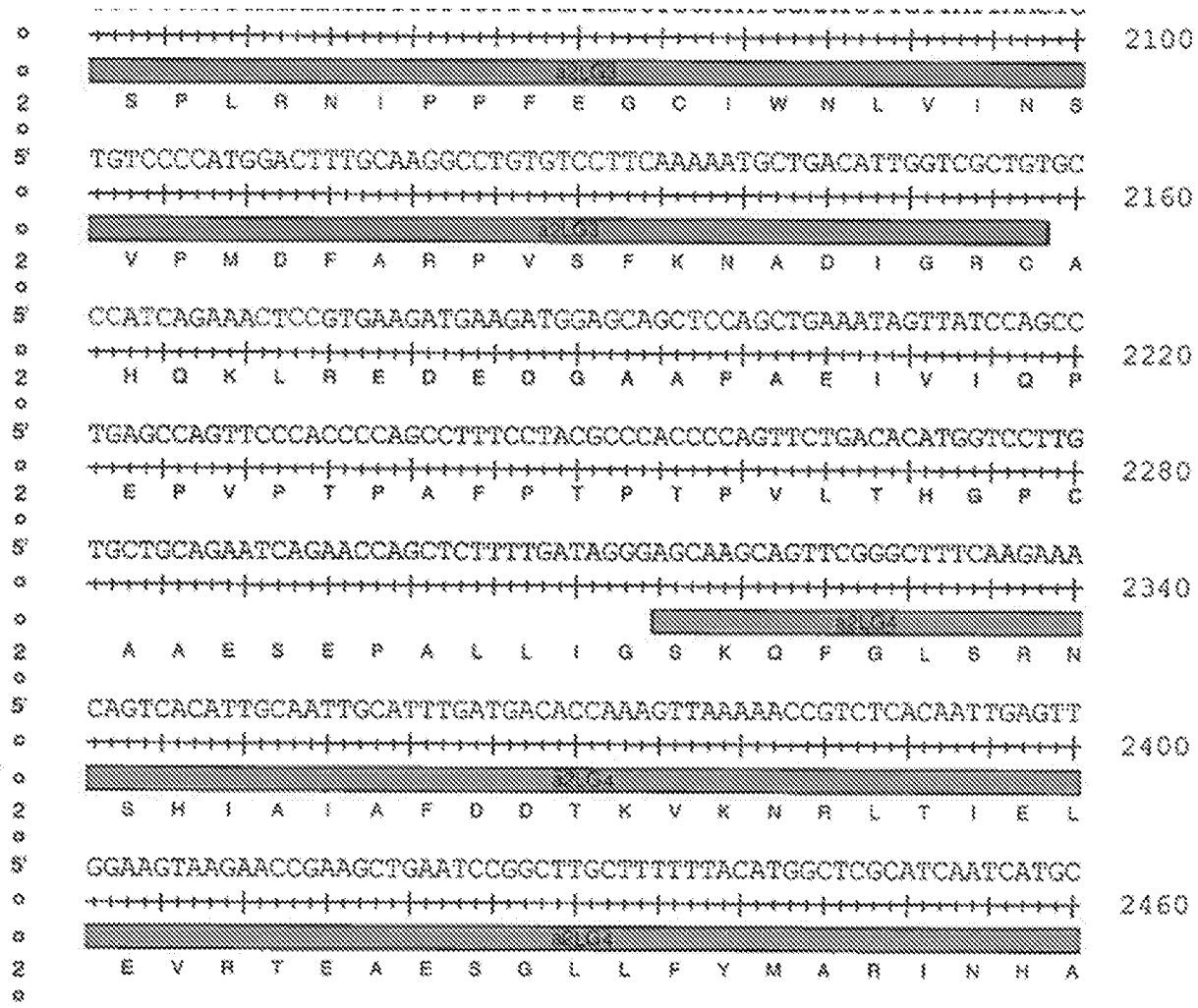


Figure 4 (Continued)

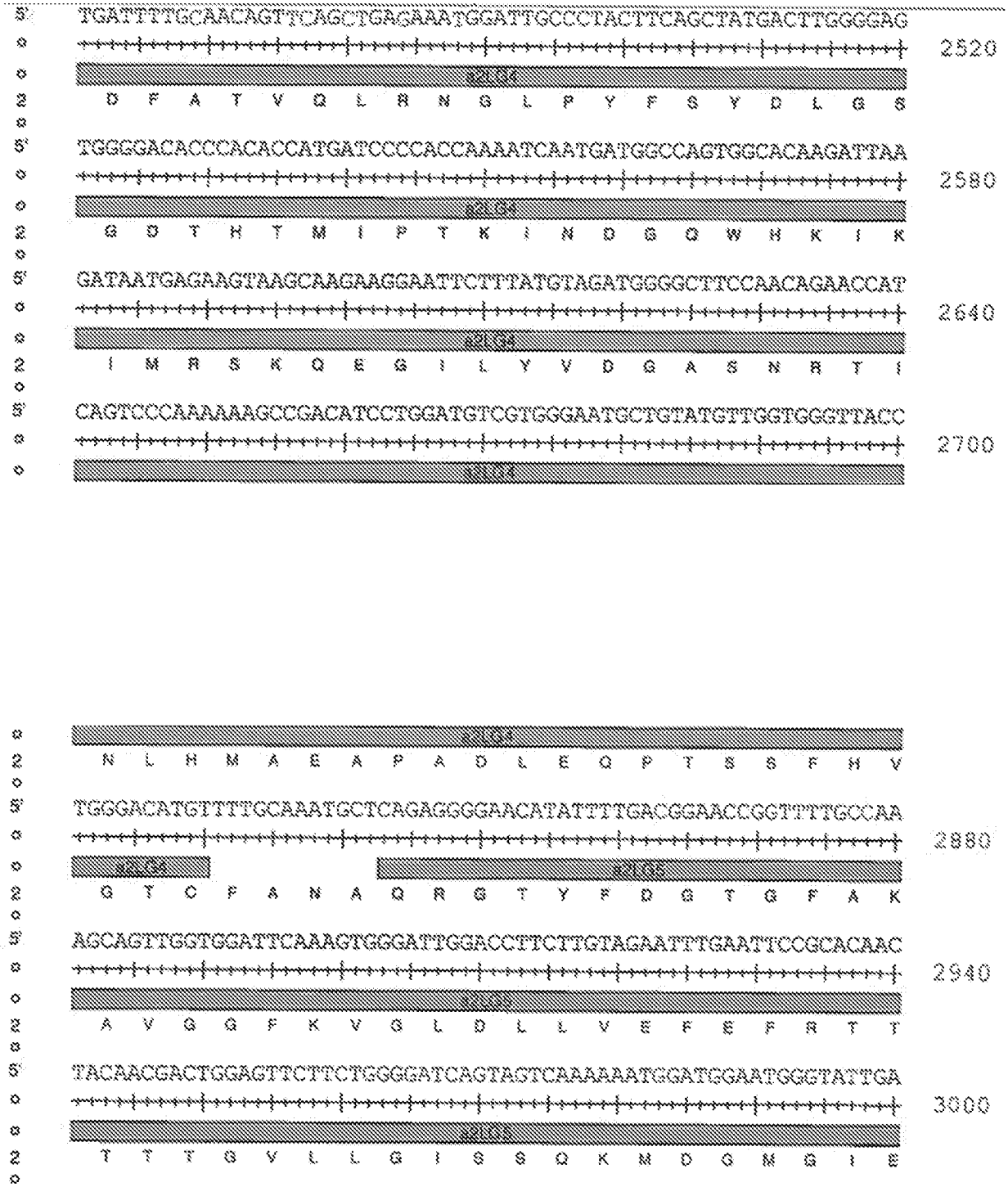


Figure 4 (Continued)

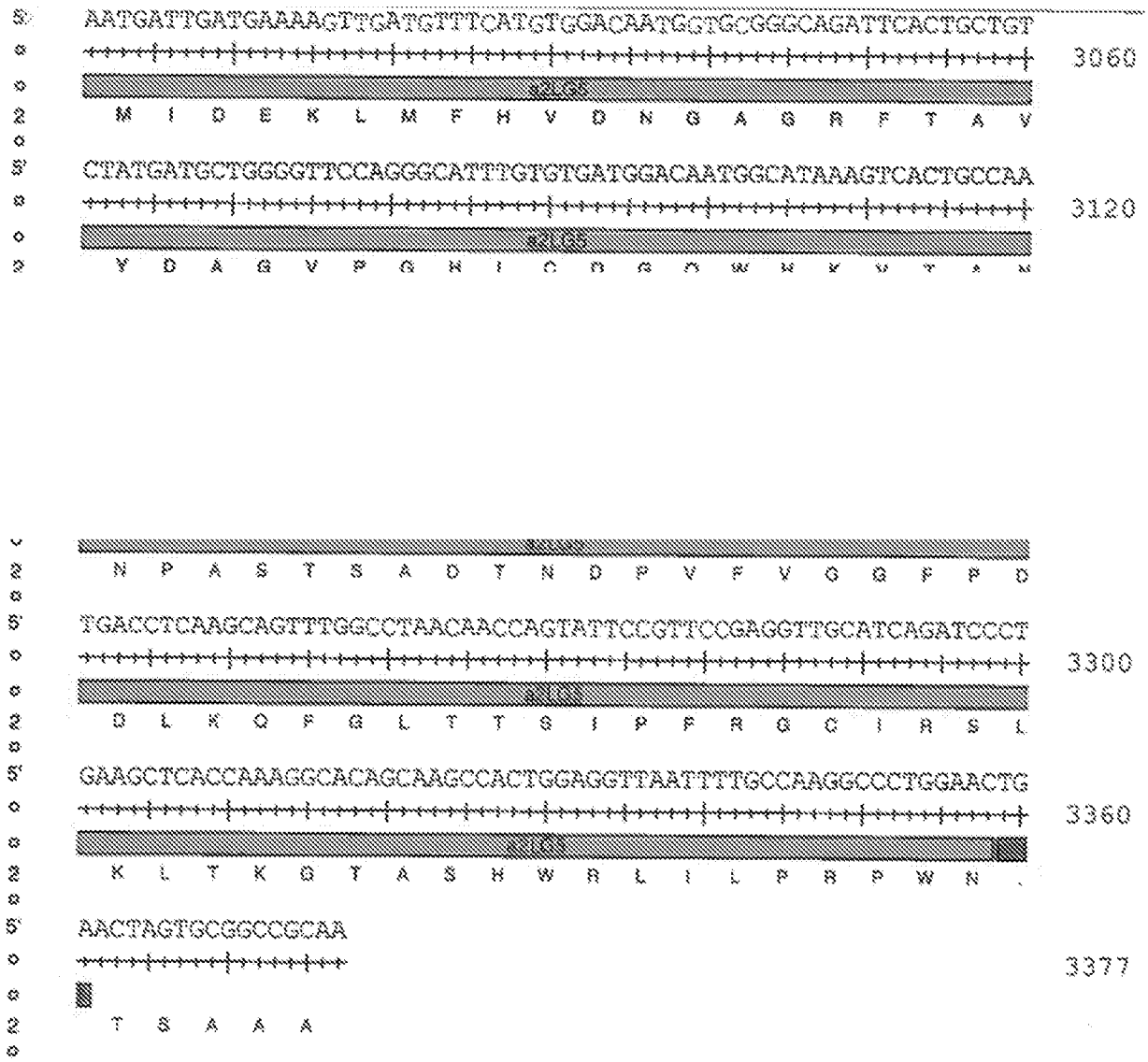


Figure 5

HB-EGF (ending at heparin binding domain)-LAMA2 G3-G5

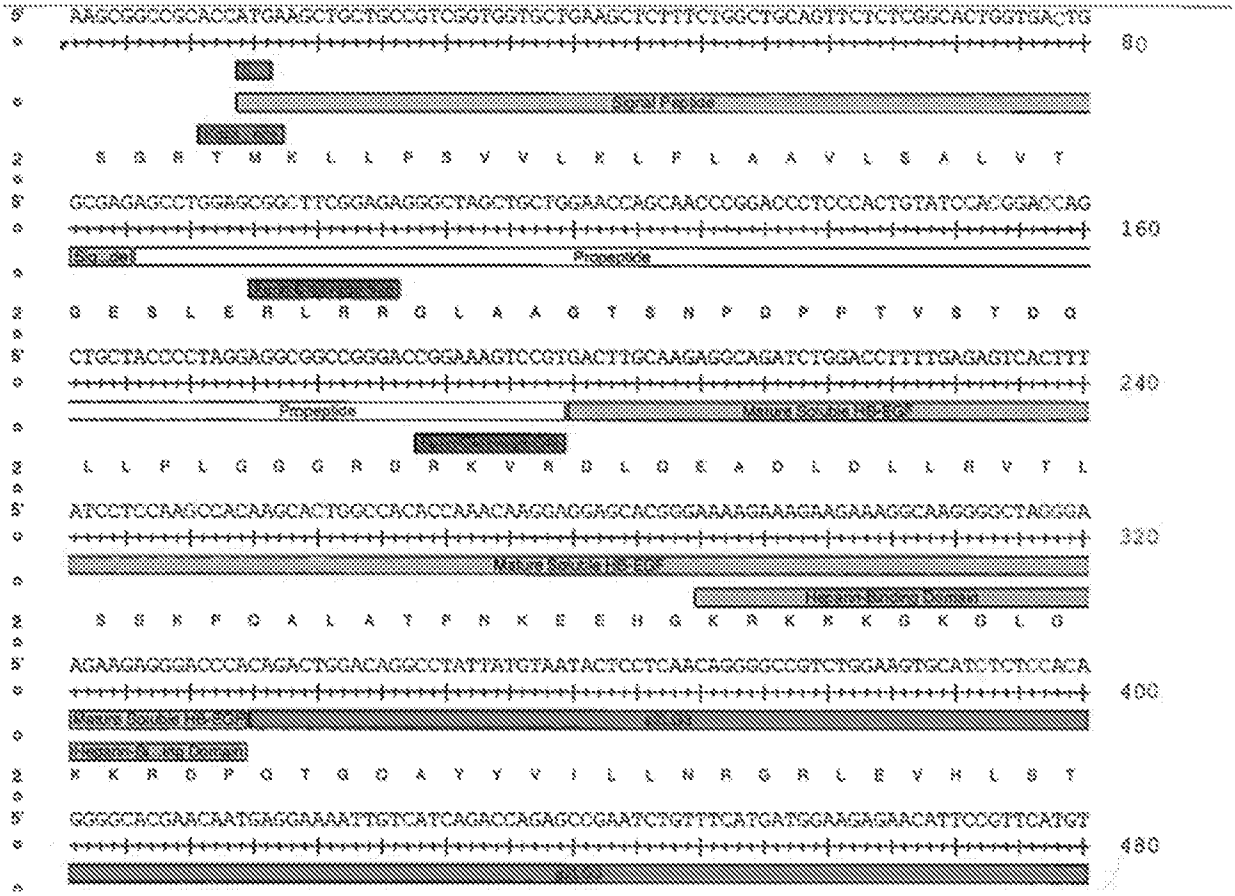


Figure 5 (Continued)

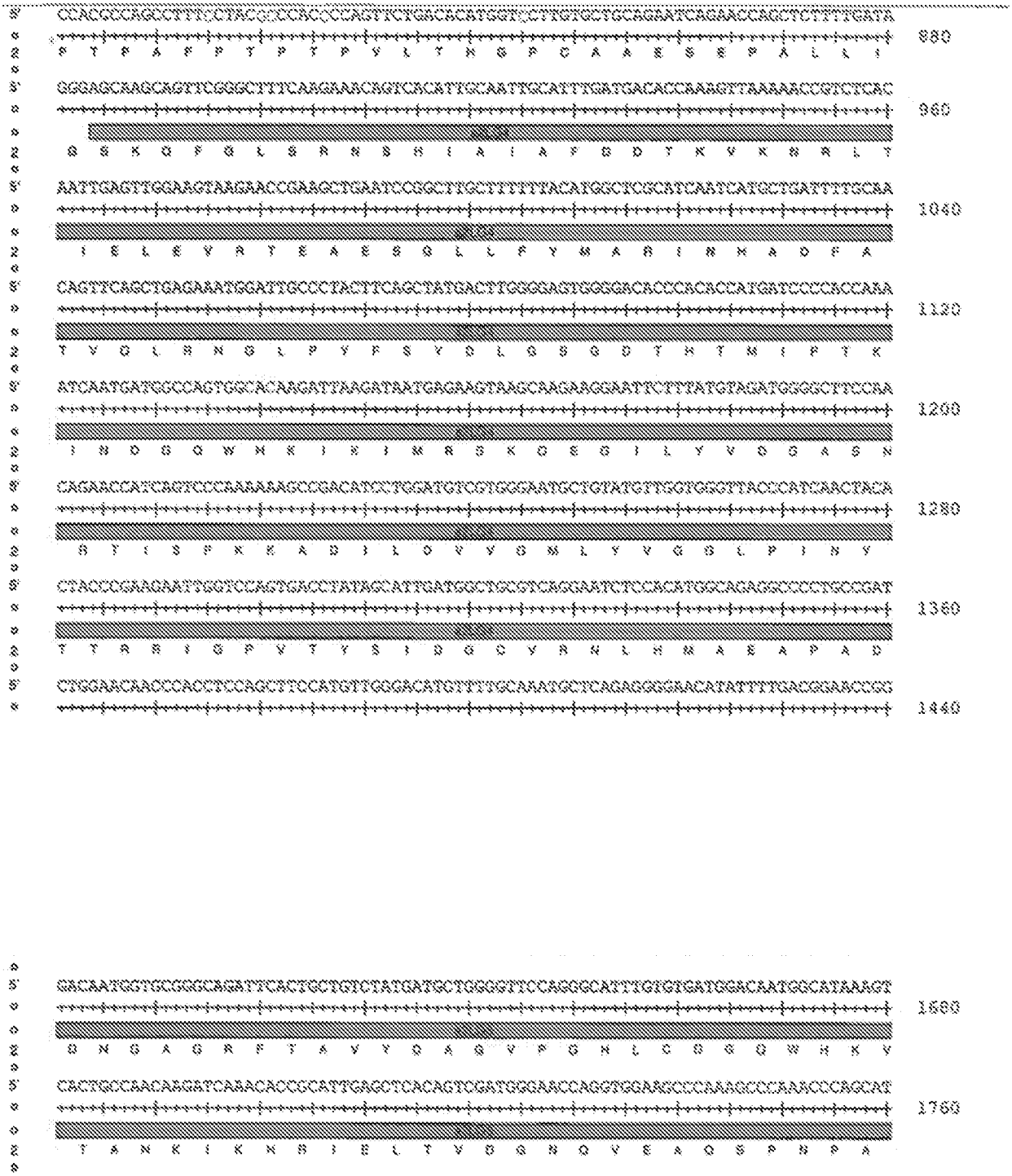


Figure 5 (Continued)

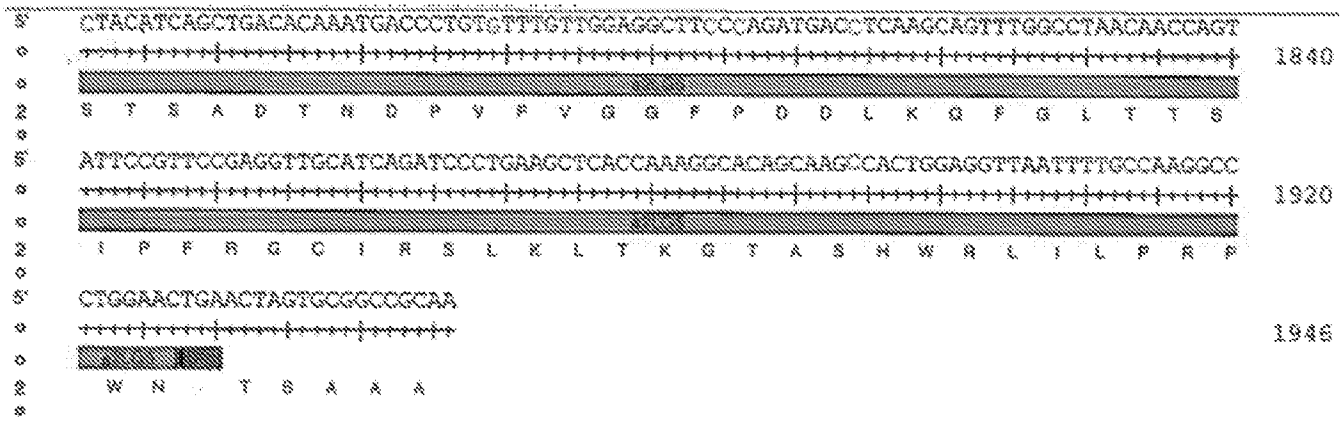


Figure 6

HB-EGF (complete soluble form)-LAMA2 G3-G5

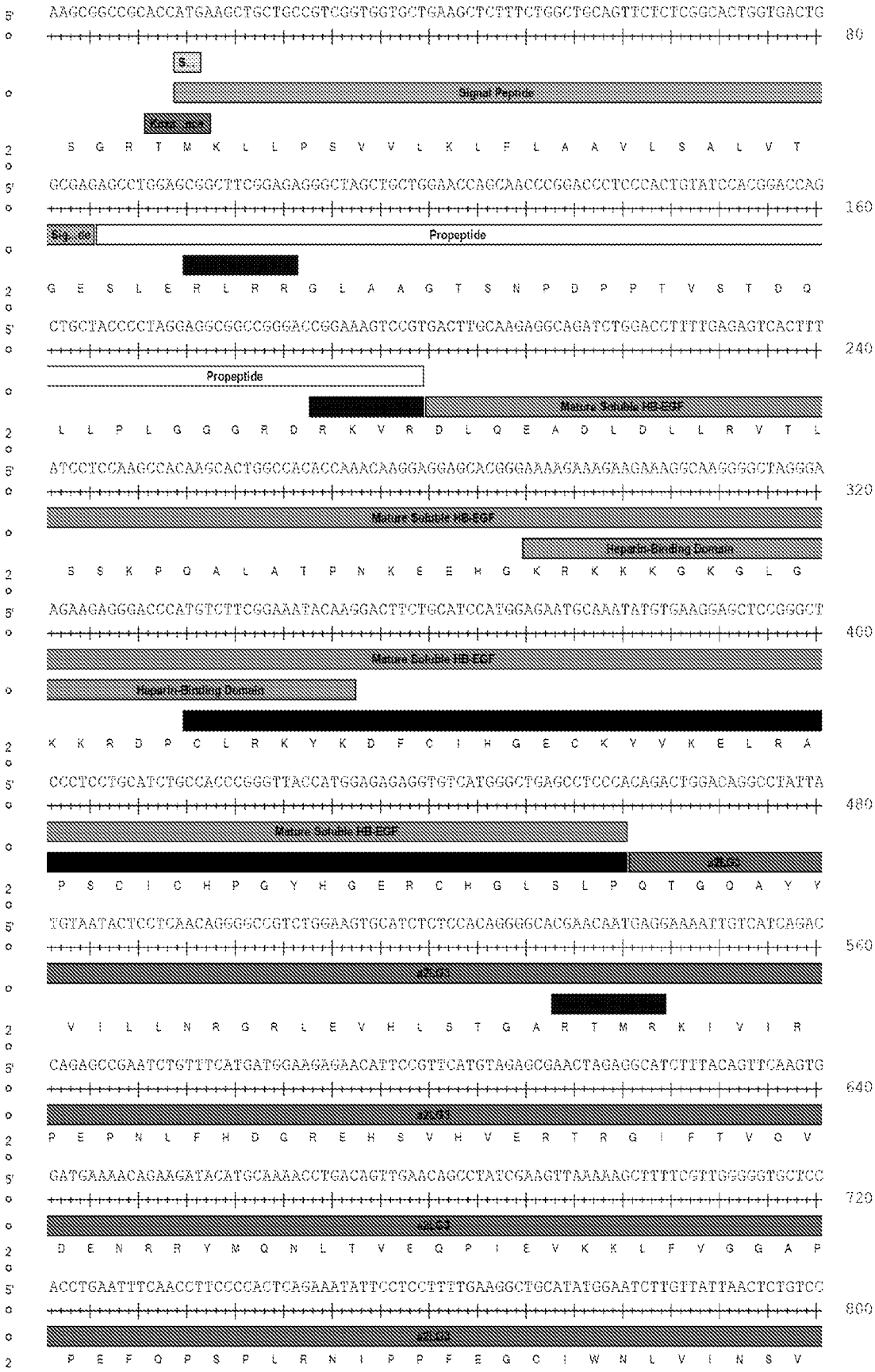


Figure 6 (Continued)

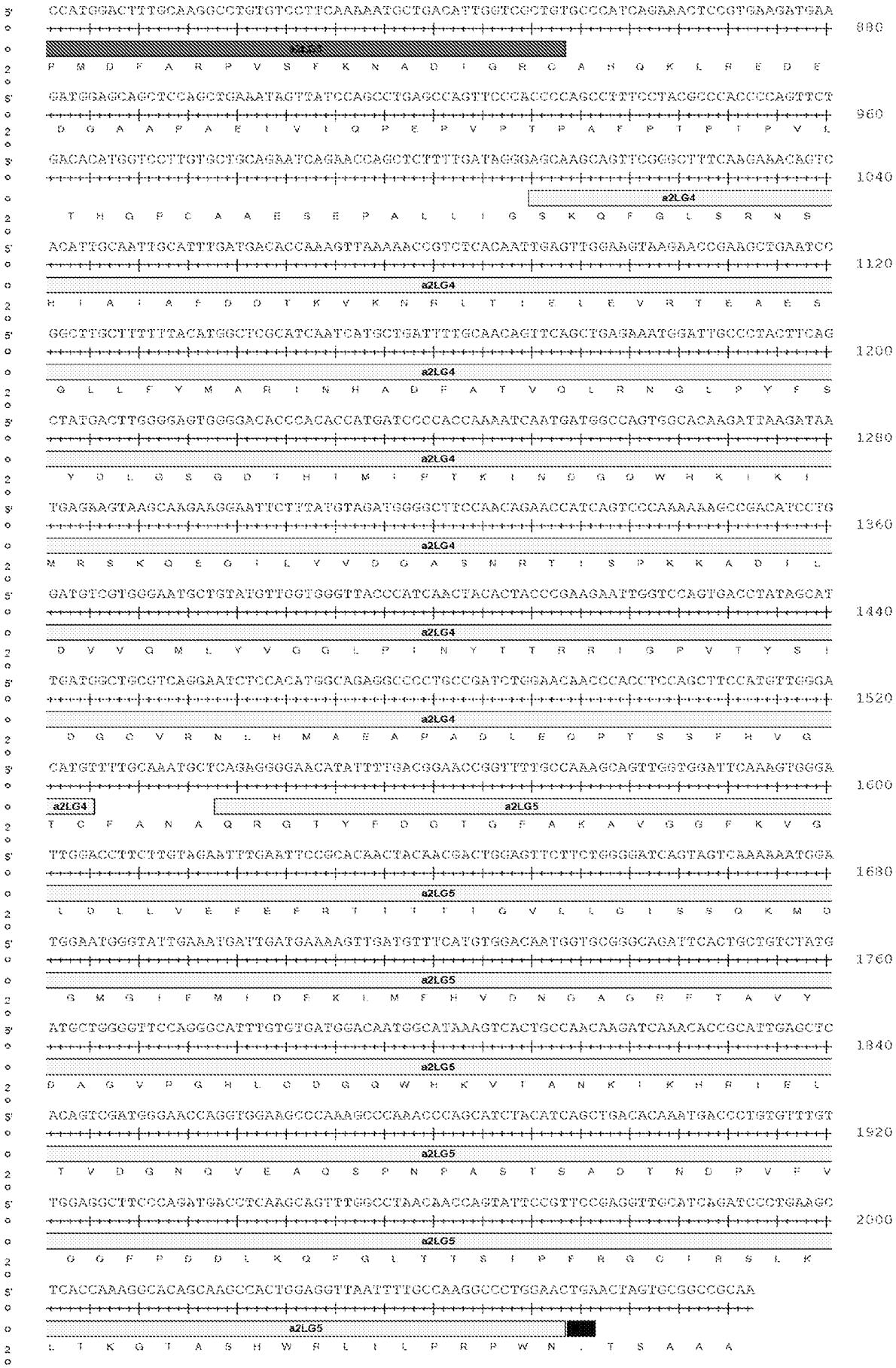


Figure 7

HB-EGF (ending at heparin binding domain)-DAG1 (native processed alpha DG gene)

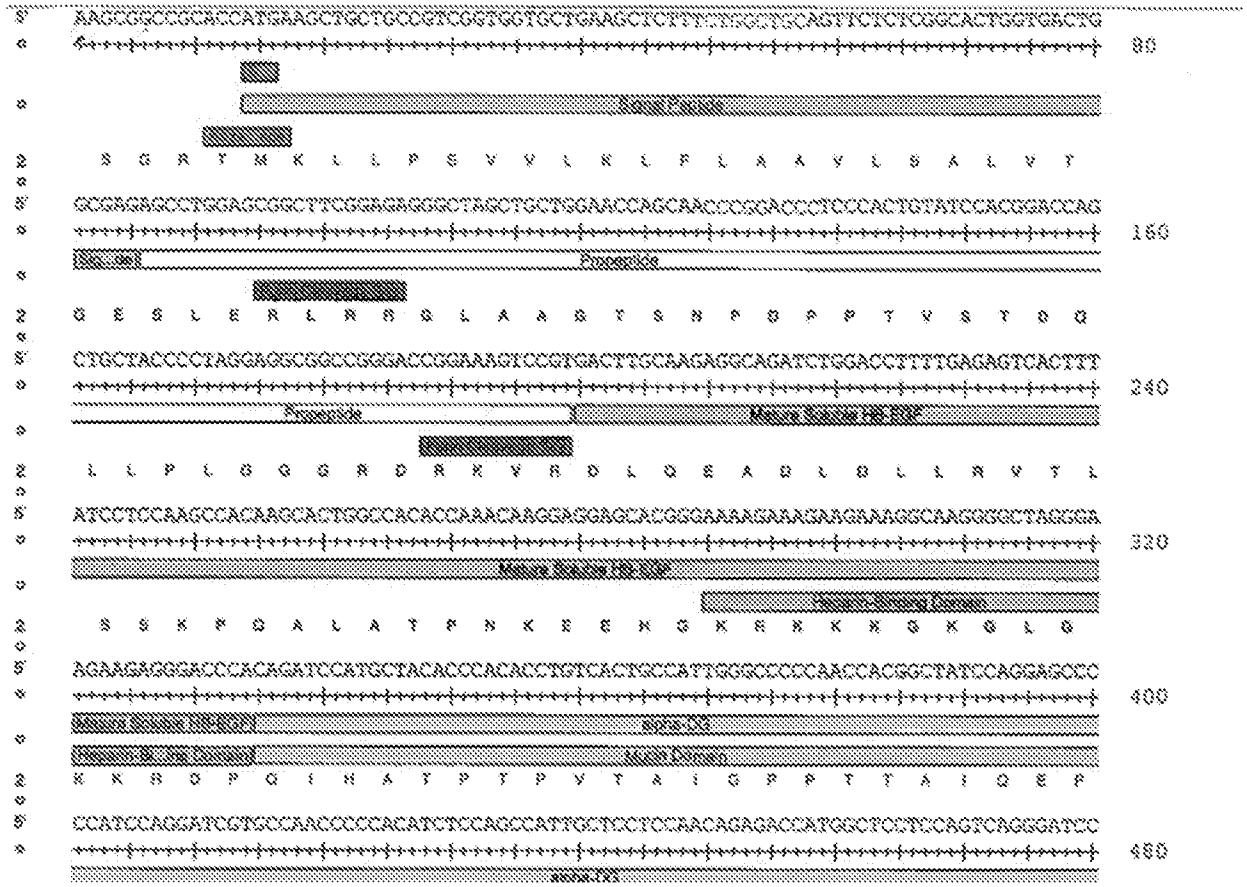


Figure 7 (Continued)

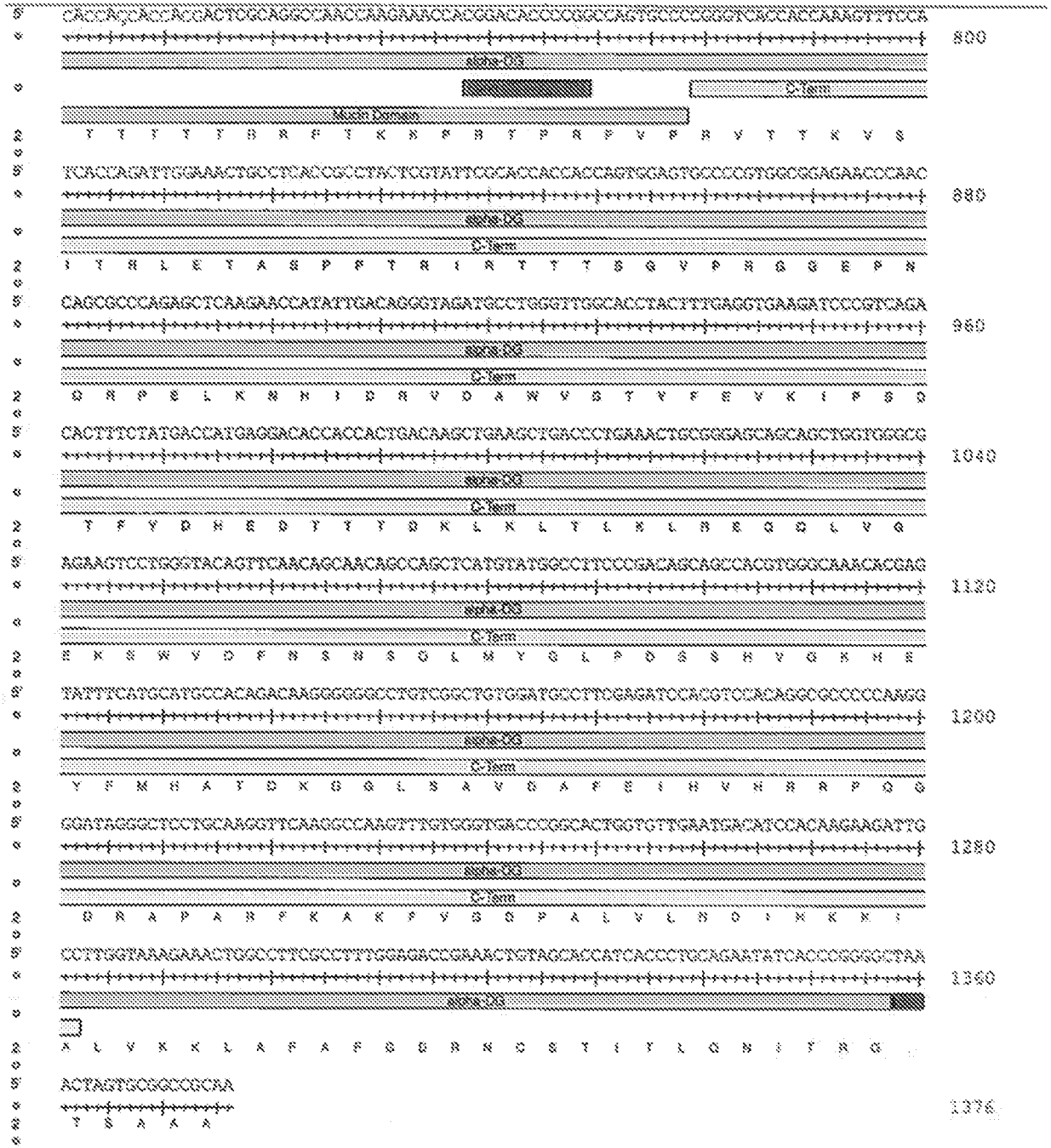


Figure 8

HB-EGF (complete soluble form)-DAG1 (native processed alpha DG gene)

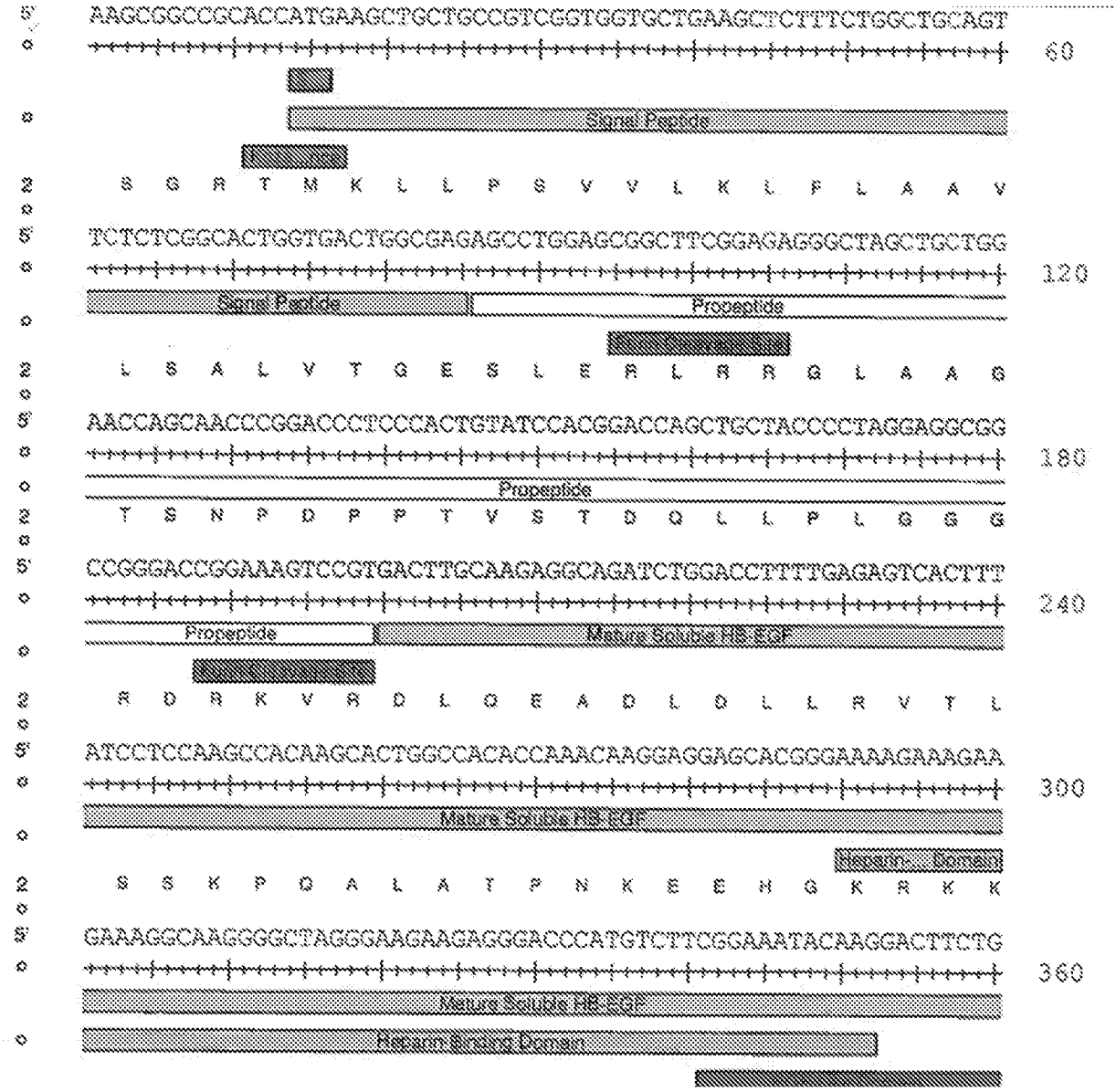
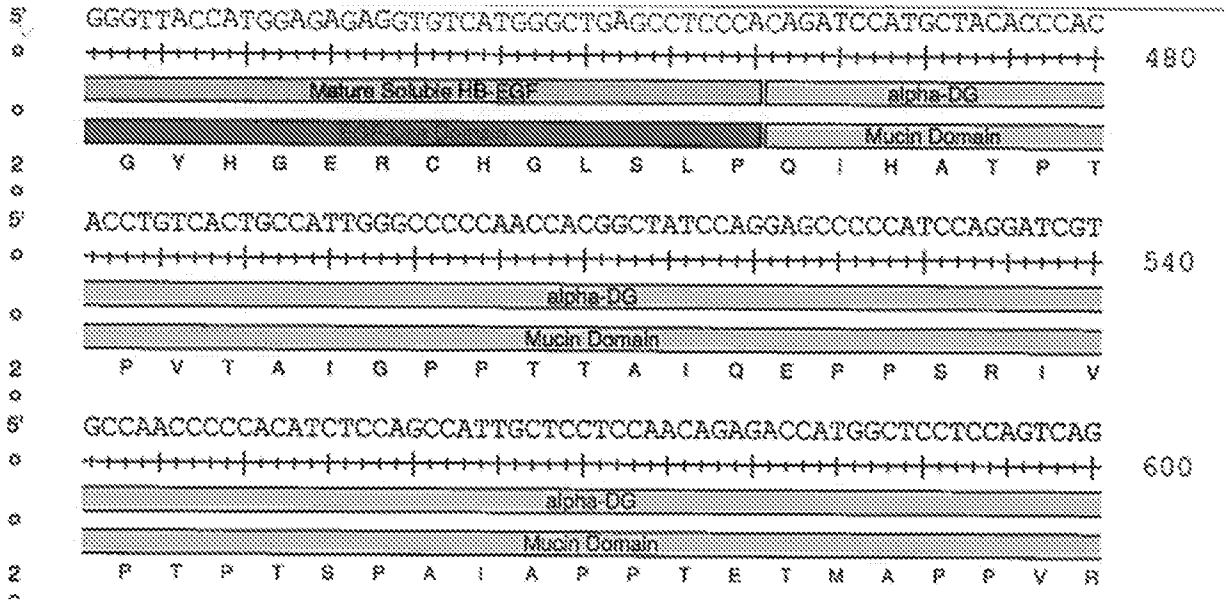


Figure 8 (Continued)



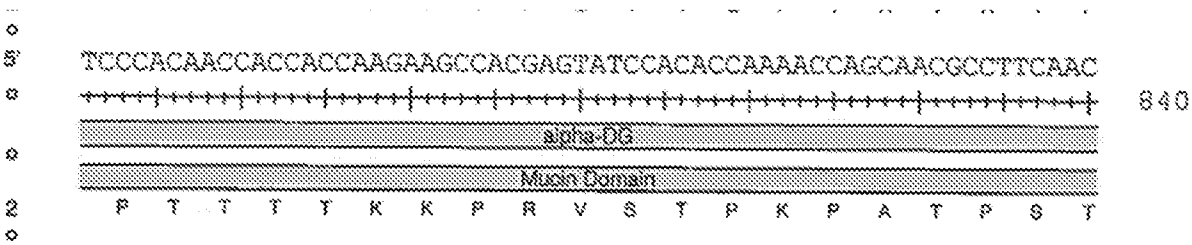


Figure 8 (Continued)

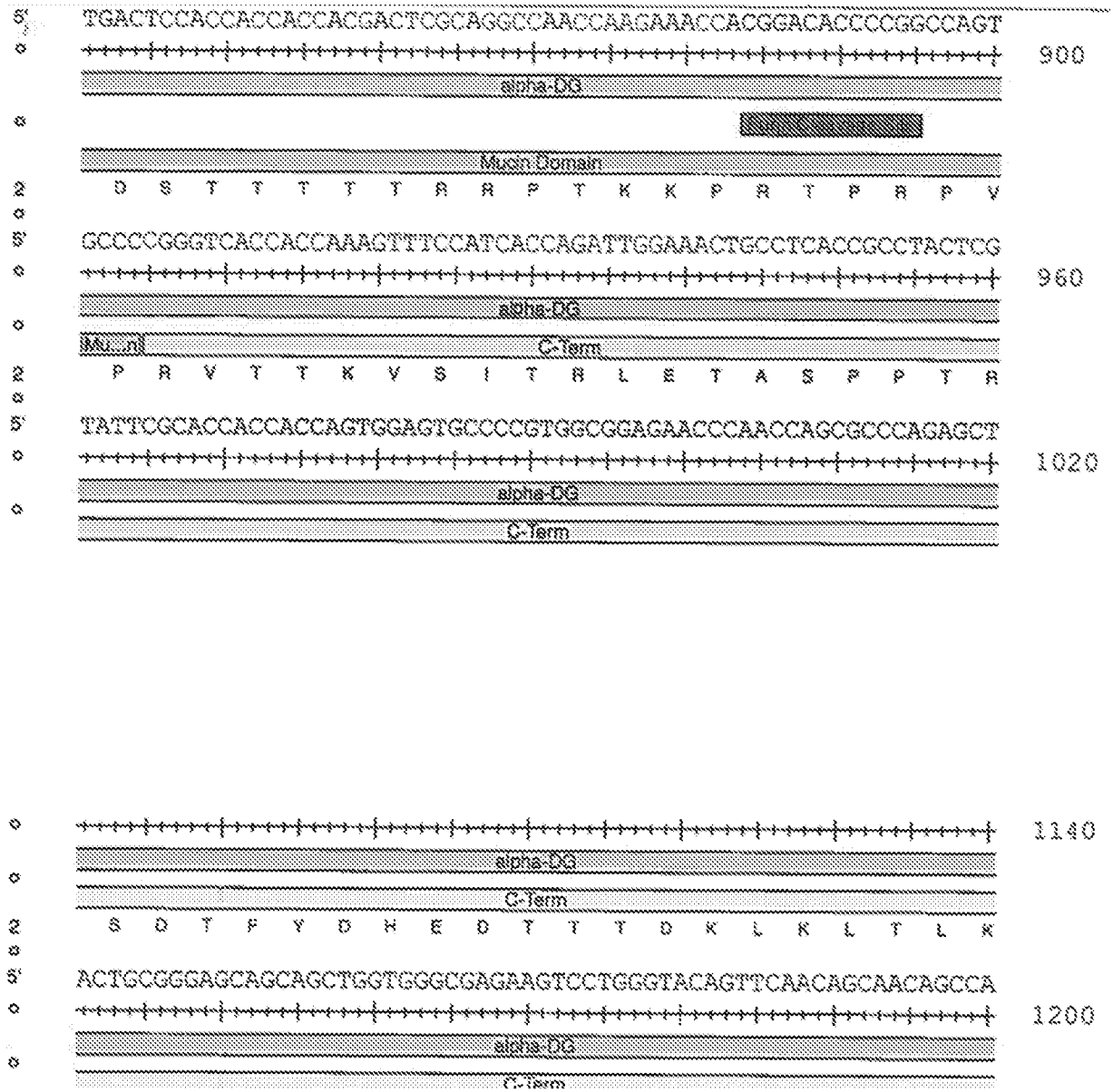


Figure 8 (Continued)

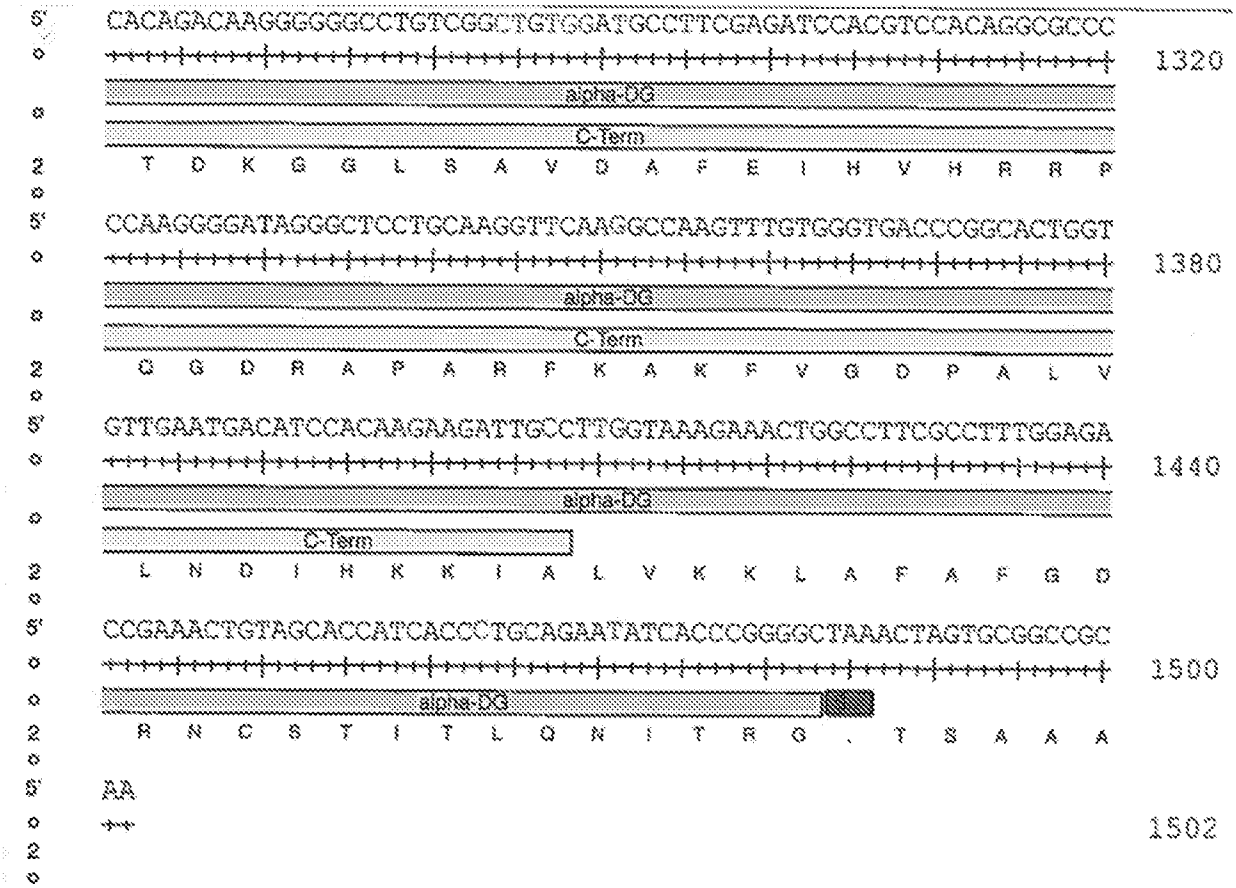


Figure 9

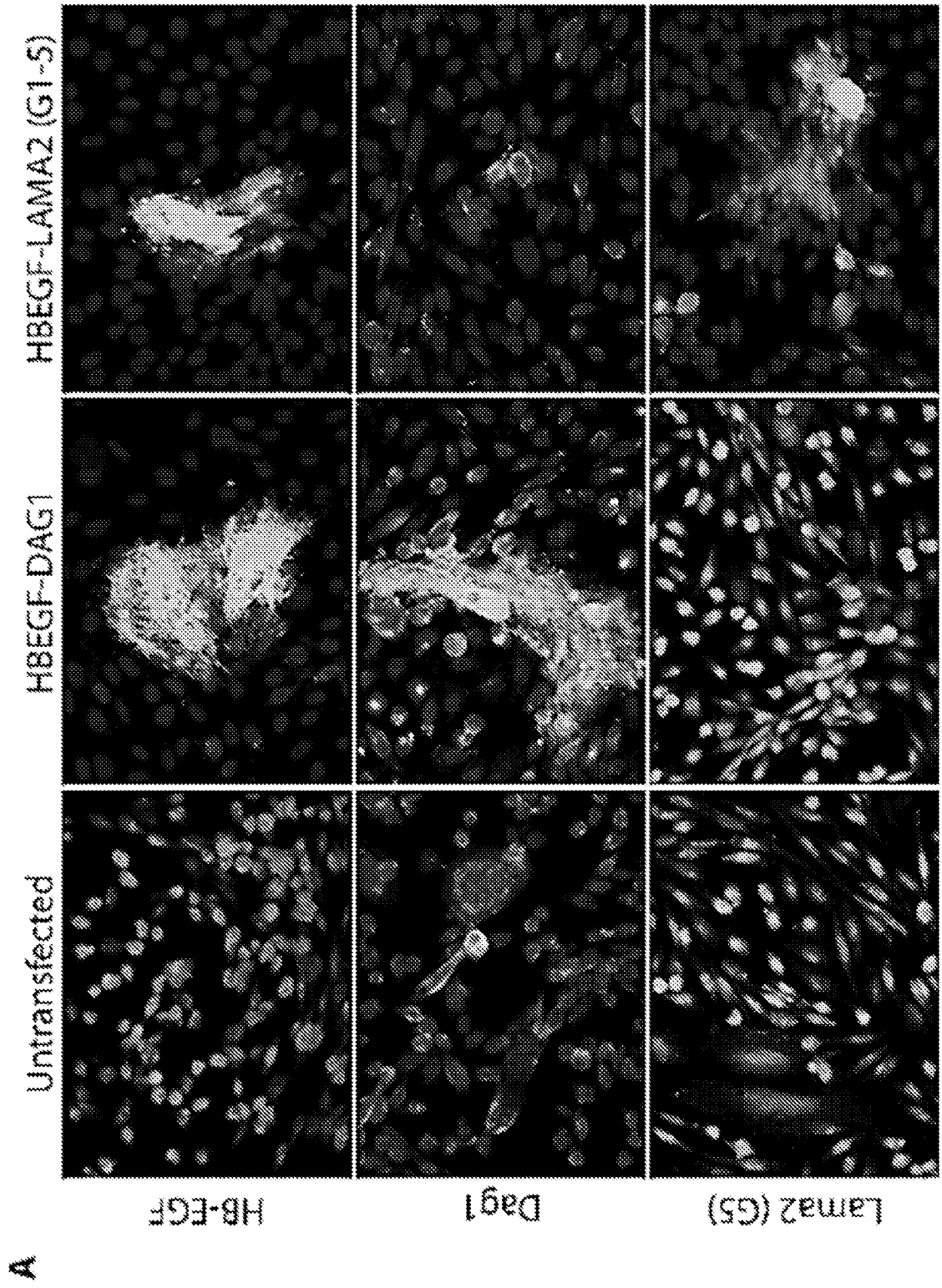

A
HB-EGF
Dag1
Lama2 (G5)
Untransfected
HBEGF-DAG1
HBEGF-LAMA2 (G1-5)

Figure 9 Continued

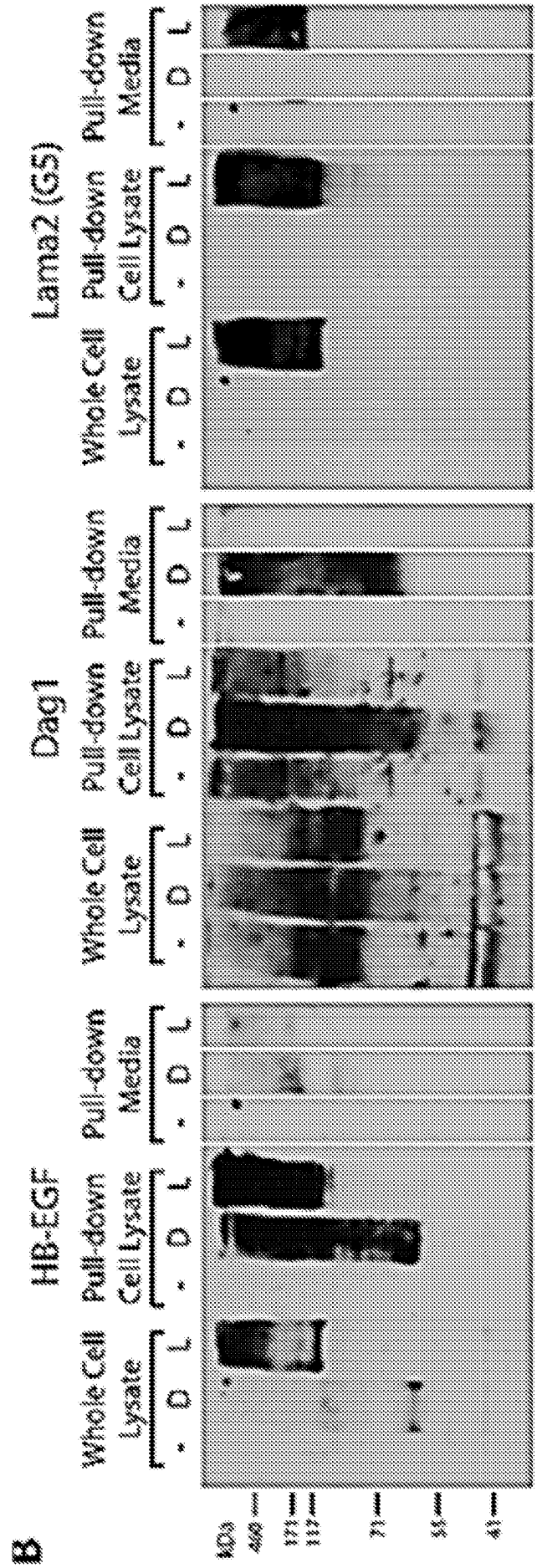


Figure 10

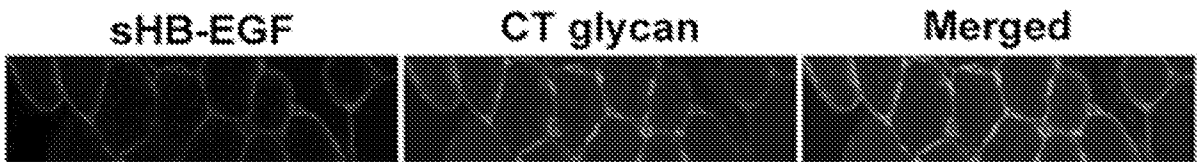


Figure 11

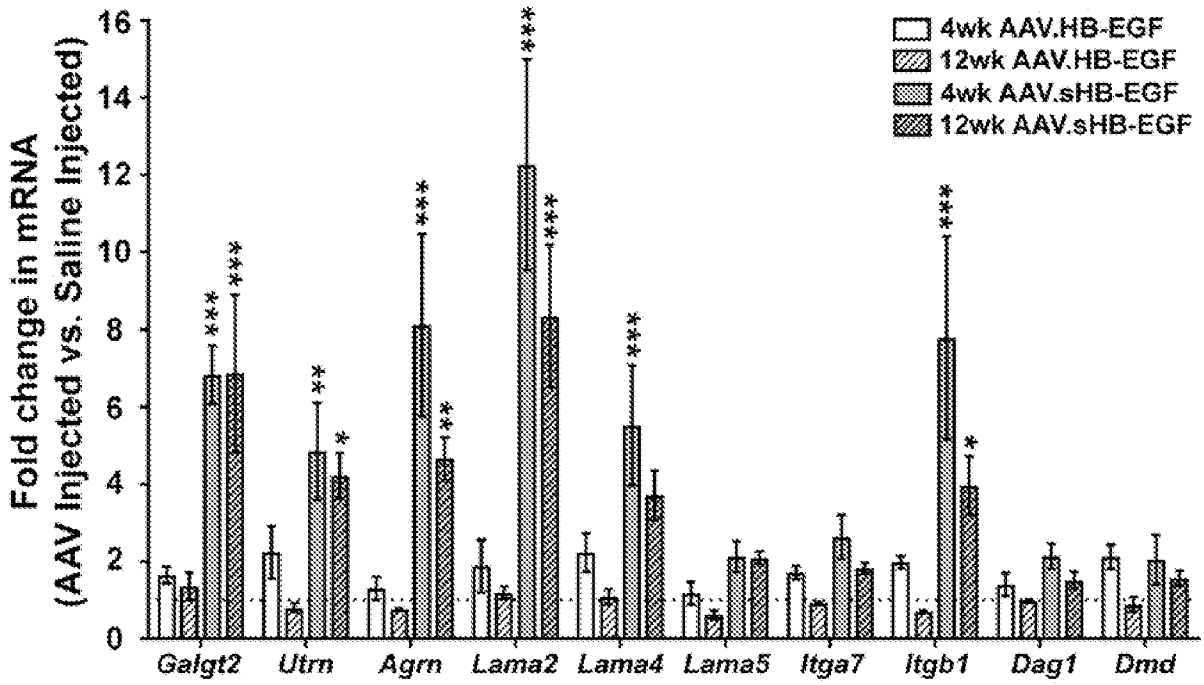

Fold change in mRNA
(AAV Injected vs. Saline Injected)
0
2
4
6
8
10
12
14
16
4wk AAV.HB-EGF
12wk AAV.HB-EGF
4wk AAV.sHB-EGF
12wk AAV.sHB-EGF
Galgt2
Utrn
Agrn
Lama2
Lama4
Lama5
Itga7
Itgb1
Dag1
Dmd

Figure 12

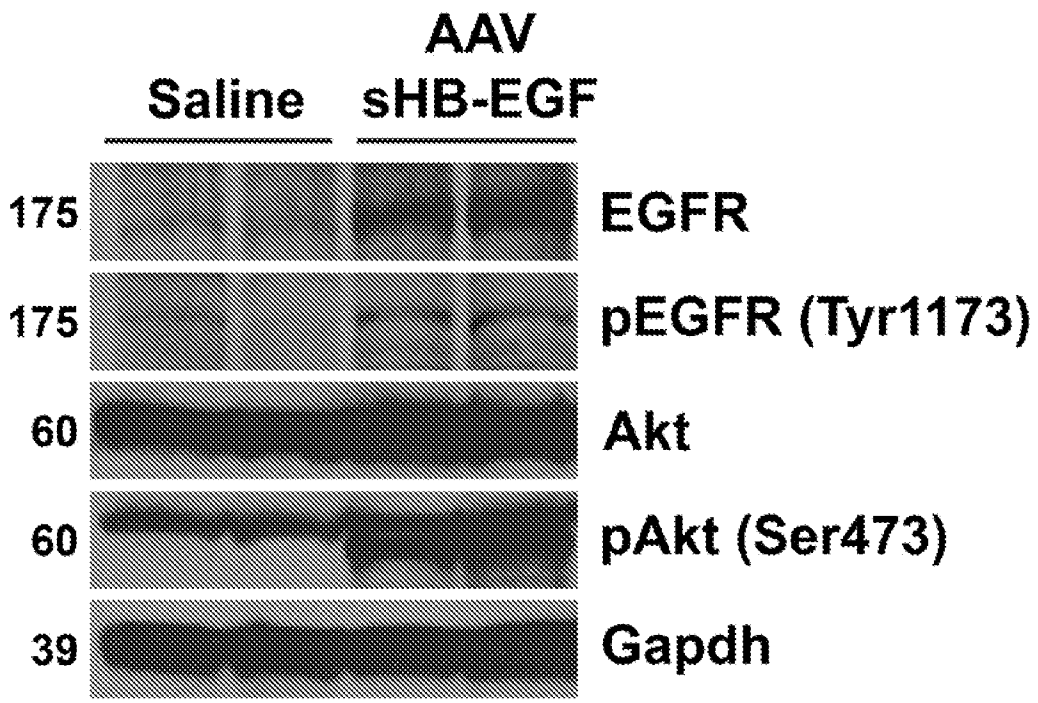

AAV
Saline
sHB-EGF
175
EGFR
175
pEGFR (Tyr1173)
60
Akt
60
pAkt (Ser473)
39
Gapdh

Figure 13

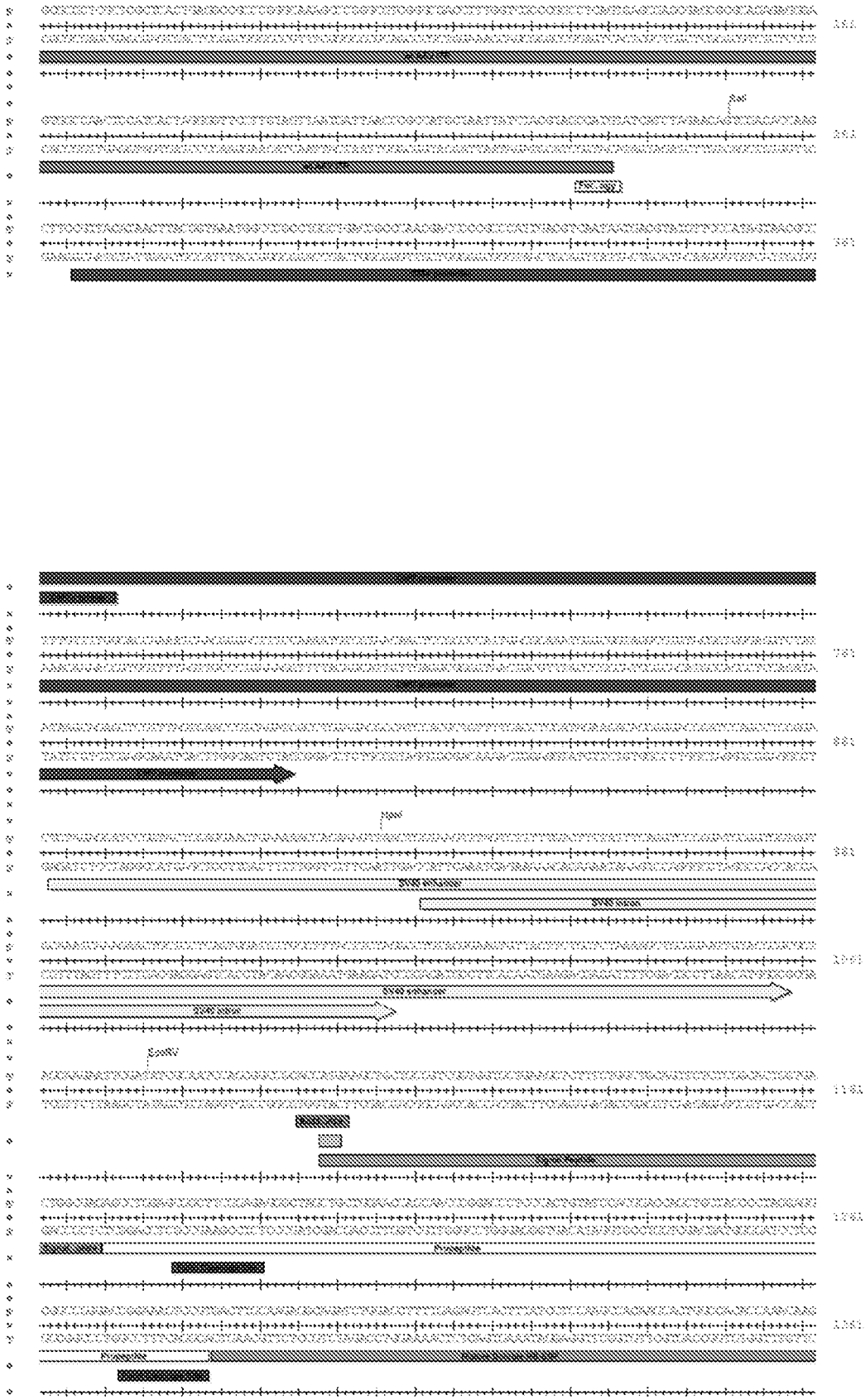


Figure 13 (Continued)

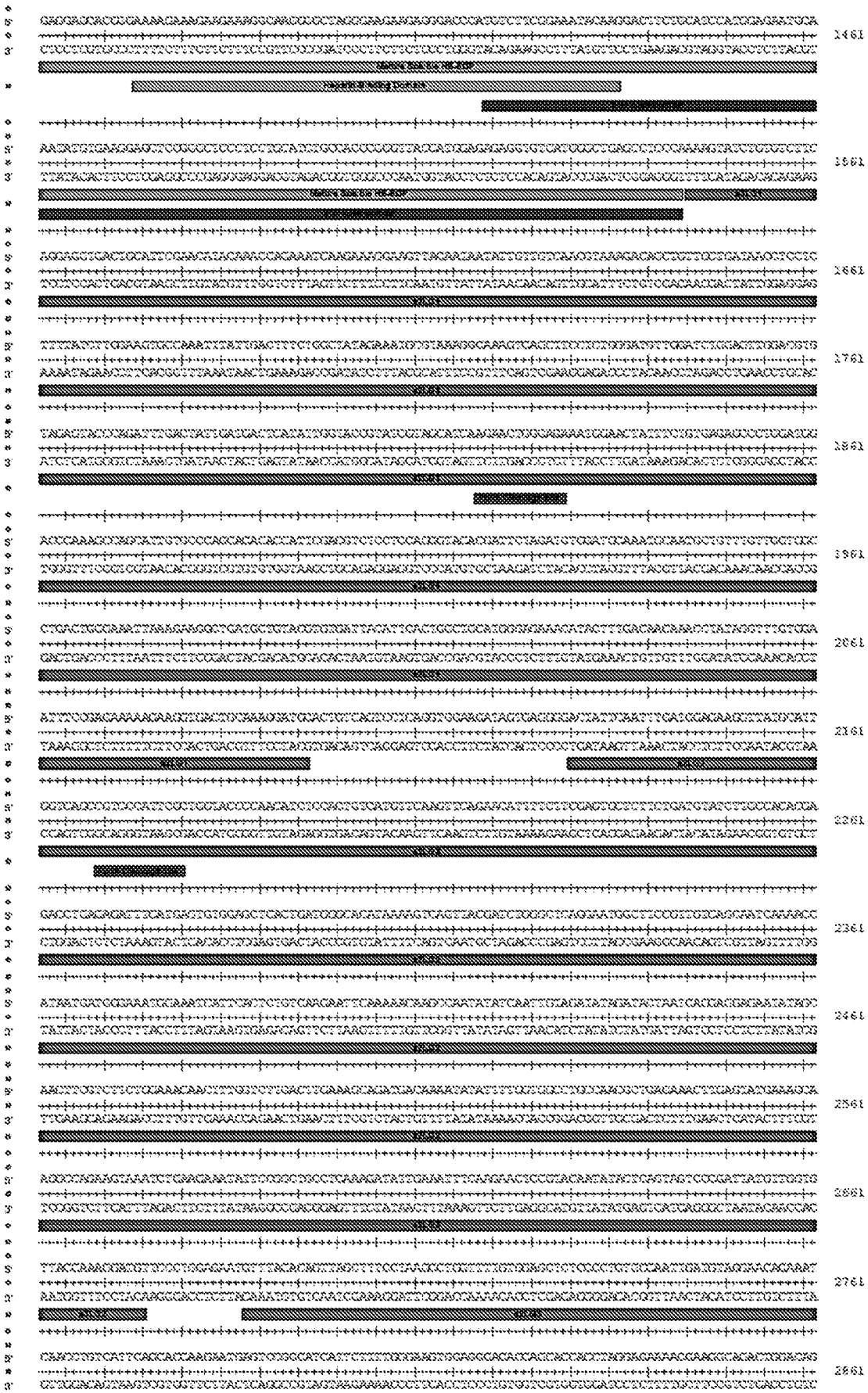


Figure 13 (Continued)

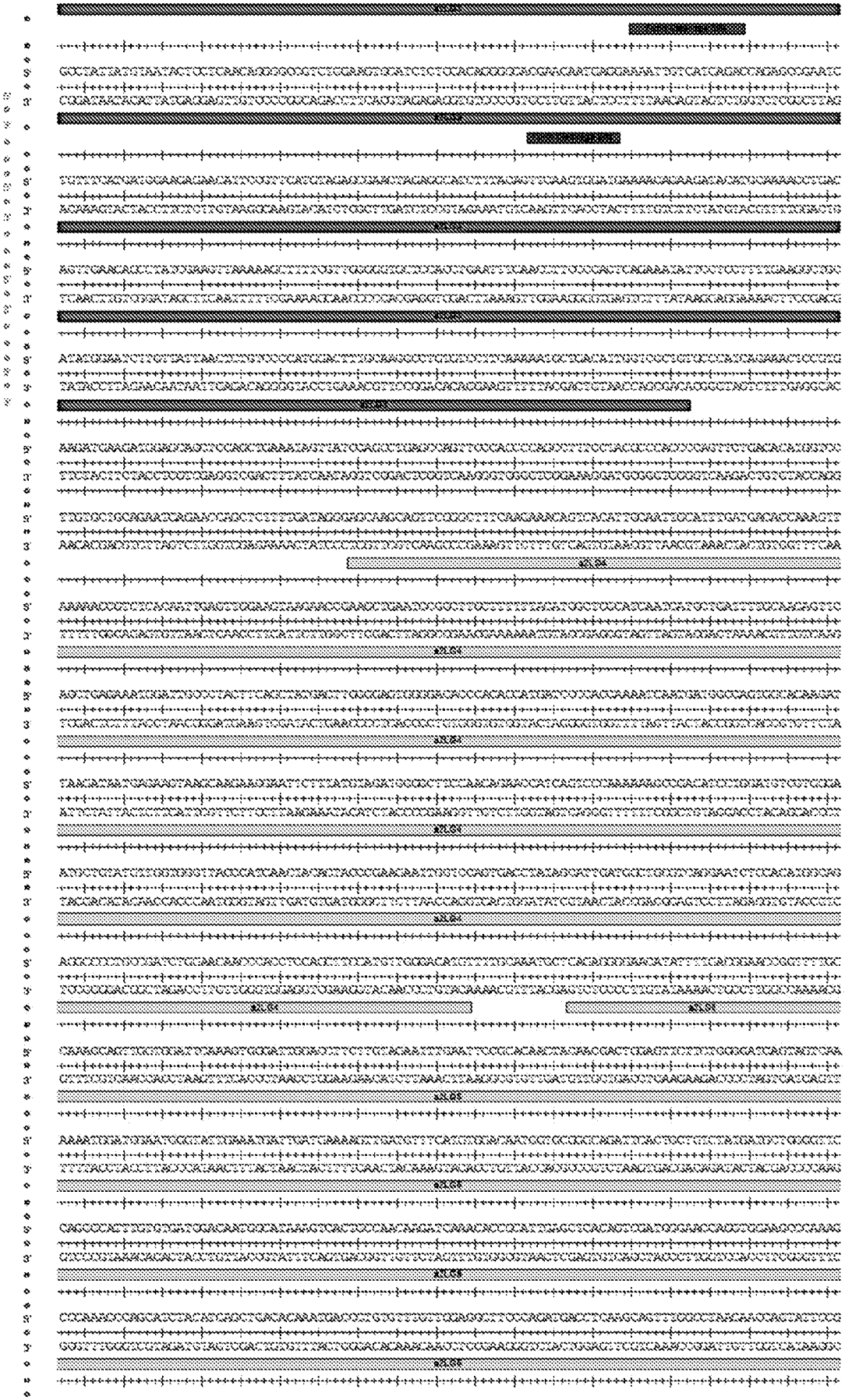


Figure 13 (Continued)

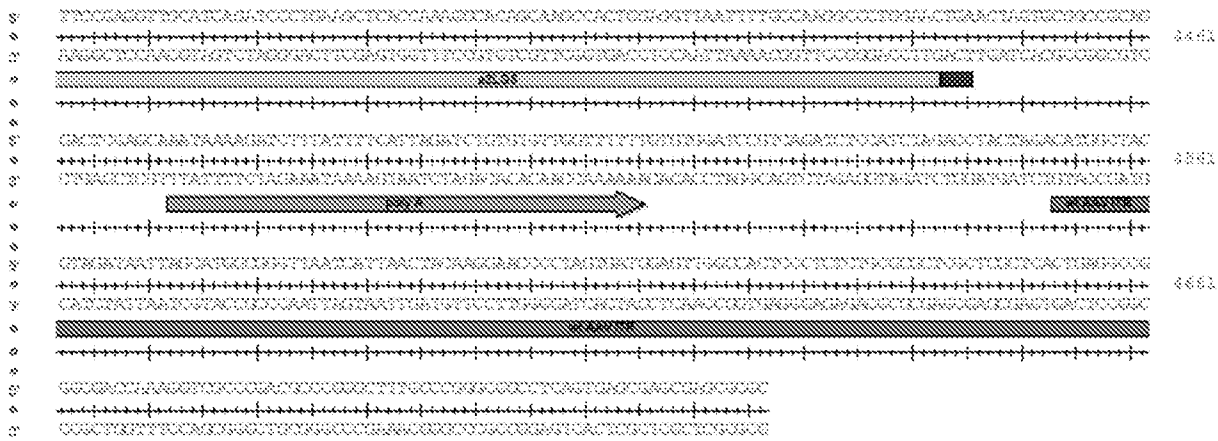


Figure 14

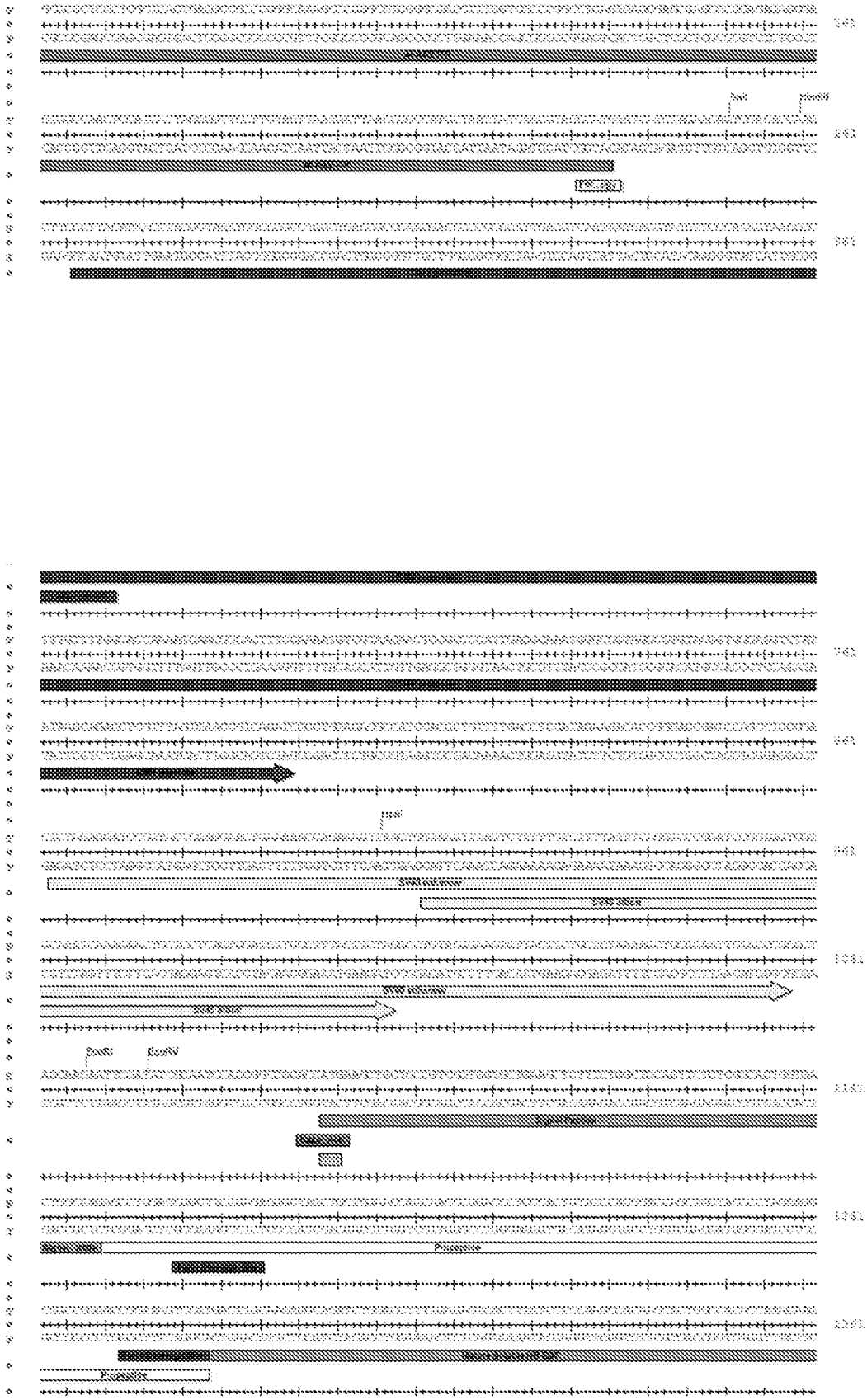


Figure 14 (Continued)

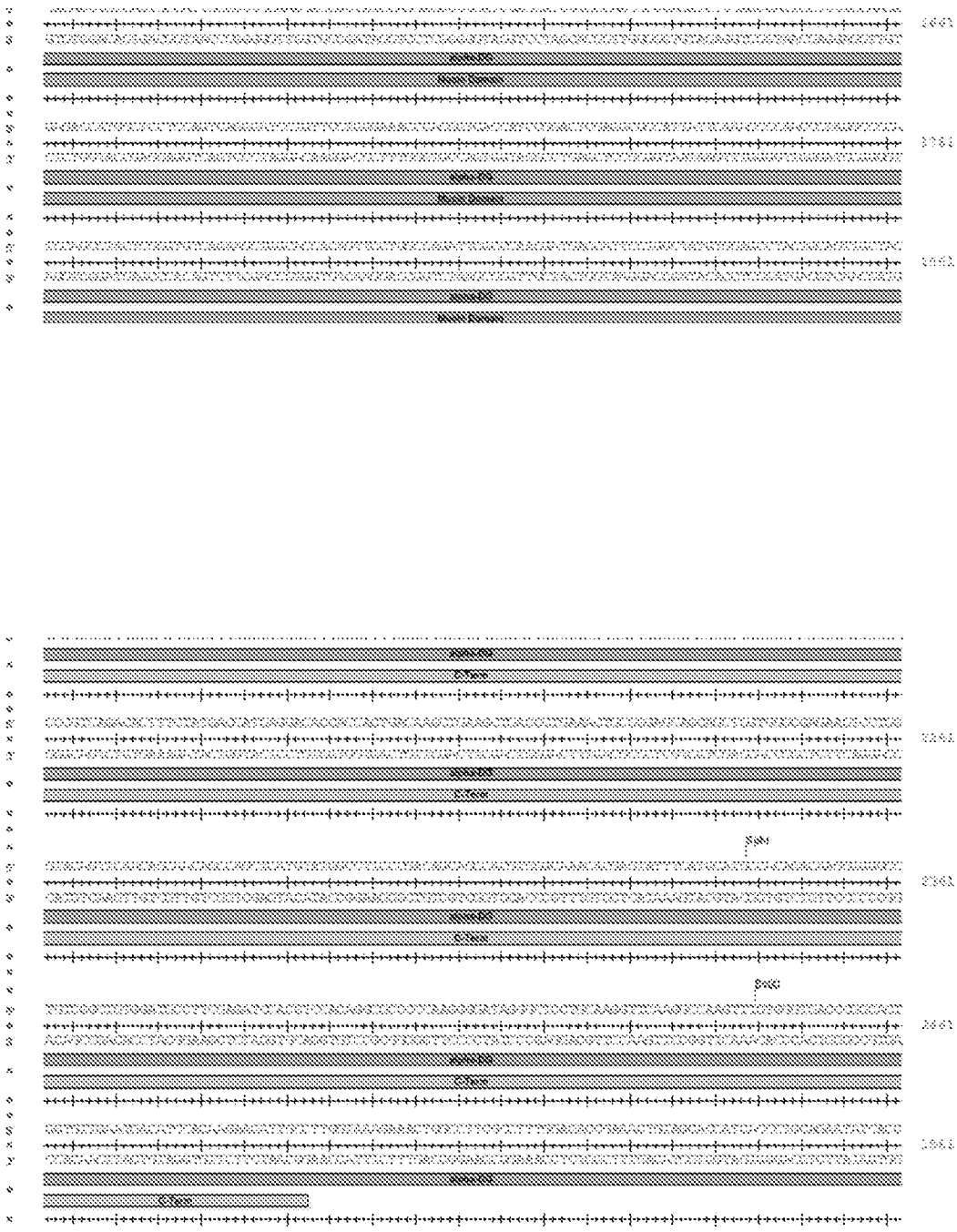


Figure 14 (Continued)

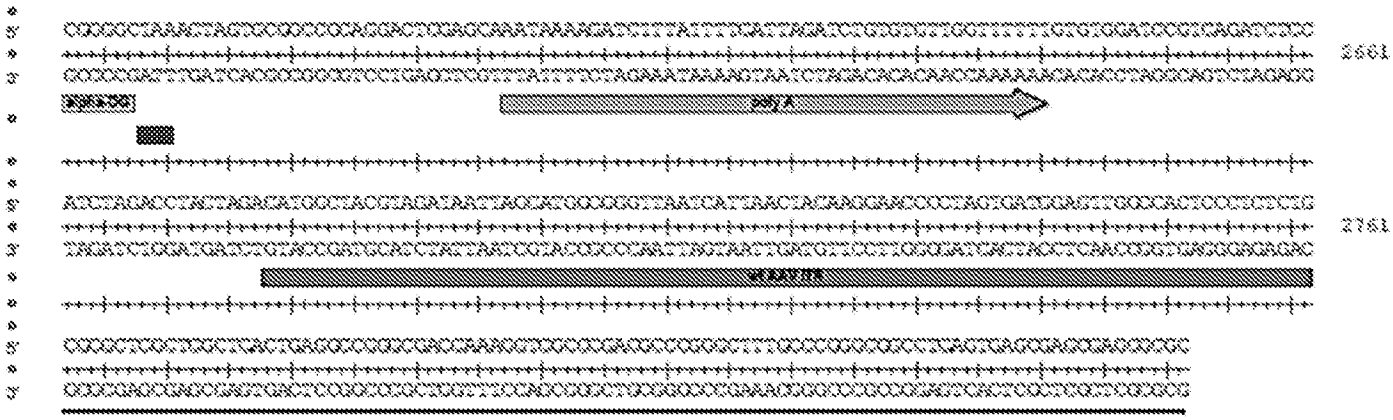


Figure 15

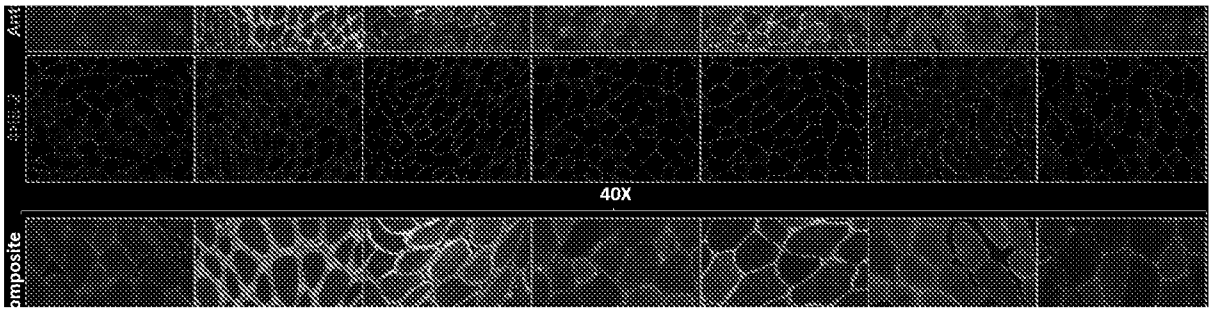


Figure 16

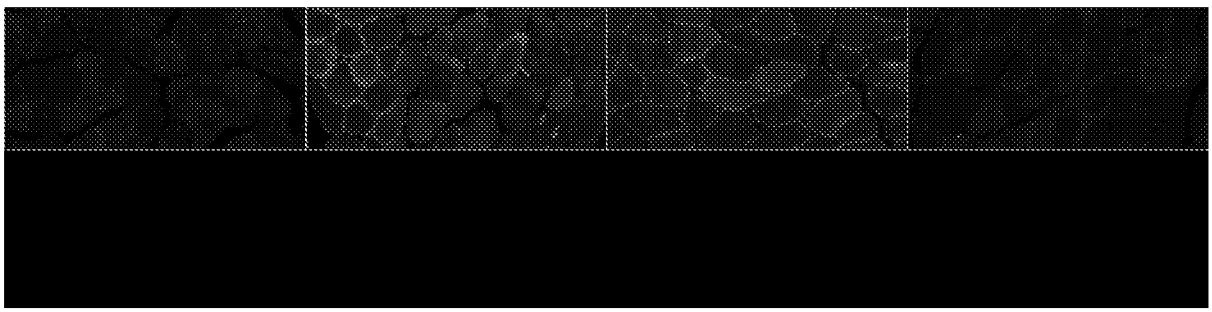


Figure 17

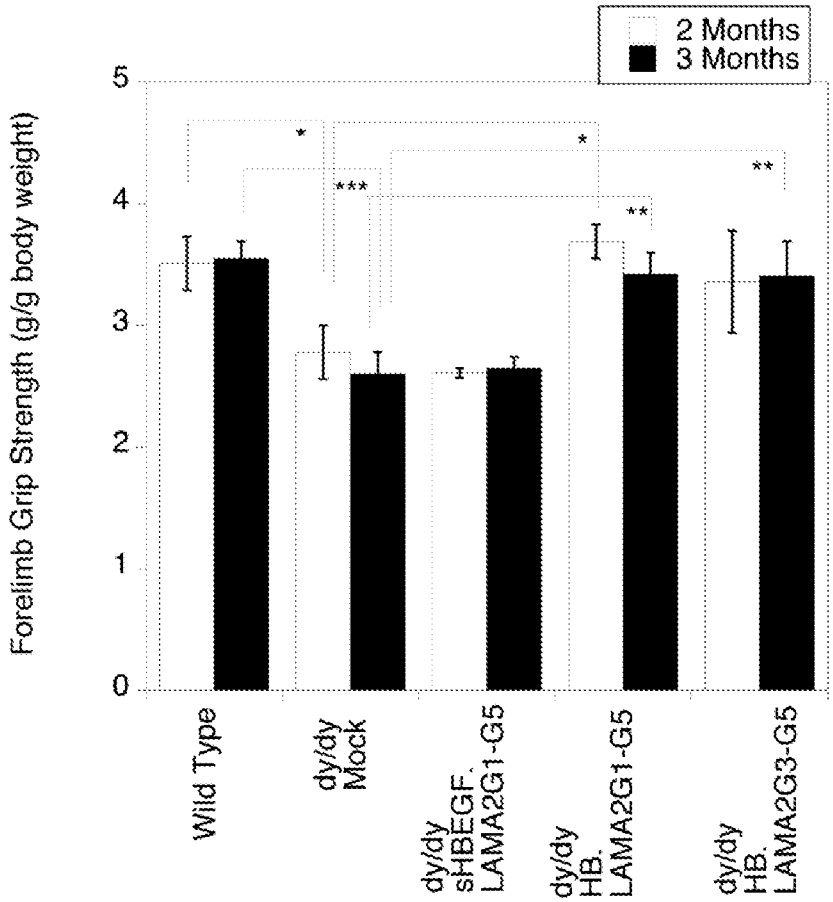

2 Months
3 Months
Forelimb Grip Strength (g/g body weight)
0
1
2
3
4
5
*
***
*
**
**
Wild Type
dy/dy Mock
dy/dy sHBEGF. LAMA2G1-G5
dy/dy HB. LAMA2G1-G5
dy/dy HB. LAMA2G3-G5

Figure 18

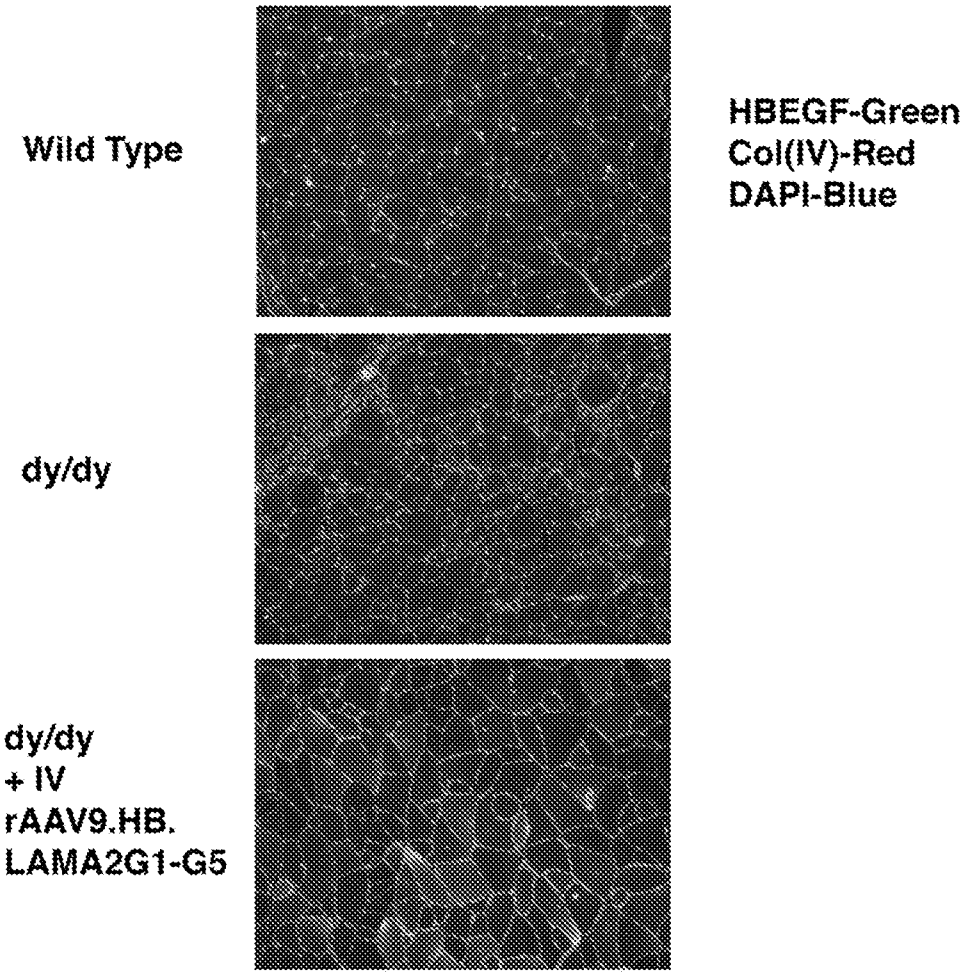

Wild Type
dy/dy
dy/dy
+ IV
rAAV9.HB.
LAMA2G1-G5
HBEGF-Green
Col(IV)-Red
DAPI-Blue

Figure 19

rAAV.CMV.HB.LAMA2(G1-G5) (SEQ ID NO: 2)

Ampicillin Resistance Gene
ITRs
CMV Promoter
SV40 Enhancer
HBEGF Signal Peptide
HBEGF Propeptide Domain
HBEGF Mature Peptide Domain
HBEGF HB Domain
LAMA2 G1-3
LAMA2 G3-5
SV40 PolyA Tail

GCTCTTCCGCTTGGTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAACCCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTACGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA
CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAA
AATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC

TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAACTTCC
AACATCCAATAAATCATACAGGCAAGGCAAAGAATTAGCAAAATTAAGCAATAA
AGCCTCAGAGCATAAAGCTAAATCGGTTGTACCAAAAACATTATGACCCTGTAAT
ACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAGGATAAAAATTTTTAGAACCCT
CATATATTTTAAATGCAATGCCTGAGTAATGTGTAGGTAAAGATTCAAACGGGTG
AGAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATTCAACCGTTCTAGC
TGATAAATTCATGCCGGAGAGGGTAGCTATTTTTGAGAGGTCTCTACAAAGGCTA
TCAGGTCATTGCCTGAGAGTCTGGAGCAAACAAGAGAATCGATGAACGGTAATC
GTAAAACTAGCATGTCAATCATATGTACCCCGGTTGATAATCAGAAAAGCCCCA
AAAACAGGAAGATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAATCACTAAATCGGAACCCTAAAGGGA
GCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAA
GGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAC
GCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTA
CTATGGTTGCTTTGACGAGCACGTATAACGTGCTTTCCTCGTTAGAATCAGAGCG
GGAGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGCCAGA
ATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCA
TCACGCAAATTAACCGTTGTCGCAATACTTCTTTGATTAGTAATAACATCACTTGC
CTGAGTAGAAGAACTCAAACTATCGGCCTTGCTGGTAATATCCAGAACAATATTA
CCGCCAGCCATTGCAACAGGAAAAACGCTCATGGAAATACCTACATTTTGACGCT
CAATCGTCTGGAACTTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG
ATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG**NNNNNNGCGCGCTCGCTCGC**
TCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCC
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA
CTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTAATTATCTACGTA
GCCATGTCTAGGGT*CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC*
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC
GCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG
TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCAT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT
*AGTGAACCGTCAGATC*GCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAA
GACACCGGGACCGATCCAGCCTCCGGAC*TCTAGAGGATCCGGTACTCGAGGAACTG*
AAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGAT
CCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGG
*CCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGC*GGCCGCAC

CATGAAGCTGCTGCCGTCGGTGGTGCTGAAGCTCTTTCTGGCTGCAGTTCTCTCG
GCACTGGTGACTGGCGAG**AGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTG**
GAACCAGCAACCCGGACCCTCCCACTGTATCCACGGACCAGCTGCTACCCC
TAGGAGGCGGCCGGGACCGGAAAGTCCGTGACTTGCAAGAGGCAGATCTGG
ACCTTTTGAGAGTCACTTTATCCTCCAAGCCACAAGCACTGGCCACACCAAA
CAAGGAGGAGCACGGGAAAAGAAAGAAGAAAGGCAAGGGGCTAGGGAAGAA
GAGGGACCCAAAAGTATCTGTGTCTTCAGGAGGTGACTGCATTCGAACATACAA
ACCAGAAATCAAGAAAGGAAGTTACAATAATATTGTTGTCAACGTAAAGACAGC
TGTTGCTGATAACCTCCTCTTTTATCTTGGAAGTGCCAAATTTATTGACTTTCTGG
CTATAGAAATGCGTAAAGGCAAAGTCAGCTTCCTCTGGGATGTTGGATCTGGAGT
TGGACGTGTAGAGTACCCAGATTTGACTATTGATGACTCATATTGGTACCGTATC
GTAGCATCAAGAACTGGGAGAAATGGAACTATTTCTGTGAGAGCCCTGGATGGA
CCCAAAGCCAGCATTGTGCCCAGCACACACCATTCGACGTCTCCTCCAGGGTACA
CGATTCTAGATGTGGATGCAAATGCAATGCTGTTTGTTGGTGGCCTGACTGGGAA
ATTAAAGAAGGCTGATGCTGTACGTGTGATTACATTCACTGGCTGCATGGGAGAA
ACATACTTTGACAACAAACCTATAGGTTTGTGGAATTTCCGAGAAAAAGAAGGT
GACTGCAAAGGATGCACTGTCAGTCCTCAGGTGGAAGATAGTGAGGGGACTATT
CAATTTGATGGAGAAGGTTATGCATTGGTCAGCCGTCCCATTCGCTGGTACCCCA
ACATCTCCACTGTCATGTTCAAGTTCAGAACATTTTCTTCGAGTGCTCTTCTGATG
TATCTTGCCACACGAGACCTGAGAGATTTCATGAGTGTGGAGCTCACTGATGGGC
ACATAAAAGTCAGTTACGATCTGGGCTCAGGAATGGCTTCCGTTGTCAGCAATCA
AAACCATAATGATGGGAAATGGAAATCATTCACTCTGTCAAGAATTCAAAAACA
AGCCAATATATCAATTGTAGATATAGATACTAATCAGGAGGAGAATATAGCAAC
TTCGTCTTCTGGAAACAACTTTGGTCTTGACTTGAAAGCAGATGACAAAATATAT
TTTGGTGGCCTGCCAACGCTGAGAAACTTGAGTATGAAAGCAAGGCCAGAAGTA
AATCTGAAGAAATATTCCGGCTGCCTCAAAGATATTGAAATTTCAAGAACTCCGT
ACAATATACTCAGTAGTCCCGATTATGTTGGTGTTACCAAAGGATGTTCCCTGGA
GAATGTTTACACAGTTAGCTTTCCTAAGCCTGGTTTTGTGGAGCTCTCCCCTGTGC
CAATTGATGTAGGAACAGAAATCAACCTGTCATTCAGCACCAAGAATGAGTCCG
GCATCATTCTTTTGGGAAGTGGAGGGACACCAGCACCACCTAGGAGAAAACGAA
GG**CAGACTGGACAGGCCTATTATGTAATACTCCTCAACAGGGGCCGTCTGG**
AAGTGCATCTCTCCACAGGGGCACGAACAATGAGGAAAATTGTCATCAGAC
CAGAGCCGAATCTGTTTCATGATGGAAGAGAACATTCCGTTCATGTAGAGCG
AACTAGAGGCATCTTTACAGTTCAAGTGGATGAAAACAGAAGATACATGCAA
AACCTGACAGTTGAACAGCCTATCGAAGTTAAAAAGCTTTTCGTTGGGGGTG
CTCCACCTGAATTTCAACCTTCCCCACTCAGAAATATTCCTCCTTTTGAAGG
CTGCATATGGAATCTTGTTATTAACTCTGTCCCCATGGACTTTGCAAGGCCT
GTGTCCTTCAAAAATGCTGACATTGGTCGCTGTGCCCATCAGAAACTCCGTG
AAGATGAAGATGGAGCAGCTCCAGCTGAAATAGTTATCCAGCCTGAGCCAG
TTCCCACCCCAGCCTTTCCTACGCCCACCCCAGTTCTGACACATGGTCCTTG
TGCTGCAGAATCAGAACCAGCTCTTTTGATAGGGAGCAAGCAGTTCGGGCT
TTCAAGAAACAGTCACATTGCAATTGCATTTGATGACACCAAAGTTAAAAAC
CGTCTCACAATTGAGTTGGAAGTAAGAACCGAAGCTGAATCCGGCTTGCTTT
TTTACATGGCTCGCATCAATCATGCTGATTTTGCAACAGTTCAGCTGAGAAA
TGGATTGCCCTACTTCAGCTATGACTTGGGGAGTGGGGACACCCACACCAT
GATCCCCACCAAAATCAATGATGGCCAGTGGCACAAGATTAAGATAATGAG
AAGTAAGCAAGAAGGAATTCTTTATGTAGATGGGGCTTCCAACAGAACCATC
AGTCCCAAAAAAGCCGACATCCTGGATGTCGTGGGAATGCTGTATGTTGGT
GGGTTACCCATCAACTACACTACCCGAAGAATTGGTCCAGTGACCTATAGCA
TTGATGGCTGCGTCAGGAATCTCCACATGGCAGAGGCCCCTGCCGATCTGG

**AACAACCCACCTCCAGCTTCCATGTTGGGACATGTTTTGCAAATGCTCAGAG
GGGAACATATTTTGACGGAACCGGTTTTGCCAAAGCAGTTGGTGGATTCAAA
GTGGGATTGGACCTTCTTGTAGAATTTGAATTCCGCACAACTACAACGACTG
GAGTTCTTCTGGGGATCAGTAGTCAAAAAATGGATGGAATGGGTATTGAAAT
GATTGATGAAAAGTTGATGTTTCATGTGGACAATGGTGCGGGCAGATTCACT
GCTGTCTATGATGCTGGGGTTCCAGGGCATTTGTGTGATGGACAATGGCAT
AAAGTCACTGCCAACAAGATCAAACACCGCATTGAGCTCACAGTCGATGGG
AACCAGGTGGAAGCCCAAAGCCCAAACCCAGCATCTACATCAGCTGACACA
AATGACCCTGTGTTTGTTGGAGGCTTCCCAGATGACCTCAAGCAGTTTGGCC
TAACAACCAGTATTCCGTTCCGAGGTTGCATCAGATCCCTGAAGCTCACCAA
AGGCACAGCAAGCCACTGGAGGTTAATTTTGCCAAGGCCCTGGAACTGA**AC
TAGTGCGGCCGCGGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAA
ACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTA
TTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTG
CATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTCGGATCCTC
TAGAGTCGACCA**CATGGCTACGTAGATAATTAGCATGGCGGGTTAATCATTA
ACTACAAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTG
CCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCNNNNNN**CAGCTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG

Figure 20

pAAV.CMV.HBEGF.LAMA2(G1-G5) (SEQ ID NO: 4)

Main Characteristics

Ampicillin Resistance Gene
ITRs
CMV Promoter
SV40 Enhancer
HBEGF Signal Peptide
HBEGF Propeptide Domain
HBEGF Mature Peptide Domain
HBEGF HB Domain
HBEGF EGF-Like Domain
LAMA2 G1-3
LAMA2 G3-5
SV40 PolyA Tail

GCTCTTCCGCTTGGTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAACCCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTACGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA
CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA

TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAA
AATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAACTTCC
AACATCCAATAAATCATACAGGCAAGGCAAAGAATTAGCAAAATTAAGCAATAA
AGCCTCAGAGCATAAAGCTAAATCGGTTGTACCAAAAACATTATGACCCTGTAAT
ACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAGGATAAAAATTTTTAGAACCCT
CATATATTTTAAATGCAATGCCTGAGTAATGTGTAGGTAAAGATTCAAACGGGTG
AGAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATTCAACCGTTCTAGC
TGATAAATTCATGCCGGAGAGGGTAGCTATTTTTGAGAGGTCTCTACAAAGGCTA
TCAGGTCATTGCCTGAGAGTCTGGAGCAAACAAGAGAATCGATGAACGGTAATC
GTAAAACTAGCATGTCAATCATATGTACCCCGGTTGATAATCAGAAAAGCCCCA
AAAACAGGAAGATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAATCACTAAATCGGAACCCTAAAGGGA
GCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAA
GGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAC
GCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTA
CTATGGTTGCTTTGACGAGCACGTATAACGTGCTTTCCTCGTTAGAATCAGAGCG
GGAGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGCCAGA
ATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCA
TCACGCAAATTAACCGTTGTCGCAATACTTCTTTGATTAGTAATAACATCACTTGC
CTGAGTAGAAGAACTCAAACTATCGGCCTTGCTGGTAATATCCAGAACAATATTA
CCGCCAGCCATTGCAACAGGAAAAACGCTCATGGAAATACCTACATTTTGACGCT
CAATCGTCTGGAACTTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG
ATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG**NNNNNNGCGCGCTCGCTCGC**
TCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCC
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA
CTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTAATTATCTACGTA
GCCATGTCTAGGGT*CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC*
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC
GCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG
TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCAT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT
*AGTGAACCGTCAGATC*GCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAA
GACACCGGGACCGATCCAGCCTCCGGAC*TCTAGAGGATCCGGTACTCGAGGAACTG*

AAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGAT
CCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGG
*CCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGC*GGCCGCAC
CATGAAGCTGCTGCCGTCGGTGGTGCTGAAGCTCTTTCTGGCTGCAGTTCTCTCG
GCACTGGTGACTGGCGAG**AGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTG**
GAACCAGCAACCCGGACCCTCCCACTGTATCCACGGACCAGCTGCTACCCC
TAGGAGGCGGCCGGGACCGGAAAGTCCGTGACTTGCAAGAGGCAGATCTGG
ACCTTTTGAGAGTCACTTTATCCTCCAAGCCACAAGCACTGGCCACACCAAA
CAAGGAGGAGCACGGGAAAAGAAAGAAGAAAGGCAAGGGGCTAGGGAAGAA
GAGGGACCCA**TGTCTTCGGAAATACAAGGACTTCTGCATCCATGGAGAATGC**
AAATATGTGAAGGAGCTCCGGGCTCCCTCCTGCATCTGCCACCCGGGTTAC
CATGGAGAGAGGTGTCATGGGCTGAGCCTCCCAAAAGTATCTGTGTCTTCAG
GAGGTGACTGCATTCGAACATACAAACCAGAAATCAAGAAAGGAAGTTACAATA
ATATTGTTGTCAACGTAAAGACAGCTGTTGCTGATAACCTCCTCTTTTATCTTGGA
AGTGCCAAATTTATTGACTTTCTGGCTATAGAAATGCGTAAAGGCAAAGTCAGCT
TCCTCTGGGATGTTGGATCTGGAGTTGGACGTGTAGAGTACCCAGATTTGACTAT
TGATGACTCATATTGGTACCGTATCGTAGCATCAAGAACTGGGAGAAATGGAAC
TATTTCTGTGAGAGCCCTGGATGGACCCAAAGCCAGCATTGTGCCCAGCACACAC
CATTCGACGTCTCCTCCAGGGTACACGATTCTAGATGTGGATGCAAATGCAATGC
TGTTTGTTGGTGGCCTGACTGGGAAATTAAAGAAGGCTGATGCTGTACGTGTGAT
TACATTCACTGGCTGCATGGGAGAAACATACTTTGACAACAAACCTATAGGTTTG
TGGAATTTCCGAGAAAAAGAAGGTGACTGCAAAGGATGCACTGTCAGTCCTCAG
GTGGAAGATAGTGAGGGGACTATTCAATTTGATGGAGAAGGTTATGCATTGGTC
AGCCGTCCCATTCGCTGGTACCCCAACATCTCCACTGTCATGTTCAAGTTCAGAA
CATTTTCTTCGAGTGCTCTTCTGATGTATCTTGCCACACGAGACCTGAGAGATTTC
ATGAGTGTGGAGCTCACTGATGGGCACATAAAAGTCAGTTACGATCTGGGCTCA
GGAATGGCTTCCGTTGTCAGCAATCAAAACCATAATGATGGGAAATGGAAATCA
TTCACTCTGTCAAGAATTCAAAAACAAGCCAATATATCAATTGTAGATATAGATA
CTAATCAGGAGGAGAATATAGCAACTTCGTCTTCTGGAAACAACTTTGGTCTTGA
CTTGAAAGCAGATGACAAAATATATTTTGGTGGCCTGCCAACGCTGAGAAACTTG
AGTATGAAAGCAAGGCCAGAAGTAAATCTGAAGAAATATTCCGGCTGCCTCAAA
GATATTGAAATTTCAAGAACTCCGTACAATATACTCAGTAGTCCCGATTATGTTG
GTGTTACCAAAGGATGTTCCCTGGAGAATGTTTACACAGTTAGCTTTCCTAAGCC
TGGTTTTGTGGAGCTCTCCCCTGTGCCAATTGATGTAGGAACAGAAATCAACCTG
TCATTCAGCACCAAGAATGAGTCCGGCATCATTCTTTTGGGAAGTGGAGGGACAC
CAGCACCACCTAGGAGAAAACGAAGG**CAGACTGGACAGGCCTATTATGTAAT**
ACTCCTCAACAGGGGCCGTCTGGAAGTGCATCTCTCCACAGGGGCACGAAC
AATGAGGAAAATTGTCATCAGACCAGAGCCGAATCTGTTTCATGATGGAAGA
GAACATTCCGTTCATGTAGAGCGAACTAGAGGCATCTTTACAGTTCAAGTGG
ATGAAAACAGAAGATACATGCAAAACCTGACAGTTGAACAGCCTATCGAAGT
TAAAAAGCTTTTCGTTGGGGGTGCTCCACCTGAATTTCAACCTTCCCCACTC
AGAAATATTCCTCCTTTTGAAGGCTGCATATGGAATCTTGTTATTAACTCTGT
CCCCATGGACTTTGCAAGGCCTGTGTCCTTCAAAAATGCTGACATTGGTCGC
TGTGCCCATCAGAAACTCCGTGAAGATGAAGATGGAGCAGCTCCAGCTGAA
ATAGTTATCCAGCCTGAGCCAGTTCCCACCCCAGCCTTTCCTACGCCCACCC
CAGTTCTGACACATGGTCCTTGTGCTGCAGAATCAGAACCAGCTCTTTTGAT
AGGGAGCAAGCAGTTCGGGCTTTCAAGAAACAGTCACATTGCAATTGCATTT
GATGACACCAAAGTTAAAAACCGTCTCACAATTGAGTTGGAAGTAAGAACCG
AAGCTGAATCCGGCTTGCTTTTTTACATGGCTCGCATCAATCATGCTGATTT
TGCAACAGTTCAGCTGAGAAATGGATTGCCCTACTTCAGCTATGACTTGGGG

AGTGGGGACACCCACACCATGATCCCCACCAAAATCAATGATGGCCAGTGG
CACAAGATTAAGATAATGAGAAGTAAGCAAGAAGGAATTCTTTATGTAGATG
GGGCTTCCAACAGAACCATCAGTCCCAAAAAAGCCGACATCCTGGATGTCG
TGGGAATGCTGTATGTTGGTGGGTTACCCATCAACTACACTACCCGAAGAAT
TGGTCCAGTGACCTATAGCATTGATGGCTGCGTCAGGAATCTCCACATGGCA
GAGGCCCCTGCCGATCTGGAACAACCCACCTCCAGCTTCCATGTTGGGACA
TGTTTTGCAAATGCTCAGAGGGGAACATATTTTGACGGAACCGGTTTTGCCA
AAGCAGTTGGTGGATTCAAAGTGGGATTGGACCTTCTTGTAGAATTTGAATT
CCGCACAACTACAACGACTGGAGTTCTTCTGGGGATCAGTAGTCAAAAAATG
GATGGAATGGGTATTGAAATGATTGATGAAAAGTTGATGTTTCATGTGGACA
ATGGTGCGGGCAGATTCACTGCTGTCTATGATGCTGGGGTTCCAGGGCATT
TGTGTGATGGACAATGGCATAAAGTCACTGCCAACAAGATCAAACACCGCAT
TGAGCTCACAGTCGATGGGAACCAGGTGGAAGCCCAAAGCCCAAACCCAGC
ATCTACATCAGCTGACACAAATGACCCTGTGTTTGTTGGAGGCTTCCCAGAT
GACCTCAAGCAGTTTGGCCTAACAACCAGTATTCCGTTCCGAGGTTGCATCA
GATCCCTGAAGCTCACCAAAGGCACAGCAAGCCACTGGAGGTTAATTTTGC
CAAGGCCCTGGAACTGAACTAGTGCGGCCGCGGGGATCCAGACATGATAAGAT
ACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTA
TTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAA
CAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGT
GGGAGGTTTTTTTCGGATCCTCTAGAGTCGACCA**CATGGCTACGTAGATAATTAG**
CATGGCGGGTTAATCATTAACTACAAGGAACCCCTAGTGATGGAGTTGGCC
ACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC
GCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCG
CNNNNNNCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCG
TATTGGGCMG

Figure 21

SEQ ID NO: 6

pAAV.CMV.HB.LAMA2(G3-G5)

Main Characteristics
Ampicillin Resistance Gene
ITRs
CMV Promoter
SV40 Enhancer
HBEGF Signal Peptide
HBEGF Propeptide Domain
HBEGF Mature Peptide Domain
HBEGF HB Domain
LAMA2 G3-G5
SV40 PolyA Tail

GCTCTTCCGCTTGGTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAACCCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTACGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA

CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAA
AATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
CGAAAAGTGCCACCTGACGTCTAAGAAACCATCGTTACATAACTTACGGTATATT
ATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGC
GTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCA
CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAG
CGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGT
ACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA
ATACCGCATCAGGAACTTCCAACATCCAATAAATCATACAGGCAAGGCAAAGAA
TTAGCAAAATTAAGCAATAAAGCCTCAGAGCATAAAGCTAAATCGGTTGTACCA
AAAACATTATGACCCTGTAATACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAG
GATAAAAATTTTTAGAACCCTCATATATTTTAAATGCAATGCCTGAGTAATGTGT
AGGTAAAGATTCAAACGGGTGAGAAAGGCCGGAGACAGTCAAATCACCATCAAT
ATGATATTCAACCGTTCTAGCTGATAAATTCATGCCGGAGAGGGTAGCTATTTTT
GAGAGGTCTCTACAAAGGCTATCAGGTCATTGCCTGAGAGTCTGGAGCAAACAA
GAGAATCGATGAACGGTAATCGTAAAACTAGCATGTCAATCATATGTACCCCGG
TTGATAATCAGAAAAGCCCCAAAAACAGGAAGATTGTATAAGCAAATATTTAAA
TTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCA
TTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGA
CCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAA
CGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACT
ACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAATCACTA
AATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCG
AACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCT
GGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGC
GCCGCTACAGGGCGCGTACTATGGTTGCTTTGACGAGCACGTATAACGTGCTTTC
CTCGTTAGAATCAGAGCGGGAGCTAAACAGGAGGCCGATTAAAGGGATTTTAGA
CAGGAACGGTACGCCAGAATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACC
GAGTAAAAGAGTCTGTCCATCACGCAAATTAACCGTTGTCGCAATACTTCTTTGA
TTAGTAATAACATCACTTGCCTGAGTAGAAGAACTCAAACTATCGGCCTTGCTGG
TAATATCCAGAACAATATTACCGCCAGCCATTGCAACAGGAAAAACGCTCATGG
AAATACCTACATTTTGACGCTCAATCGTCTGGAACTTCCATTCGCCATTCAGGCT
GCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG
NNNNNGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCG
TCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGA
GGGAGTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTAACCCG
CCATGCTAATTATCTACGTAGCCATGTCTAGGGT*CGTTACATAACTTACGGTAAAT*
GGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT
GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACG

GTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTG
ACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA
CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTT
TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC
CACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA
ATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA
*GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATC*GCCTGGAGACGCCATCC
ACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGAC*TCTA*
GAGGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTT
TTGTCTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCA
GTGGATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTG
*CGGAATTGTACCCGC*GGCCGCACCATGAAGCTGCTGCCGTCGGTGGTGCTGAAGC
TCTTTCTGGCTGCAGTTCTCTCGGCACTGGTGACTGGCGAG**AGCCTGGAGCGGC**
TTCGGAGAGGGCTAGCTGCTGGAACCAGCAACCCGGACCCTCCCACTGTAT
CCACGGACCAGCTGCTACCCCTAGGAGGCGGCCGGGACCGGAAAGTCCGTG
ACTTGCAAGAGGCAGATCTGGACCTTTTGAGAGTCACTTTATCCTCCAAGCC
ACAAGCACTGGCCACACCAAACAAGGAGGAGCACGGGAAAAGAAAGAAGAA
AGGCAAGGGGCTAGGGAAGAAGAGGGACCCACAGACTGGACAGGCCTATTATG
TAATACTCCTCAACAGGGGCCGTCTGGAAGTGCATCTCTCCACAGGGGCACGAA
CAATGAGGAAAATTGTCATCAGACCAGAGCCGAATCTGTTTCATGATGGAAGAG
AACATTCCGTTCATGTAGAGCGAACTAGAGGCATCTTTACAGTTCAAGTGGATGA
AAACAGAAGATACATGCAAAACCTGACAGTTGAACAGCCTATCGAAGTTAAAAA
GCTTTTCGTTGGGGGTGCTCCACCTGAATTTCAACCTTCCCCACTCAGAAATATTC
CTCCTTTTGAAGGCTGCATATGGAATCTTGTTATTAACTCTGTCCCCATGGACTTT
GCAAGGCCTGTGTCCTTCAAAAATGCTGACATTGGTCGCTGTGCCCATCAGAAAC
TCCGTGAAGATGAAGATGGAGCAGCTCCAGCTGAAATAGTTATCCAGCCTGAGC
CAGTTCCCACCCCAGCCTTTCCTACGCCCACCCCAGTTCTGACACATGGTCCTTGT
GCTGCAGAATCAGAACCAGCTCTTTTGATAGGGAGCAAGCAGTTCGGGCTTTCAA
GAAACAGTCACATTGCAATTGCATTTGATGACACCAAAGTTAAAAACCGTCTCAC
AATTGAGTTGGAAGTAAGAACCGAAGCTGAATCCGGCTTGCTTTTTTACATGGCT
CGCATCAATCATGCTGATTTTGCAACAGTTCAGCTGAGAAATGGATTGCCCTACT
TCAGCTATGACTTGGGGAGTGGGGACACCCACACCATGATCCCCACCAAAATCA
ATGATGGCCAGTGGCACAAGATTAAGATAATGAGAAGTAAGCAAGAAGGAATTC
TTTATGTAGATGGGGCTTCCAACAGAACCATCAGTCCCAAAAAAGCCGACATCCT
GGATGTCGTGGGAATGCTGTATGTTGGTGGGTTACCCATCAACTACACTACCCGA
AGAATTGGTCCAGTGACCTATAGCATTGATGGCTGCGTCAGGAATCTCCACATGG
CAGAGGCCCCTGCCGATCTGGAACAACCCACCTCCAGCTTCCATGTTGGGACATG
TTTTGCAAATGCTCAGAGGGGAACATATTTTGACGGAACCGGTTTTGCCAAAGCA
GTTGGTGGATTCAAAGTGGGATTGGACCTTCTTGTAGAATTTGAATTCCGCACAA
CTACAACGACTGGAGTTCTTCTGGGGATCAGTAGTCAAAAAATGGATGGAATGG
GTATTGAAATGATTGATGAAAAGTTGATGTTTCATGTGGACAATGGTGCGGGCAG
ATTCACTGCTGTCTATGATGCTGGGGTTCCAGGGCATTTGTGTGATGGACAATGG
CATAAAGTCACTGCCAACAAGATCAAACACCGCATTGAGCTCACAGTCGATGGG
AACCAGGTGGAAGCCCAAAGCCCAAACCCAGCATCTACATCAGCTGACACAAAT

GACCCTGTGTTTGTTGGAGGCTTCCCAGATGACCTCAAGCAGTTTGGCCTAACAA
CCAGTATTCCGTTCCGAGGTTGCATCAGATCCCTGAAGCTCACCAAAGGCACAGC
AAGCCACTGGAGGTTAATTTTGCCAAGGCCCTGGAACTGAACTAGTGCGGCCGC
GGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGA
ATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGT
AACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG
TTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTCGGATCCTCTAGAGTCGACCA
CATGGCTACGTAGATAATTAGCATGGCGGGTTAATCATTAACTACAAGGAAC
CCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGA
GGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCT
CAGTGAGCGAGCGAGCGCGCNNNNNNCAGCTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTGCGTATTGGGC

Figure 22

pAAV.CMV.HBEGF.LAMA2(G3-G5) (SEQ ID NO: 8)

Main Characteristics
Ampicillin Resistance Gene
ITRs
CMV Promoter
SV40 Enhancer
HBEGF Signal Peptide
HBEGF Propeptide Domain
HBEGF Mature Peptide Domain
HBEGF HB Domain
HBEGF EGF-Like Domain
LAMA2 G3-5
SV40 PolyA Tail

GCTCTTCCGCTTGGTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAACCCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTACGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA
CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAA
AATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC

TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAACTTCC
AACATCCAATAAATCATACAGGCAAGGCAAAGAATTAGCAAAATTAAGCAATAA
AGCCTCAGAGCATAAAGCTAAATCGGTTGTACCAAAAACATTATGACCCTGTAAT
ACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAGGATAAAAATTTTTAGAACCCT
CATATATTTTAAATGCAATGCCTGAGTAATGTGTAGGTAAAGATTCAAACGGGTG
AGAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATTCAACCGTTCTAGC
TGATAAATTCATGCCGGAGAGGGTAGCTATTTTTGAGAGGTCTCTACAAAGGCTA
TCAGGTCATTGCCTGAGAGTCTGGAGCAAACAAGAGAATCGATGAACGGTAATC
GTAAAACTAGCATGTCAATCATATGTACCCCGGTTGATAATCAGAAAAGCCCCA
AAAACAGGAAGATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAATCACTAAATCGGAACCCTAAAGGGA
GCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAA
GGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAC
GCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTA
CTATGGTTGCTTTGACGAGCACGTATAACGTGCTTTCCTCGTTAGAATCAGAGCG
GGAGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGCCAGA
ATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCA
TCACGCAAATTAACCGTTGTCGCAATACTTCTTTGATTAGTAATAACATCACTTGC
CTGAGTAGAAGAACTCAAACTATCGGCCTTGCTGGTAATATCCAGAACAATATTA
CCGCCAGCCATTGCAACAGGAAAAACGCTCATGGAAATACCTACATTTTGACGCT
CAATCGTCTGGAACTTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG
ATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG**NNNNNNGCGCGCTCGCTCGC**
TCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCC
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA
CTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTAATTATCTACGTA
GCCATGTCTAGGGT*CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC*
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC
GCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG
TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCAT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT
*AGTGAACCGTCAGATC*GCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAA
GACACCGGGACCGATCCAGCCTCCGGAC*TCTAGAGGATCCGGTACTCGAGGAACTG*
AAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGAT
CCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGG
*CCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGC*GGCCGCAC

CATGAAGCTGCTGCCGTCGGTGGTGCTGAAGCTCTTTCTGGCTGCAGTTCTCTCG
GCACTGGTGACTGGCGAG**AGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTG**
GAACCAGCAACCCGGACCCTCCCACTGTATCCACGGACCAGCTGCTACCCC
TAGGAGGCGGCCGGGACCGGAAAGTCCGTGACTTGCAAGAGGCAGATCTGG
ACCTTTTGAGAGTCACTTTATCCTCCAAGCCACAAGCACTGGCCACACCAAA
CAAGGAGGAGCACGGGAAAAGAAAGAAGAAAGGCAAGGGGCTAGGGAAGAA
GAGGGACCCA**TGTCTTCGGAAATACAAGGACTTCTGCATCCATGGAGAATGC**
AAATATGTGAAGGAGCTCCGGGCTCCCTCCTGCATCTGCCACCCGGGTTAC
CATGGAGAGAGGTGTCATGGGCTGAGCCTCCCACAGACTGGACAGGCCTATT
ATGTAATACTCCTCAACAGGGGCCGTCTGGAAGTGCATCTCTCCACAGGGGCACG
AACAATGAGGAAAATTGTCATCAGACCAGAGCCGAATCTGTTTCATGATGGAAG
AGAACATTCCGTTCATGTAGAGCGAACTAGAGGCATCTTTACAGTTCAAGTGGAT
GAAAACAGAAGATACATGCAAAACCTGACAGTTGAACAGCCTATCGAAGTTAAA
AAGCTTTTCGTTGGGGGTGCTCCACCTGAATTTCAACCTTCCCCACTCAGAAATA
TTCCTCCTTTTGAAGGCTGCATATGGAATCTTGTTATTAACTCTGTCCCCATGGAC
TTTGCAAGGCCTGTGTCCTTCAAAAATGCTGACATTGGTCGCTGTGCCCATCAGA
AACTCCGTGAAGATGAAGATGGAGCAGCTCCAGCTGAAATAGTTATCCAGCCTG
AGCCAGTTCCCACCCCAGCCTTTCCTACGCCCACCCCAGTTCTGACACATGGTCC
TTGTGCTGCAGAATCAGAACCAGCTCTTTTGATAGGGAGCAAGCAGTTCGGGCTT
TCAAGAAACAGTCACATTGCAATTGCATTTGATGACACCAAAGTTAAAAACCGTC
TCACAATTGAGTTGGAAGTAAGAACCGAAGCTGAATCCGGCTTGCTTTTTTACAT
GGCTCGCATCAATCATGCTGATTTTGCAACAGTTCAGCTGAGAAATGGATTGCCC
TACTTCAGCTATGACTTGGGGAGTGGGGACACCCACACCATGATCCCCACCAAA
ATCAATGATGGCCAGTGGCACAAGATTAAGATAATGAGAAGTAAGCAAGAAGG
AATTCTTTATGTAGATGGGGCTTCCAACAGAACCATCAGTCCCAAAAAAGCCGAC
ATCCTGGATGTCGTGGGAATGCTGTATGTTGGTGGGTTACCCATCAACTACACTA
CCCGAAGAATTGGTCCAGTGACCTATAGCATTGATGGCTGCGTCAGGAATCTCCA
CATGGCAGAGGCCCCTGCCGATCTGGAACAACCCACCTCCAGCTTCCATGTTGGG
ACATGTTTTGCAAATGCTCAGAGGGGAACATATTTTGACGGAACCGGTTTTGCCA
AAGCAGTTGGTGGATTCAAAGTGGGATTGGACCTTCTTGTAGAATTTGAATTCCG
CACAACTACAACGACTGGAGTTCTTCTGGGGATCAGTAGTCAAAAAATGGATGG
AATGGGTATTGAAATGATTGATGAAAAGTTGATGTTTCATGTGGACAATGGTGCG
GGCAGATTCACTGCTGTCTATGATGCTGGGGTTCCAGGGCATTTGTGTGATGGAC
AATGGCATAAAGTCACTGCCAACAAGATCAAACACCGCATTGAGCTCACAGTCG
ATGGGAACCAGGTGGAAGCCCAAAGCCCAAACCCAGCATCTACATCAGCTGACA
CAAATGACCCTGTGTTTGTTGGAGGCTTCCCAGATGACCTCAAGCAGTTTGGCCT
AACAACCAGTATTCCGTTCCGAGGTTGCATCAGATCCCTGAAGCTCACCAAAGGC
ACAGCAAGCCACTGGAGGTTAATTTTGCCAAGGCCCTGGAACTGAACTAGTGCG
GCCGCGGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAA
CTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTT
ATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAT
TTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTCGGATCCTCTAGAGT
CGACCACATGGCTACGTAGATAATTAGCATGGCGGGTTAATCATTAACTACAAG
GAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTG
AGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAG
TGAGCGAGCGAGCGCGCNNNNNNCAGCTGCATTAATGAATCGGCCAACGCGCGG
GGAGAGGCGGTTTGCGTATTGGGC

Figure 23

pAAV.CMV.HB.DAG1a (SEQ ID NO: 10)

Main Characteristics
Ampicillin Resistance Gene
ITRs
CMV Promoter
SV40 Enhancer
HBEGF Signal Peptide
HBEGF Propeptide Domain
HBEGF Mature Peptide Domain
HBEGF HB Domain
DAG1a
SV40 PolyA Tail

GCTCTTCCGCTTGGTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAACCCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTACGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA
CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAA
AATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC

TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAACTTCC
AACATCCAATAAATCATACAGGCAAGGCAAAGAATTAGCAAAATTAAGCAATAA
AGCCTCAGAGCATAAAGCTAAATCGGTTGTACCAAAAACATTATGACCCTGTAAT
ACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAGGATAAAAATTTTTAGAACCCT
CATATATTTTAAATGCAATGCCTGAGTAATGTGTAGGTAAAGATTCAAACGGGTG
AGAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATTCAACCGTTCTAGC
TGATAAATTCATGCCGGAGAGGGTAGCTATTTTTGAGAGGTCTCTACAAAGGCTA
TCAGGTCATTGCCTGAGAGTCTGGAGCAAACAAGAGAATCGATGAACGGTAATC
GTAAAACTAGCATGTCAATCATATGTACCCCGGTTGATAATCAGAAAAGCCCCA
AAAACAGGAAGATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAATCACTAAATCGGAACCCTAAAGGGA
GCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAA
GGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAC
GCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTA
CTATGGTTGCTTTGACGAGCACGTATAACGTGCTTTCCTCGTTAGAATCAGAGCG
GGAGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGCCAGA
ATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCA
TCACGCAAATTAACCGTTGTCGCAATACTTCTTTGATTAGTAATAACATCACTTGC
CTGAGTAGAAGAACTCAAACTATCGGCCTTGCTGGTAATATCCAGAACAATATTA
CCGCCAGCCATTGCAACAGGAAAAACGCTCATGGAAATACCTACATTTTGACGCT
CAATCGTCTGGAACTTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG
ATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG**NNNNNNGCGCGCTCGCTCGC**
TCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCC
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA
CTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTAATTATCTACGTA
GCCATGTCTAGGGT*CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC*
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC
GCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG
TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCAT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT
*AGTGAACCGTCAGATC*GCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAA
GACACCGGGACCGATCCAGCCTCCGGAC*TCTAGAGGATCCGGTACTCGAGGAACTG*
AAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGAT
CCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGG
*CCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGC*GGCCGCAC

CATGAAGCTGCTGCCGTCGGTGGTGCTGAAGCTCTTTCTGGCTGCAGTTCTCTCG
GCACTGGTGACTGGCGAG**AGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTG**
GAACCAGCAACCCGGACCCTCCCACTGTATCCACGGACCAGCTGCTACCCC
TAGGAGGCGGCCGGGACCGGAAAGTCCGTGACTTGCAAGAGGCAGATCTGG
ACCTTTTGAGAGTCACTTTATCCTCCAAGCCACAAGCACTGGCCACACCAAA
CAAGGAGGAGCACGGGAAAAGAAAGAAGAAAGGCAAGGGGCTAGGGAAGAA
GAGGGACCCACAGATCCATGCTACACCCACACCTGTCACTGCCATTGGGCCCCCA
ACCACGGCTATCCAGGAGCCCCCATCCAGGATCGTGCCAACCCCCACATCTCCAG
CCATTGCTCCTCCAACAGAGACCATGGCTCCTCCAGTCAGGGATCCTGTTCCTGG
GAAACCCACGGTCACCATCCGGACTCGAGGCGCCATTATTCAAACCCCAACCCTA
GGCCCCATCCAGCCTACTCGGGTGTCAGAAGCTGGCACCACAGTTCCTGGCCAGA
TTCGCCCAACGATGACCATTCCTGGCTATGTGGAGCCTACTGCAGTTGCTACCCC
TCCCACAACCACCACCAAGAAGCCACGAGTATCCACACCAAAACCAGCAACGCC
TTCAACTGACTCCACCACCACCACGACTCGCAGGCCAACCAAGAAACCACGGAC
ACCCCGGCCAGTGCCCCGGGTCACCACCAAAGTTTCCATCACCAGATTGGAAACT
GCCTCACCGCCTACTCGTATTCGCACCACCACCAGTGGAGTGCCCCGTGGCGGAG
AACCCAACCAGCGCCCAGAGCTCAAGAACCATATTGACAGGGTAGATGCCTGGG
TTGGCACCTACTTTGAGGTGAAGATCCCGTCAGACACTTTCTATGACCATGAGGA
CACCACCACTGACAAGCTGAAGCTGACCCTGAAACTGCGGGAGCAGCAGCTGGT
GGGCGAGAAGTCCTGGGTACAGTTCAACAGCAACAGCCAGCTCATGTATGGCCT
TCCCGACAGCAGCCACGTGGGCAAACACGAGTATTTCATGCATGCCACAGACAA
GGGGGGCCTGTCGGCTGTGGATGCCTTCGAGATCCACGTCCACAGGCGCCCCCA
AGGGGATAGGGCTCCTGCAAGGTTCAAGGCCAAGTTTGTGGGTGACCCGGCACT
GGTGTTGAATGACATCCACAAGAAGATTGCCTTGGTAAAGAAACTGGCCTTCGCC
TTTGGAGACCGAAACTGTAGCACCATCACCCTGCAGAATATCACCCGGGGCTAA
ACTAGTGCGGCCGCGGGGATCCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGC
TATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAT
TGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTCGGATCC
TCTAGAGTCGACCA**CATGGCTACGTAGATAATTAGCATGGCGGGTTAATCATT**
AACTACAAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTG
CCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCNNNNNNCAGCTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGC

Figure 24

pAAV.CMV.HBEGF.DAG1 (SEQ ID NO: 12)

Main Characteristics
Ampicillin Resistance Gene
ITRs
CMV Promoter
SV40 Enhancer
HBEGF Signal Peptide
HBEGF Propeptide Domain
HBEGF Mature Peptide Domain
HBEGF HB Domain
HBEGF EGF-Like Domain
DAG1a
SV40 PolyA Tail

GCTCTTCCGCTTGGTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAACCCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTACGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA
CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAA

AATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC
CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAACTTCC
AACATCCAATAAATCATACAGGCAAGGCAAAGAATTAGCAAAATTAAGCAATAA
AGCCTCAGAGCATAAAGCTAAATCGGTTGTACCAAAAACATTATGACCCTGTAAT
ACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAGGATAAAAATTTTTAGAACCCT
CATATATTTTAAATGCAATGCCTGAGTAATGTGTAGGTAAAGATTCAAACGGGTG
AGAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATTCAACCGTTCTAGC
TGATAAATTCATGCCGGAGAGGGTAGCTATTTTTGAGAGGTCTCTACAAAGGCTA
TCAGGTCATTGCCTGAGAGTCTGGAGCAAACAAGAGAATCGATGAACGGTAATC
GTAAAACTAGCATGTCAATCATATGTACCCCGGTTGATAATCAGAAAAGCCCCA
AAAACAGGAAGATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAATCACTAAATCGGAACCCTAAAGGGA
GCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAA
GGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAC
GCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTA
CTATGGTTGCTTTGACGAGCACGTATAACGTGCTTTCCTCGTTAGAATCAGAGCG
GGAGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGCCAGA
ATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCA
TCACGCAAATTAACCGTTGTCGCAATACTTCTTTGATTAGTAATAACATCACTTGC
CTGAGTAGAAGAACTCAAACTATCGGCCTTGCTGGTAATATCCAGAACAATATTA
CCGCCAGCCATTGCAACAGGAAAAACGCTCATGGAAATACCTACATTTTGACGCT
CAATCGTCTGGAACTTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG
ATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG**NNNNNNGCGCGCTCGCTCGC**
TCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCC
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA
CTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTAATTATCTACGTA
GCCATGTCTAGGGT*CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC*
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC
GCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG
TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCAT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT
*AGTGAACCGTCAGATC*GCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAA
GACACCGGGACCGATCCAGCCTCCGGAC*TCTAGAGGATCCGGTACTCGAGGAACTG*
AAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGAT
CCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGG

*CCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGC*GGCCGCAC
CATGAAGCTGCTGCCGTCGGTGGTGCTGAAGCTCTTTCTGGCTGCAGTTCTCTCG
GCACTGGTGACTGGCGAG**AGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTG**
GAACCAGCAACCCGGACCCTCCCACTGTATCCACGGACCAGCTGCTACCCC
TAGGAGGCGGCCGGGACCGGAAAGTCCGTGACTTGCAAGAGGCAGATCTGG
ACCTTTTGAGAGTCACTTTATCCTCCAAGCCACAAGCACTGGCCACACCAAA
CAAGGAGGAGCACGGGAAAAGAAAGAAGAAAGGCAAGGGGCTAGGGAAGAA
GAGGGACCCA**TGTCTTCGGAAATACAAGGACTTCTGCATCCATGGAGAATGC**
AAATATGTGAAGGAGCTCCGGGCTCCCTCCTGCATCTGCCACCCGGGTTAC
CATGGAGAGAGGTGTCATGGGCTGAGCCTCCCACAGATCCATGCTACACCCA
CACCTGTCACTGCCATTGGGCCCCCAACCACGGCTATCCAGGAGCCCCCATCCAG
GATCGTGCCAACCCCCACATCTCCAGCCATTGCTCCTCCAACAGAGACCATGGCT
CCTCCAGTCAGGGATCCTGTTCCTGGGAAACCCACGGTCACCATCCGGACTCGAG
GCGCCATTATTCAAACCCCAACCCTAGGCCCCATCCAGCCTACTCGGGTGTCAGA
AGCTGGCACCACAGTTCCTGGCCAGATTCGCCCAACGATGACCATTCCTGGCTAT
GTGGAGCCTACTGCAGTTGCTACCCCTCCCACAACCACCACCAAGAAGCCACGA
GTATCCACACCAAAACCAGCAACGCCTTCAACTGACTCCACCACCACCACGACTC
GCAGGCCAACCAAGAAACCACGGACACCCCGGCCAGTGCCCCGGGTCACCACCA
AAGTTTCCATCACCAGATTGGAAACTGCCTCACCGCCTACTCGTATTCGCACCAC
CACCAGTGGAGTGCCCCGTGGCGGAGAACCCAACCAGCGCCCAGAGCTCAAGAA
CCATATTGACAGGGTAGATGCCTGGGTTGGCACCTACTTTGAGGTGAAGATCCCG
TCAGACACTTTCTATGACCATGAGGACACCACCACTGACAAGCTGAAGCTGACCC
TGAAACTGCGGGAGCAGCAGCTGGTGGGCGAGAAGTCCTGGGTACAGTTCAACA
GCAACAGCCAGCTCATGTATGGCCTTCCCGACAGCAGCCACGTGGGCAAACACG
AGTATTTCATGCATGCCACAGACAAGGGGGGCCTGTCGGCTGTGGATGCCTTCGA
GATCCACGTCCACAGGCGCCCCCAAGGGGATAGGGCTCCTGCAAGGTTCAAGGC
CAAGTTTGTGGGTGACCCGGCACTGGTGTTGAATGACATCCACAAGAAGATTGCC
TTGGTAAAGAAACTGGCCTTCGCCTTTGGAGACCGAAACTGTAGCACCATCACCC
TGCAGAATATCACCCGGGGCTAAACTAGTGCGGCCGCGGGGATCCAGACATGAT
AAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATG
CTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA
ATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGA
GGTGTGGGAGGTTTTTTCGGATCCTCTAGAGTCGACCA**CATGGCTACGTAGATA**
ATTAGCATGGCGGGTTAATCATTAACTACAAGGAACCCCTAGTGATGGAGTT
GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAG
CGCGCNNNNNNCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGT
TTGCGTATTGGGC

SEQUENCE LISTING

<110> RESEARCH INSTITUTE AT NATIONWIDE CHILDRENS HOSPITAL

<120> RECOMBINANT ADENO-ASSOCIATED VIRUS PRODUCTS AND METHODS FOR TREAING DYSTROGLYCANOPATHIES AND LAMININ-DEFICIENT MUSCULAR DYSTROPHIES

<130> 28335/53147A

<150> US 62/686,522
<151> 2018-06-18

<160> 24

<170> PatentIn version 3.5

<210> 1
<211> 3222
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 1
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc aaaagtatct gtgtcttcag gaggtgactg cattcgaaca 360

tacaaaccag aaatcaagaa aggaagttac aataatattg ttgtcaacgt aaagacagct 420

gttgctgata acctcctctt ttatcttgga agtgccaaat ttattgactt tctggctata 480

gaaatgcgta aaggcaaagt cagcttcctc tgggatgttg gatctggagt tggacgtgta 540

gagtacccag atttgactat tgatgactca tattggtacc gtatcgtagc atcaagaact 600

gggagaaatg gaactatttc tgtgagagcc ctggatggac ccaaagccag cattgtgccc 660

agcacacacc attcgacgtc tcctccaggg tacacgattc tagatgtgga tgcaaatgca 720

atgctgtttg ttggtggcct gactgggaaa ttaaagaagg ctgatgctgt acgtgtgatt 780

acattcactg gctgcatggg agaaacatac tttgacaaca aacctatagg tttgtggaat 840

ttccgagaaa aagaaggtga ctgcaaagga tgcactgtca gtcctcaggt ggaagatagt 900

gaggggacta ttcaatttga tggagaaggt tatgcattgg tcagccgtcc cattcgctgg 960

taccccaaca tctccactgt catgttcaag ttcagaacat tttcttcgag tgctcttctg 1020

atgtatcttg ccacacgaga cctgagagat ttcatgagtg tggagctcac tgatgggcac 1080

ataaaagtca gttacgatct gggctcagga atggcttccg ttgtcagcaa tcaaaaccat 1140

aatgatggga aatggaaatc attcactctg tcaagaattc aaaaacaagc caatatatca 1200

attgtagata tagatactaa tcaggaggag aatatagcaa cttcgtcttc tggaaacaac 1260

tttggtcttg acttgaaagc agatgacaaa atatattttg gtggcctgcc aacgctgaga 1320

aacttgagta tgaaagcaag gccagaagta aatctgaaga aatattccgg ctgcctcaaa 1380

gatattgaaa tttcaagaac tccgtacaat atactcagta gtcccgatta tgttggtgtt 1440

accaaaggat gttccctgga gaatgtttac acagttagct ttcctaagcc tggttttgtg 1500

gagctctccc ctgtgccaat tgatgtagga acagaaatca acctgtcatt cagcaccaag 1560

aatgagtccg gcatcattct tttgggaagt ggagggacac cagcaccacc taggagaaaa 1620

cgaaggcaga ctggacaggc ctattatgta atactcctca acaggggccg tctggaagtg 1680

catctctcca caggggcacg aacaatgagg aaaattgtca tcagaccaga gccgaatctg 1740

tttcatgatg gaagagaaca ttccgttcat gtagagcgaa ctagaggcat ctttacagtt 1800

caagtggatg aaaacagaag atacatgcaa aacctgacag ttgaacagcc tatcgaagtt 1860

aaaaagcttt tcgttggggg tgctccacct gaatttcaac cttccccact cagaaatatt 1920

cctccttttg aaggctgcat atggaatctt gttattaact ctgtccccat ggactttgca 1980

aggcctgtgt ccttcaaaaa tgctgacatt ggtcgctgtg cccatcagaa actccgtgaa 2040

gatgaagatg gagcagctcc agctgaaata gttatccagc ctgagccagt tcccacccca 2100

gcctttccta cgcccacccc agttctgaca catggtcctt gtgctgcaga atcagaacca 2160

gctcttttga tagggagcaa gcagttcggg ctttcaagaa acagtcacat tgcaattgca 2220

tttgatgaca ccaaagttaa aaaccgtctc acaattgagt tggaagtaag aaccgaagct 2280

gaatccggct tgctttttta catggctcgc atcaatcatg ctgattttgc aacagttcag 2340

ctgagaaatg gattgcccta cttcagctat gacttgggga gtggggacac ccacaccatg 2400

atccccacca aaatcaatga tggccagtgg cacaagatta agataatgag aagtaagcaa 2460

gaaggaattc tttatgtaga tggggcttcc aacagaacca tcagtcccaa aaaagccgac 2520

atcctggatg tcgtgggaat gctgtatgtt ggtgggttac ccatcaacta cactacccga 2580

agaattggtc cagtgaccta tagcattgat ggctgcgtca ggaatctcca catggcagag 2640

gcccctgccg atctggaaca acccacctcc agcttccatg ttgggacatg ttttgcaaat 2700

gctcagaggg gaacatattt tgacggaacc ggttttgcca aagcagttgg tggattcaaa 2760

gtgggattgg accttcttgt agaatttgaa ttccgcacaa ctacaacgac tggagttctt 2820

ctggggatca gtagtcaaaa aatggatgga atgggtattg aaatgattga tgaaaagttg 2880

atgtttcatg tggacaatgg tgcgggcaga ttcactgctg tctatgatgc tggggttcca 2940

gggcatttgt gtgatggaca atggcataaa gtcactgcca acaagatcaa acaccgcatt 3000

gagctcacag tcgatgggaa ccaggtggaa gcccaaagcc caaacccagc atctacatca 3060

gctgacacaa atgaccctgt gtttgttgga ggcttcccag atgacctcaa gcagtttggc 3120

ctaacaacca gtattccgtt ccgaggttgc atcagatccc tgaagctcac caaaggcaca 3180

gcaagccact ggaggttaat tttgccaagg ccctggaact ga 3222

<210> 2
<211> 8270
<212> DNA
<213> rAAV

<220>
<221> misc_feature
<222> (3590)..(3595)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (8210)..(8215)
<223> n is a, c, g, t or u

<400> 2
gctcttccgc ttggtcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 60

tatcagctca ctcaaacccg gtaatacggt tatccacaga atcaggggat aacgcaggaa 120

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 180

cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga 240

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 300

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 360

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc 420

gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 480

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca 540

ctggtaacag gattagcaga gcgaggtatg tacgcggtgc tacagagttc ttgaagtggt 600

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 660

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 720

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 780

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 840

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 900

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 960

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg 1020

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac 1080

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 1140

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 1200

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta 1260

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 1320

gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc 1380

ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac 1440

tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 1500

caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa 1560

tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 1620

cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca 1680

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 1740

aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac 1800

tcatactctt ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 1860

gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc 1920

gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata 1980

ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 2040

acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 2100

cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 2160

cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa 2220

ggagaaaata ccgcatcagg aacttccaac atccaataaa tcatacaggc aaggcaaaga 2280

attagcaaaa ttaagcaata aagcctcaga gcataaagct aaatcggttg taccaaaaac 2340

attatgaccc tgtaatactt ttgcgggaga agcctttatt tcaacgcaag gataaaaatt 2400

tttagaaccc tcatatattt taaatgcaat gcctgagtaa tgtgtaggta aagattcaaa 2460

cgggtgagaa aggccggaga cagtcaaatc accatcaata tgatattcaa ccgttctagc 2520

tgataaattc atgccggaga gggtagctat ttttgagagg tctctacaaa ggctatcagg 2580

tcattgcctg agagtctgga gcaaacaaga gaatcgatga acggtaatcg taaaactagc 2640

atgtcaatca tatgtacccc ggttgataat cagaaaagcc ccaaaaacag gaagattgta 2700

taagcaaata tttaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt 2760

taaatcagct catttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa 2820

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 2880

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 2940

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaatcact aaatcggaac 3000

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 3060

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 3120

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgta ctatggttgc 3180

tttgacgagc acgtataacg tgctttcctc gttagaatca gagcgggagc taaacaggag 3240

gccgattaaa gggattttag acaggaacgg tacgccagaa tcctgagaag tgtttttata 3300

atcagtgagg ccaccgagta aaagagtctg tccatcacgc aaattaaccg ttgtcgcaat 3360

acttctttga ttagtaataa catcacttgc ctgagtagaa gaactcaaac tatcggcctt 3420

gctggtaata tccagaacaa tattaccgcc agccattgca acaggaaaaa cgctcatgga 3480

aatacctaca ttttgacgct caatcgtctg gaacttccat tcgccattca ggctgcgcaa 3540

ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgn nnnnngcgcg 3600

ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc 3660

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 3720

ccttgtagtt aatgattaac ccgccatgct aattatctac gtagccatgt ctagggtcgt 3780

tacataactt acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac 3840

gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg 3900

ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaag 3960

tacgcccccc attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtacat 4020

gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat 4080

ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt 4140

tccaagtctc caccccattg acgtcaatgg gagtttgttt tggcaccaaa atcaacggga 4200

ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg 4260

gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcgcct ggagacgcca 4320

tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc ggactctaga 4380

ggatccggta ctcgaggaac tgaaaaacca gaaagttaac tggtaagttt agtctttttg 4440

tcttttattt caggtcccgg atccggtggt ggtgcaaatc aaagaactgc tcctcagtgg 4500

atgttgcctt tacttctagg cctgtacgga agtgttactt ctgctctaaa agctgcggaa 4560

ttgtacccgc ggccgcacca tgaagctgct gccgtcggtg gtgctgaagc tctttctggc 4620

tgcagttctc tcggcactgg tgactggcga gagcctggag cggcttcgga gagggctagc 4680

tgctggaacc agcaacccgg accctcccac tgtatccacg gaccagctgc taccctagg 4740

aggcggccgg gaccggaaag tccgtgactt gcaagaggca gatctggacc ttttgagagt 4800

cactttatcc tccaagccac aagcactggc cacaccaaac aaggaggagc acgggaaaag 4860

aaagaagaaa ggcaaggggc tagggaagaa gagggaccca aaagtatctg tgtcttcagg 4920

aggtgactgc attcgaacat acaaaccaga aatcaagaaa ggaagttaca ataatattgt 4980

tgtcaacgta aagacagctg ttgctgataa cctcctcttt tatcttggaa gtgccaaatt 5040

tattgacttt ctggctatag aaatgcgtaa aggcaaagtc agcttcctct gggatgttgg 5100

atctggagtt ggacgtgtag agtacccaga tttgactatt gatgactcat attggtaccg 5160

tatcgtagca tcaagaactg ggagaaatgg aactatttct gtgagagccc tggatggacc 5220

caaagccagc attgtgccca gcacacacca ttcgacgtct cctccagggt acacgattct 5280

agatgtggat gcaaatgcaa tgctgtttgt tggtggcctg actgggaaat taaagaaggc 5340

tgatgctgta cgtgtgatta cattcactgg ctgcatggga gaaacatact ttgacaacaa 5400

acctataggt ttgtggaatt tccgagaaaa agaaggtgac tgcaaaggat gcactgtcag 5460

tcctcaggtg gaagatagtg aggggactat tcaatttgat ggagaaggtt atgcattggt 5520

cagccgtccc attcgctggt accccaacat ctccactgtc atgttcaagt tcagaacatt 5580

ttcttcgagt gctcttctga tgtatcttgc cacacgagac ctgagagatt tcatgagtgt 5640

ggagctcact gatgggcaca taaaagtcag ttacgatctg ggctcaggaa tggcttccgt 5700

tgtcagcaat caaaaccata atgatgggaa atggaaatca ttcactctgt caagaattca 5760

aaaacaagcc aatatatcaa ttgtagatat agatactaat caggaggaga atatagcaac 5820

ttcgtcttct ggaaacaact ttggtcttga cttgaaagca gatgacaaaa tatattttgg 5880

tggcctgcca acgctgagaa acttgagtat gaaagcaagg ccagaagtaa atctgaagaa 5940

atattccggc tgcctcaaag atattgaaat ttcaagaact ccgtacaata tactcagtag 6000

tcccgattat gttggtgtta ccaaaggatg ttccctggag aatgtttaca cagttagctt 6060

tcctaagcct ggttttgtgg agctctcccc tgtgccaatt gatgtaggaa cagaaatcaa 6120

cctgtcattc agcaccaaga atgagtccgg catcattctt ttgggaagtg gagggacacc 6180

agcaccacct aggagaaaac gaaggcagac tggacaggcc tattatgtaa tactcctcaa 6240

caggggccgt ctggaagtgc atctctccac aggggcacga acaatgagga aaattgtcat 6300

cagaccagag ccgaatctgt ttcatgatgg aagagaacat tccgttcatg tagagcgaac 6360

tagaggcatc tttacagttc aagtggatga aaacagaaga tacatgcaaa acctgacagt 6420

tgaacagcct atcgaagtta aaaagctttt cgttgggggt gctccacctg aatttcaacc 6480

ttccccactc agaaatattc ctccttttga aggctgcata tggaatcttg ttattaactc 6540

tgtccccatg gactttgcaa ggcctgtgtc cttcaaaaat gctgacattg gtcgctgtgc 6600

ccatcagaaa ctccgtgaag atgaagatgg agcagctcca gctgaaatag ttatccagcc 6660

tgagccagtt cccaccccag cctttcctac gcccacccca gttctgacac atggtccttg 6720

tgctgcagaa tcagaaccag ctcttttgat agggagcaag cagttcgggc tttcaagaaa 6780

cagtcacatt gcaattgcat ttgatgacac caaagttaaa aaccgtctca caattgagtt 6840

ggaagtaaga accgaagctg aatccggctt gctttttttac atggctcgca tcaatcatgc 6900

tgattttgca acagttcagc tgagaaatgg attgccctac ttcagctatg acttggggag 6960

tggggacacc cacaccatga tccccaccaa aatcaatgat ggccagtggc acaagattaa 7020

gataatgaga agtaagcaag aaggaattct ttatgtagat ggggcttcca acagaaccat 7080

cagtcccaaa aaagccgaca tcctggatgt cgtgggaatg ctgtatgttg gtgggttacc 7140

catcaactac actacccgaa gaattggtcc agtgacctat agcattgatg gctgcgtcag 7200

gaatctccac atggcagagg cccctgccga tctggaacaa cccacctcca gcttccatgt 7260

tgggacatgt tttgcaaatg ctcagagggg aacatatttt gacggaaccg gttttgccaa 7320

agcagttggt ggattcaaag tgggattgga ccttcttgta gaatttgaat tccgcacaac 7380

tacaacgact ggagttcttc tggggatcag tagtcaaaaa atggatggaa tgggtattga 7440

aatgattgat gaaaagttga tgtttcatgt ggacaatggt gcgggcagat tcactgctgt 7500

ctatgatgct ggggttccag ggcatttgtg tgatggacaa tggcataaag tcactgccaa 7560

caagatcaaa caccgcattg agctcacagt cgatgggaac caggtggaag cccaaagccc 7620

aaacccagca tctacatcag ctgacacaaa tgaccctgtg tttgttggag gcttcccaga 7680

tgacctcaag cagtttggcc taacaaccag tattccgttc cgaggttgca tcagatccct 7740

gaagctcacc aaaggcacag caagccactg gaggttaatt ttgccaaggc cctggaactg 7800

aactagtgcg gccgcgggga tccagacatg ataagataca ttgatgagtt tggacaaacc 7860

acaactagaa tgcagtgaaa aaaatgcttt atttgtgaaa tttgtgatgc tattgcttta 7920

tttgtaacca ttataagctg caataaacaa gttaacaaca acaattgcat tcattttatg 7980

tttcaggttc agggggaggt gtgggaggtt ttttcggatc ctctagagtc gaccacatgg 8040

ctacgtagat aattagcatg gcgggttaat cattaactac aaggaacccc tagtgatgga 8100

gttggccact ccctctctgc gcgctcgctc gctcactgag gccgggcgac caaaggtcgc 8160

ccgacgcccg ggctttgccc gggcggcctc agtgagcgag cgagcgcgcn nnnnncagct 8220

gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc 8270

<210> 3
<211> 3348
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 3
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc atgtcttcgg aaatacaagg acttctgcat ccatggagaa 360

tgcaaatatg tgaaggagct ccgggctccc tcctgcatct gccacccggg ttaccatgga 420

gagaggtgtc atgggctgag cctcccaaaa gtatctgtgt cttcaggagg tgactgcatt 480

cgaacataca aaccagaaat caagaaagga agttacaata atattgttgt caacgtaaag 540

acagctgttg ctgataacct cctcttttat cttggaagtg ccaaatttat tgactttctg 600

gctatagaaa tgcgtaaagg caaagtcagc ttcctctggg atgttggatc tggagttgga 660

cgtgtagagt acccagattt gactattgat gactcatatt ggtaccgtat cgtagcatca 720

agaactggga gaaatggaac tatttctgtg agagccctgg atggacccaa agccagcatt 780

gtgcccagca cacaccattc gacgtctcct ccagggtaca cgattctaga tgtggatgca 840

aatgcaatgc tgtttgttgg tggcctgact gggaaattaa agaaggctga tgctgtacgt 900

gtgattacat tcactggctg catgggagaa acatactttg acaacaaacc tataggtttg 960

tggaatttcc gagaaaaaga aggtgactgc aaaggatgca ctgtcagtcc tcaggtggaa 1020

gatagtgagg ggactattca atttgatgga gaaggttatg cattggtcag ccgtcccatt 1080

cgctggtacc ccaacatctc cactgtcatg ttcaagttca gaacattttc ttcgagtgct 1140

cttctgatgt atcttgccac acgagacctg agagatttca tgagtgtgga gctcactgat 1200

gggcacataa aagtcagtta cgatctgggc tcaggaatgg cttccgttgt cagcaatcaa 1260

aaccataatg atgggaaatg gaaatcattc actctgtcaa gaattcaaaa acaagccaat 1320

atatcaattg tagatataga tactaatcag gaggagaata tagcaacttc gtcttctgga 1380

aacaactttg gtcttgactt gaaagcagat gacaaaatat attttggtgg cctgccaacg 1440

ctgagaaact tgagtatgaa agcaaggcca gaagtaaatc tgaagaaata ttccggctgc 1500

ctcaaagata ttgaaatttc aagaactccg tacaatatac tcagtagtcc cgattatgtt 1560

ggtgttacca aaggatgttc cctggagaat gtttacacag ttagctttcc taagcctggt 1620

tttgtggagc tctcccctgt gccaattgat gtaggaacag aaatcaacct gtcattcagc 1680

accaagaatg agtccggcat cattcttttg ggaagtggag ggacaccagc accacctagg 1740

agaaaacgaa ggcagactgg acaggcctat tatgtaatac tcctcaacag gggccgtctg 1800

gaagtgcatc tctccacagg ggcacgaaca atgaggaaaa ttgtcatcag accagagccg 1860

aatctgtttc atgatggaag agaacattcc gttcatgtag agcgaactag aggcatcttt 1920

acagttcaag tggatgaaaa cagaagatac atgcaaaacc tgacagttga acagcctatc 1980

gaagttaaaa agcttttcgt tgggggtgct ccacctgaat ttcaaccttc cccactcaga 2040

aatattcctc cttttgaagg ctgcatatgg aatcttgtta ttaactctgt ccccatggac 2100

tttgcaaggc ctgtgtcctt caaaaatgct gacattggtc gctgtgccca tcagaaactc 2160

cgtgaagatg aagatggagc agctccagct gaaatagtta tccagcctga gccagttccc 2220

accccagcct ttcctacgcc caccccagtt ctgacacatg gtccttgtgc tgcagaatca 2280

gaaccagctc ttttgatagg gagcaagcag ttcgggcttt caagaaacag tcacattgca 2340

attgcatttg atgacaccaa agttaaaaac cgtctcacaa ttgagttgga agtaagaacc 2400

gaagctgaat ccggcttgct tttttacatg gctcgcatca atcatgctga ttttgcaaca 2460

gttcagctga gaaatggatt gccctacttc agctatgact tggggagtgg ggacacccac 2520

accatgatcc ccaccaaaat caatgatggc cagtggcaca agattaagat aatgagaagt 2580

aagcaagaag gaattcttta tgtagatggg gcttccaaca gaaccatcag tcccaaaaaa 2640

gccgacatcc tggatgtcgt gggaatgctg tatgttggtg ggttacccat caactacact 2700

acccgaagaa ttggtccagt gacctatagc attgatggct gcgtcaggaa tctccacatg 2760

gcagaggccc ctgccgatct ggaacaaccc acctccagct tccatgttgg gacatgtttt 2820

gcaaatgctc agaggggaac atattttgac ggaaccggtt ttgccaaagc agttggtgga 2880

ttcaaagtgg gattggacct tcttgtagaa tttgaattcc gcacaactac aacgactgga 2940

gttcttctgg ggatcagtag tcaaaaaatg gatggaatgg gtattgaaat gattgatgaa 3000

aagttgatgt ttcatgtgga caatggtgcg ggcagattca ctgctgtcta tgatgctggg 3060

gttccagggc atttgtgtga tggacaatgg cataaagtca ctgccaacaa gatcaaacac 3120

cgcattgagc tcacagtcga tgggaaccag gtggaagccc aaagcccaaa cccagcatct 3180

acatcagctg acacaaatga ccctgtgttt gttggaggct tcccagatga cctcaagcag 3240

tttggcctaa caaccagtat tccgttccga ggttgcatca gatccctgaa gctcaccaaa 3300

ggcacagcaa gccactggag gttaattttg ccaaggccct ggaactga 3348

<210> 4
<211> 8396
<212> DNA
<213> rAAV

<220>
<221> misc_feature
<222> (3590)..(3595)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (8336)..(8341)
<223> n is a, c, g, t or u

<400> 4
gctcttccgc ttggtcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 60

tatcagctca ctcaaacccg gtaatacggt tatccacaga atcaggggat aacgcaggaa 120

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 180

cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga 240

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 300

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 360

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc 420

gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 480

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca 540

ctggtaacag gattagcaga gcgaggtatg tacgcggtgc tacagagttc ttgaagtggt 600

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 660

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 720

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 780

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 840

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 900

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 960

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg 1020

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac 1080

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 1140

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 1200

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta 1260

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 1320

gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc 1380

ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac 1440

tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 1500

caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa 1560

tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 1620

cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca 1680

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 1740

aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac 1800

tcatactctt ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 1860

gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc 1920

gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata 1980

ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 2040

acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 2100

cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 2160

cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa 2220

ggagaaaata ccgcatcagg aacttccaac atccaataaa tcatacaggc aaggcaaaga 2280

attagcaaaa ttaagcaata aagcctcaga gcataaagct aaatcggttg taccaaaaac 2340

attatgaccc tgtaatactt ttgcgggaga agcctttatt tcaacgcaag gataaaaatt 2400

tttagaaccc tcatatattt taaatgcaat gcctgagtaa tgtgtaggta aagattcaaa 2460

cgggtgagaa aggccggaga cagtcaaatc accatcaata tgatattcaa ccgttctagc 2520

tgataaattc atgccggaga gggtagctat ttttgagagg tctctacaaa ggctatcagg 2580

tcattgcctg agagtctgga gcaaacaaga gaatcgatga acggtaatcg taaaactagc 2640

atgtcaatca tatgtacccc ggttgataat cagaaaagcc ccaaaaacag gaagattgta 2700

taagcaaata tttaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt 2760

taaatcagct catttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa 2820

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 2880

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 2940

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaatcact aaatcggaac 3000

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 3060

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 3120

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgta ctatggttgc 3180

tttgacgagc acgtataacg tgctttcctc gttagaatca gagcgggagc taaacaggag 3240

gccgattaaa gggattttag acaggaacgg tacgccagaa tcctgagaag tgtttttata 3300

atcagtgagg ccaccgagta aaagagtctg tccatcacgc aaattaaccg ttgtcgcaat 3360

acttctttga ttagtaataa catcacttgc ctgagtagaa gaactcaaac tatcggcctt 3420

gctggtaata tccagaacaa tattaccgcc agccattgca acaggaaaaa cgctcatgga 3480

aatacctaca ttttgacgct caatcgtctg gaacttccat tcgccattca ggctgcgcaa 3540

ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgn nnnnngcgcg 3600

ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc 3660

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 3720

ccttgtagtt aatgattaac ccgccatgct aattatctac gtagccatgt ctagggtcgt 3780

tacataactt acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac 3840

gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg 3900

ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaag 3960

tacgcccccc attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtacat 4020

gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat 4080

ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt 4140

tccaagtctc caccccattg acgtcaatgg gagtttgttt tggcaccaaa atcaacggga 4200

ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg 4260

gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcgcct ggagacgcca 4320

tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc ggactctaga 4380

ggatccggta ctcgaggaac tgaaaaacca gaaagttaac tggtaagttt agtctttttg 4440

tcttttattt caggtcccgg atccggtggt ggtgcaaatc aaagaactgc tcctcagtgg 4500

atgttgcctt tacttctagg cctgtacgga agtgttactt ctgctctaaa agctgcggaa 4560

ttgtacccgc ggccgcacca tgaagctgct gccgtcggtg gtgctgaagc tctttctggc 4620

tgcagttctc tcggcactgg tgactggcga gagcctggag cggcttcgga gagggctagc 4680

tgctggaacc agcaacccgg accctcccac tgtatccacg gaccagctgc taccctagg 4740

aggcggccgg gaccggaaag tccgtgactt gcaagaggca gatctggacc ttttgagagt 4800

cactttatcc tccaagccac aagcactggc cacaccaaac aaggaggagc acgggaaaag 4860

aaagaagaaa ggcaaggggc tagggaagaa gagggaccca tgtcttcgga aatacaagga 4920

cttctgcatc catggagaat gcaaatatgt gaaggagctc cgggctccct cctgcatctg 4980

ccacccgggt taccatggag agaggtgtca tgggctgagc ctcccaaaag tatctgtgtc 5040

ttcaggaggt gactgcattc gaacatacaa accagaaatc aagaaaggaa gttacaataa 5100

tattgttgtc aacgtaaaga cagctgttgc tgataacctc ctcttttatc ttggaagtgc 5160

caaatttatt gactttctgg ctatagaaat gcgtaaaggc aaagtcagct tcctctggga 5220

tgttggatct ggagttggac gtgtagagta cccagatttg actattgatg actcatattg 5280

gtaccgtatc gtagcatcaa gaactgggag aaatggaact atttctgtga gagccctgga 5340

tggacccaaa gccagcattg tgcccagcac acaccattcg acgtctcctc cagggtacac 5400

gattctagat gtggatgcaa atgcaatgct gtttgttggt ggcctgactg ggaaattaaa 5460

gaaggctgat gctgtacgtg tgattacatt cactggctgc atgggagaaa catactttga 5520

caacaaacct ataggtttgt ggaatttccg agaaaaagaa ggtgactgca aaggatgcac 5580

tgtcagtcct caggtggaag atagtgaggg gactattcaa tttgatggag aaggttatgc 5640

attggtcagc cgtcccattc gctggtaccc caacatctcc actgtcatgt tcaagttcag 5700

aacattttct tcgagtgctc ttctgatgta tcttgccaca cgagacctga gagatttcat 5760

gagtgtggag ctcactgatg ggcacataaa agtcagttac gatctgggct caggaatggc 5820

ttccgttgtc agcaatcaaa accataatga tgggaaatgg aaatcattca ctctgtcaag 5880

aattcaaaaa caagccaata tatcaattgt agatatagat actaatcagg aggagaatat 5940

agcaacttcg tcttctggaa acaactttgg tcttgacttg aaagcagatg acaaaatata 6000

ttttggtggc ctgccaacgc tgagaaactt gagtatgaaa gcaaggccag aagtaaatct 6060

gaagaaatat tccggctgcc tcaaagatat tgaaatttca agaactccgt acaatatact 6120

cagtagtccc gattatgttg gtgttaccaa aggatgttcc ctggagaatg tttacacagt 6180

tagctttcct aagcctggtt ttgtggagct ctcccctgtg ccaattgatg taggaacaga 6240

aatcaacctg tcattcagca ccaagaatga gtccggcatc attcttttgg gaagtggagg 6300

gacaccagca ccacctagga gaaaacgaag gcagactgga caggcctatt atgtaatact 6360

cctcaacagg ggccgtctgg aagtgcatct ctccacaggg gcacgaacaa tgaggaaaat 6420

tgtcatcaga ccagagccga atctgtttca tgatggaaga gaacattccg ttcatgtaga 6480

gcgaactaga ggcatcttta cagttcaagt ggatgaaaac agaagataca tgcaaaacct 6540

gacagttgaa cagcctatcg aagttaaaaa gcttttcgtt gggggtgctc cacctgaatt 6600

tcaaccttcc ccactcagaa atattcctcc ttttgaaggc tgcatatgga atcttgttat 6660

taactctgtc cccatggact ttgcaaggcc tgtgtccttc aaaaatgctg acattggtcg 6720

ctgtgcccat cagaaactcc gtgaagatga agatggagca gctccagctg aaatagttat 6780

ccagcctgag ccagttccca ccccagcctt tcctacgccc accccagttc tgacacatgg 6840

tccttgtgct gcagaatcag aaccagctct tttgataggg agcaagcagt tcgggctttc 6900

aagaaacagt cacattgcaa ttgcatttga tgacaccaaa gttaaaaacc gtctcacaat 6960

tgagttggaa gtaagaaccg aagctgaatc cggcttgctt ttttacatgg ctcgcatcaa 7020

tcatgctgat tttgcaacag ttcagctgag aaatggattg ccctacttca gctatgactt 7080

ggggagtggg gacacccaca ccatgatccc caccaaaatc aatgatggcc agtggcacaa 7140

gattaagata atgagaagta agcaagaagg aattctttat gtagatgggg cttccaacag 7200

aaccatcagt cccaaaaaag ccgacatcct ggatgtcgtg ggaatgctgt atgttggtgg 7260

gttacccatc aactacacta cccgaagaat tggtccagtg acctatagca ttgatggctg 7320

cgtcaggaat ctccacatgg cagaggcccc tgccgatctg gaacaaccca cctccagctt 7380

ccatgttggg acatgttttg caaatgctca gaggggaaca tattttgacg gaaccggttt 7440

tgccaaagca gttggtggat tcaaagtggg attggacctt cttgtagaat ttgaattccg 7500

cacaactaca acgactggag ttcttctggg gatcagtagt caaaaaatgg atggaatggg 7560

tattgaaatg attgatgaaa agttgatgtt tcatgtggac aatggtgcgg gcagattcac 7620

tgctgtctat gatgctgggg ttccagggca tttgtgtgat ggacaatggc ataaagtcac 7680

tgccaacaag atcaaacacc gcattgagct cacagtcgat gggaaccagg tggaagccca 7740

aagcccaaac ccagcatcta catcagctga cacaaatgac cctgtgtttg ttggaggctt 7800

cccagatgac ctcaagcagt ttggcctaac aaccagtatt ccgttccgag gttgcatcag 7860

atccctgaag ctcaccaaag gcacagcaag ccactggagg ttaattttgc caaggccctg 7920

gaactgaact agtgcggccg cggggatcca gacatgataa gatacattga tgagtttgga 7980

caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt gtgaaatttg tgatgctatt 8040

gctttatttg taaccattat aagctgcaat aaacaagtta acaacaacaa ttgcattcat 8100

tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt cggatcctct agagtcgacc 8160

acatggctac gtagataatt agcatggcgg gttaatcatt aactacaagg aacccctagt 8220

gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaaa 8280

ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag cgcgcnnnnn 8340

ncagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggc 8396

<210> 5
<211> 1917
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 5
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc acagactgga caggcctatt atgtaatact cctcaacagg 360

ggccgtctgg aagtgcatct ctccacaggg gcacgaacaa tgaggaaaat tgtcatcaga 420

ccagagccga atctgtttca tgatggaaga gaacattccg ttcatgtaga gcgaactaga 480

ggcatcttta cagttcaagt ggatgaaaac agaagataca tgcaaaacct gacagttgaa 540

cagcctatcg aagttaaaaa gcttttcgtt gggggtgctc cacctgaatt tcaaccttcc 600

ccactcagaa atattcctcc ttttgaaggc tgcatatgga atcttgttat taactctgtc 660

cccatggact ttgcaaggcc tgtgtccttc aaaaatgctg acattggtcg ctgtgcccat 720

cagaaactcc gtgaagatga agatggagca gctccagctg aaatagttat ccagcctgag 780

ccagttccca ccccagcctt tcctacgccc accccagttc tgacacatgg tccttgtgct 840

gcagaatcag aaccagctct tttgataggg agcaagcagt tcgggctttc aagaaacagt 900

cacattgcaa ttgcatttga tgacaccaaa gttaaaaacc gtctcacaat tgagttggaa 960

gtaagaaccg aagctgaatc cggcttgctt ttttacatgg ctcgcatcaa tcatgctgat 1020

tttgcaacag ttcagctgag aaatggattg ccctacttca gctatgactt ggggagtggg 1080

gacacccaca ccatgatccc caccaaaatc aatgatggcc agtggcacaa gattaagata 1140

atgagaagta agcaagaagg aattctttat gtagatgggg cttccaacag aaccatcagt 1200

cccaaaaaag ccgacatcct ggatgtcgtg ggaatgctgt atgttggtgg gttacccatc 1260

aactacacta cccgaagaat tggtccagtg acctatagca ttgatggctg cgtcaggaat 1320

ctccacatgg cagaggcccc tgccgatctg gaacaaccca cctccagctt ccatgttggg 1380

acatgttttg caaatgctca gaggggaaca tattttgacg gaaccggttt tgccaaagca 1440

gttggtggat tcaaagtggg attggacctt cttgtagaat ttgaattccg cacaactaca 1500

acgactggag ttcttctggg gatcagtagt caaaaaatgg atggaatggg tattgaaatg 1560

attgatgaaa agttgatgtt tcatgtggac aatggtgcgg gcagattcac tgctgtctat 1620

gatgctgggg ttccagggca tttgtgtgat ggacaatggc ataaagtcac tgccaacaag 1680

atcaaacacc gcattgagct cacagtcgat gggaaccagg tggaagccca aagcccaaac 1740

ccagcatcta catcagctga cacaaatgac cctgtgtttg ttggaggctt cccagatgac 1800

ctcaagcagt ttggcctaac aaccagtatt ccgttccgag gttgcatcag atccctgaag 1860

ctcaccaaag gcacagcaag ccactggagg ttaattttgc caaggccctg gaactga 1917

<210> 6
<211> 6984
<212> DNA
<213> rAAV

<220>
<221> misc_feature
<222> (3609)..(3614)
<223> n is a, c, g, t or u

<220>
<221> misc_feature

<222> (6924)..(6929)
<223> n is a, c, g, t or u

<400> 6
gctcttccgc ttggtcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 60

tatcagctca ctcaaacccg gtaatacggt tatccacaga atcaggggat aacgcaggaa 120

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 180

cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga 240

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 300

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 360

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc 420

gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 480

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca 540

ctggtaacag gattagcaga gcgaggtatg tacgcggtgc tacagagttc ttgaagtggt 600

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 660

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 720

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 780

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 840

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 900

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 960

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg 1020

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac 1080

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 1140

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 1200

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta 1260

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 1320

gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc 1380

ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac 1440

tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 1500

caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa 1560

tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 1620

cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca 1680

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 1740

aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac 1800

tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 1860

gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc 1920

gaaaagtgcc acctgacgtc taagaaacca tcgttacata acttacggta tattatcatg 1980

acattaacct ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg tttcggtgat 2040

gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg tctgtaagcg 2100

gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg gtgtcggggc 2160

tggcttaact atgcggcatc agagcagatt gtactgagag tgcaccatat gcggtgtgaa 2220

ataccgcaca gatgcgtaag gagaaaatac cgcatcagga acttccaaca tccaataaat 2280

catacaggca aggcaaagaa ttagcaaaat taagcaataa agcctcagag cataaagcta 2340

aatcggttgt accaaaaaca ttatgaccct gtaatacttt tgcgggagaa gcctttattt 2400

caacgcaagg ataaaaattt ttagaaccct catatatttt aaatgcaatg cctgagtaat 2460

gtgtaggtaa agattcaaac gggtgagaaa ggccggagac agtcaaatca ccatcaatat 2520

gatattcaac cgttctagct gataaattca tgccggagag ggtagctatt tttgagaggt 2580

ctctacaaag gctatcaggt cattgcctga gagtctggag caaacaagag aatcgatgaa 2640

cggtaatcgt aaaactagca tgtcaatcat atgtaccccg gttgataatc agaaaagccc 2700

caaaaacagg aagattgtat aagcaaatat ttaaattgta aacgttaata ttttgttaaa 2760

attcgcgtta aatttttgtt aaatcagctc atttttttaac caataggccg aaatcggcaa 2820

aatcccttat aaatcaaaag aatagaccga gatagggttg agtgttgttc cagtttggaa 2880

caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca 2940

gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt cgaggtgccg 3000

taaatcacta aatcggaacc ctaaagggag cccccgattt agagcttgac ggggaaagcc 3060

ggcgaacgtg gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta gggcgctggc 3120

aagtgtagcg gtcacgctgc gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca 3180

gggcgcgtac tatggttgct ttgacgagca cgtataacgt gctttcctcg ttagaatcag 3240

agcgggagct aaacaggagg ccgattaaag ggattttaga caggaacggt acgccagaat 3300

cctgagaagt gtttttataa tcagtgaggc caccgagtaa aagagtctgt ccatcacgca 3360

aattaaccgt tgtcgcaata cttctttgat tagtaataac atcacttgcc tgagtagaag 3420

aactcaaact atcggccttg ctggtaatat ccagaacaat attaccgcca gccattgcaa 3480

caggaaaaac gctcatggaa atacctacat tttgacgctc aatcgtctgg aacttccatt 3540

cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct tcgctattac 3600

gccagctgnn nnnngcgcgc tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt 3660

cgggcgacct ttggtcgccc ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc 3720

aactccatca ctaggggttc cttgtagtta atgattaacc cgccatgcta attatctacg 3780

tagccatgtc tagggtcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca 3840

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 3900

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 3960

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct 4020

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 4080

tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 4140

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 4200

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa 4260

tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctcgttta gtgaaccgtc 4320

agatcgcctg gagacgccat ccacgctgtt ttgacctcca tagaagacac cgggaccgat 4380

ccagcctccg gactctagag gatccggtac tcgaggaact gaaaaaccag aaagttaact 4440

ggtaagttta gtctttttgt cttttatttc aggtcccgga tccggtggtg gtgcaaatca 4500

aagaactgct cctcagtgga tgttgccttt acttctaggc ctgtacggaa gtgttacttc 4560

tgctctaaaa gctgcggaat tgtacccgcg gccgcaccat gaagctgctg ccgtcggtgg 4620

tgctgaagct ctttctggct gcagttctct cggcactggt gactggcgag agcctggagc 4680

ggcttcggag agggctagct gctggaacca gcaacccgga ccctcccact gtatccacgg 4740

accagctgct acccctagga ggcggccggg accggaaagt ccgtgacttg caagaggcag 4800

atctggacct tttgagagtc actttatcct ccaagccaca agcactggcc acaccaaaca 4860

aggaggagca cgggaaaaga aagaagaaag gcaaggggct agggaagaag agggacccac 4920

agactggaca ggcctattat gtaatactcc tcaacagggg ccgtctggaa gtgcatctct 4980

ccacaggggc acgaacaatg aggaaaattg tcatcagacc agagccgaat ctgtttcatg 5040

atggaagaga acattccgtt catgtagagc gaactagagg catctttaca gttcaagtgg 5100

atgaaaacag aagatacatg caaaacctga cagttgaaca gcctatcgaa gttaaaaagc 5160

ttttcgttgg gggtgctcca cctgaatttc aaccttcccc actcagaaat attcctcctt 5220

ttgaaggctg catatggaat cttgttatta actctgtccc catggacttt gcaaggcctg 5280

tgtccttcaa aaatgctgac attggtcgct gtgcccatca gaaactccgt gaagatgaag 5340

atggagcagc tccagctgaa atagttatcc agcctgagcc agttcccacc ccagcctttc 5400

ctacgcccac cccagttctg acacatggtc cttgtgctgc agaatcagaa ccagctcttt 5460

tgatagggag caagcagttc gggctttcaa gaaacagtca cattgcaatt gcatttgatg 5520

acaccaaagt taaaaaccgt ctcacaattg agttggaagt aagaaccgaa gctgaatccg 5580

gcttgctttt ttacatggct cgcatcaatc atgctgattt tgcaacagtt cagctgagaa 5640

atggattgcc ctacttcagc tatgacttgg ggagtgggga cacccacacc atgatcccca 5700

ccaaaatcaa tgatggccag tggcacaaga ttaagataat gagaagtaag caagaaggaa 5760

ttctttatgt agatggggct tccaacagaa ccatcagtcc caaaaaagcc gacatcctgg 5820

atgtcgtggg aatgctgtat gttggtgggt tacccatcaa ctacactacc cgaagaattg 5880

gtccagtgac ctatagcatt gatggctgcg tcaggaatct ccacatggca gaggcccctg 5940

ccgatctgga acaacccacc tccagcttcc atgttgggac atgttttgca aatgctcaga 6000

ggggaacata ttttgacgga accggttttg ccaaagcagt tggtggattc aaagtgggat 6060

tggaccttct tgtagaattt gaattccgca caactacaac gactggagtt cttctgggga 6120

tcagtagtca aaaaatggat ggaatgggta ttgaaatgat tgatgaaaag ttgatgtttc 6180

atgtggacaa tggtgcgggc agattcactg ctgtctatga tgctggggtt ccagggcatt 6240

tgtgtgatgg acaatggcat aaagtcactg ccaacaagat caaacaccgc attgagctca 6300

cagtcgatgg gaaccaggtg gaagcccaaa gcccaaaccc agcatctaca tcagctgaca 6360

caaatgaccc tgtgtttgtt ggaggcttcc cagatgacct caagcagttt ggcctaacaa 6420

ccagtattcc gttccgaggt tgcatcagat ccctgaagct caccaaaggc acagcaagcc 6480

actggaggtt aattttgcca aggccctgga actgaactag tgcggccgcg gggatccaga 6540

catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg 6600

ctttatttgt gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa 6660

acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga 6720

ggtttttttcg gatcctctag agtcgaccac atggctacgt agataattag catggcgggt 6780

taatcattaa ctacaaggaa cccctagtga tggagttggc cactccctct ctgcgcgctc 6840

gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg cccgggcttt gcccgggcgg 6900

cctcagtgag cgagcgagcg cgcnnnnnnc agctgcatta atgaatcggc caacgcgcgg 6960

ggagaggcgg tttgcgtatt gggc 6984

<210> 7
<211> 2043
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 7
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc atgtcttcgg aaatacaagg acttctgcat ccatggagaa 360

tgcaaatatg tgaaggagct ccgggctccc tcctgcatct gccacccggg ttaccatgga 420

gagaggtgtc atgggctgag cctcccacag actggacagg cctattatgt aatactcctc 480

aacaggggcc gtctggaagt gcatctctcc acaggggcac gaacaatgag gaaaattgtc 540

atcagaccag agccgaatct gtttcatgat ggaagagaac attccgttca tgtagagcga 600

actagaggca tctttacagt tcaagtggat gaaaacagaa gatacatgca aaacctgaca 660

gttgaacagc ctatcgaagt taaaaagctt ttcgttgggg gtgctccacc tgaatttcaa 720

ccttccccac tcagaaatat tcctcctttt gaaggctgca tatggaatct tgttattaac 780

tctgtcccca tggactttgc aaggcctgtg tccttcaaaa atgctgacat tggtcgctgt 840

gcccatcaga aactccgtga agatgaagat ggagcagctc cagctgaaat agttatccag 900

cctgagccag ttcccacccc agcctttcct acgcccaccc cagttctgac acatggtcct 960

tgtgctgcag aatcagaacc agctcttttg atagggagca agcagttcgg gctttcaaga 1020

aacagtcaca ttgcaattgc atttgatgac accaaagtta aaaaccgtct cacaattgag 1080

ttggaagtaa gaaccgaagc tgaatccggc ttgctttttt acatggctcg catcaatcat 1140

gctgattttg caacagttca gctgagaaat ggattgccct acttcagcta tgacttgggg 1200

agtggggaca cccacaccat gatccccacc aaaatcaatg atggccagtg gcacaagatt 1260

aagataatga gaagtaagca agaaggaatt ctttatgtag atggggcttc caacagaacc 1320

atcagtccca aaaaagccga catcctggat gtcgtgggaa tgctgtatgt tggtgggtta 1380

cccatcaact acactacccg aagaattggt ccagtgacct atagcattga tggctgcgtc 1440

aggaatctcc acatggcaga ggcccctgcc gatctggaac aacccacctc cagcttccat 1500

gttgggacat gttttgcaaa tgctcagagg ggaacatatt ttgacggaac cggttttgcc 1560

aaagcagttg gtggattcaa agtgggattg gaccttcttg tagaatttga attccgcaca 1620

actacaacga ctggagttct tctggggatc agtagtcaaa aaatggatgg aatgggtatt 1680

gaaatgattg atgaaaagtt gatgtttcat gtggacaatg gtgcgggcag attcactgct 1740

gtctatgatg ctggggttcc agggcatttg tgtgatggac aatggcataa agtcactgcc 1800

aacaagatca aacaccgcat tgagctcaca gtcgatggga accaggtgga agcccaaagc 1860

ccaaacccag catctacatc agctgacaca aatgaccctg tgtttgttgg aggcttccca 1920

gatgacctca agcagtttgg cctaacaacc agtattccgt tccgaggttg catcagatcc 1980

ctgaagctca ccaaaggcac agcaagccac tggaggttaa ttttgccaag gccctggaac 2040

tga 2043

<210> 8
<211> 7091
<212> DNA
<213> rAAV

<220>
<221> misc_feature
<222> (3590)..(3595)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (7031)..(7036)
<223> n is a, c, g, t or u

<400> 8
gctcttccgc ttggtcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 60

tatcagctca ctcaaacccg gtaatacggt tatccacaga atcaggggat aacgcaggaa 120

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 180

cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga 240

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 300

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 360

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc 420

gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 480

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca 540

ctggtaacag gattagcaga gcgaggtatg tacgcggtgc tacagagttc ttgaagtggt 600

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 660

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 720

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 780

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 840

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 900

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 960

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg 1020

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac 1080

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 1140

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 1200

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta 1260

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 1320

gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc 1380

ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac 1440

tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 1500

caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa 1560

tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 1620

cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca 1680

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 1740

aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac 1800

tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 1860

gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc 1920

gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata 1980

ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 2040

acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 2100

cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 2160

cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa 2220

ggagaaaata ccgcatcagg aacttccaac atccaataaa tcatacaggc aaggcaaaga 2280

attagcaaaa ttaagcaata aagcctcaga gcataaagct aaatcggttg taccaaaaac 2340

attatgaccc tgtaatactt ttgcgggaga agcctttatt tcaacgcaag gataaaaatt 2400

tttagaaccc tcatatattt taaatgcaat gcctgagtaa tgtgtaggta aagattcaaa 2460

cgggtgagaa aggccggaga cagtcaaatc accatcaata tgatattcaa ccgttctagc 2520

tgataaattc atgccggaga gggtagctat ttttgagagg tctctacaaa ggctatcagg 2580

tcattgcctg agagtctgga gcaaacaaga gaatcgatga acggtaatcg taaaactagc 2640

atgtcaatca tatgtacccc ggttgataat cagaaaagcc ccaaaaacag gaagattgta 2700

taagcaaata tttaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt 2760

taaatcagct cattttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa 2820

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 2880

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 2940

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaatcact aaatcggaac 3000

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 3060

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 3120

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgta ctatggttgc 3180

tttgacgagc acgtataacg tgctttcctc gttagaatca gagcgggagc taaacaggag 3240

gccgattaaa gggattttag acaggaacgg tacgccagaa tcctgagaag tgtttttata 3300

atcagtgagg ccaccgagta aaagagtctg tccatcacgc aaattaaccg ttgtcgcaat 3360

acttctttga ttagtaataa catcacttgc ctgagtagaa gaactcaaac tatcggcctt 3420

gctggtaata tccagaacaa tattaccgcc agccattgca acaggaaaaa cgctcatgga 3480

aatacctaca ttttgacgct caatcgtctg gaacttccat tcgccattca ggctgcgcaa 3540

ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgn nnnnngcgcg 3600

ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc 3660

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 3720

ccttgtagtt aatgattaac ccgccatgct aattatctac gtagccatgt ctagggtcgt 3780

tacataactt acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac 3840

gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg 3900

ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaag 3960

tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtacat 4020

gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat 4080

ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt 4140

tccaagtctc cacccattg acgtcaatgg gagtttgttt tggcaccaaa atcaacggga 4200

ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg 4260

gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcgcct ggagacgcca 4320

tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc ggactctaga 4380

ggatccggta ctcgaggaac tgaaaaacca gaaagttaac tggtaagttt agtctttttg 4440

tcttttattt caggtcccgg atccggtggt ggtgcaaatc aaagaactgc tcctcagtgg 4500

atgttgcctt tacttctagg cctgtacgga agtgttactt ctgctctaaa agctgcggaa 4560

ttgtacccgc ggccgcacca tgaagctgct gccgtcggtg gtgctgaagc tctttctggc 4620

tgcagttctc tcggcactgg tgactggcga gagcctggag cggcttcgga gagggctagc 4680

tgctggaacc agcaacccgg accctcccac tgtatccacg gaccagctgc tacccctagg 4740

aggcggccgg gaccggaaag tccgtgactt gcaagaggca gatctggacc ttttgagagt 4800

cactttatcc tccaagccac aagcactggc cacaccaaac aaggaggagc acgggaaaag 4860

aaagaagaaa ggcaaggggc tagggaagaa gagggaccca tgtcttcgga aatacaagga 4920

cttctgcatc catggagaat gcaaatatgt gaaggagctc cgggctccct cctgcatctg 4980

ccacccgggt taccatggag agaggtgtca tgggctgagc ctcccacaga ctggacaggc 5040

ctattatgta atactcctca acaggggccg tctggaagtg catctctcca caggggcacg 5100

aacaatgagg aaaattgtca tcagaccaga gccgaatctg tttcatgatg gaagagaaca 5160

ttccgttcat gtagagcgaa ctagaggcat ctttacagtt caagtggatg aaaacagaag 5220

atacatgcaa aacctgacag ttgaacagcc tatcgaagtt aaaaagcttt tcgttggggg 5280

tgctccacct gaatttcaac cttccccact cagaaatatt cctccttttg aaggctgcat 5340

atggaatctt gttattaact ctgtccccat ggactttgca aggcctgtgt ccttcaaaaa 5400

tgctgacatt ggtcgctgtg cccatcagaa actccgtgaa gatgaagatg gagcagctcc 5460

agctgaaata gttatccagc ctgagccagt tcccacccca gcctttccta cgcccacccc 5520

agttctgaca catggtcctt gtgctgcaga atcagaacca gctcttttga tagggagcaa 5580

gcagttcggg ctttcaagaa acagtcacat tgcaattgca tttgatgaca ccaaagttaa 5640

aaaccgtctc acaattgagt tggaagtaag aaccgaagct gaatccggct tgctttttta 5700

catggctcgc atcaatcatg ctgattttgc aacagttcag ctgagaaatg gattgcccta 5760

cttcagctat gacttgggga gtggggacac ccacaccatg atccccacca aaatcaatga 5820

tggccagtgg cacaagatta agataatgag aagtaagcaa gaaggaattc tttatgtaga 5880

tggggcttcc aacagaacca tcagtcccaa aaaagccgac atcctggatg tcgtgggaat 5940

gctgtatgtt ggtgggttac ccatcaacta cactacccga agaattggtc cagtgaccta 6000

tagcattgat ggctgcgtca ggaatctcca catggcagag gcccctgccg atctggaaca 6060

acccacctcc agcttccatg ttgggacatg ttttgcaaat gctcagaggg gaacatattt 6120

tgacggaacc ggttttgcca aagcagttgg tggattcaaa gtgggattgg accttcttgt 6180

agaatttgaa ttccgcacaa ctacaacgac tggagttctt ctggggatca gtagtcaaaa 6240

aatggatgga atgggtattg aaatgattga tgaaaagttg atgtttcatg tggacaatgg 6300

tgcgggcaga ttcactgctg tctatgatgc tggggttcca gggcatttgt gtgatggaca 6360

atggcataaa gtcactgcca acaagatcaa acaccgcatt gagctcacag tcgatgggaa 6420

ccaggtggaa gcccaaagcc caaacccagc atctacatca gctgacacaa atgaccctgt 6480

gtttgttgga ggcttcccag atgacctcaa gcagtttggc ctaacaacca gtattccgtt 6540

ccgaggttgc atcagatccc tgaagctcac caaaggcaca gcaagccact ggaggttaat 6600

tttgccaagg ccctggaact gaactagtgc ggccgcgggg atccagacat gataagatac 6660

attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa 6720

atttgtgatg ctattgcttt atttgtaacc attataagct gcaataaaca agttaacaac 6780

aacaattgca ttcattttat gtttcaggtt cagggggagg tgtgggaggt tttttcggat 6840

cctctagagt cgaccacatg gctacgtaga taattagcat ggcgggttaa tcattaacta 6900

caaggaaccc ctagtgatgg agttggccac tccctctctg cgcgctcgct cgctcactga 6960

ggccgggcga ccaaaggtcg cccgacgccc gggctttgcc cgggcggcct cagtgagcga 7020

gcgagcgcgc nnnnnncagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt 7080

gcgtattggg c 7091

<210> 9
<211> 1347
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 9
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc acagatccat gctacaccca cacctgtcac tgccattggg 360

cccccaacca cggctatcca ggagccccca tccaggatcg tgccaacccc cacatctcca 420

gccattgctc ctccaacaga gaccatggct cctccagtca gggatcctgt tcctgggaaa 480

cccacggtca ccatccggac tcgaggcgcc attattcaaa ccccaaccct aggccccatc 540

cagcctactc gggtgtcaga agctggcacc acagttcctg gccagattcg cccaacgatg 600

accattcctg gctatgtgga gcctactgca gttgctaccc ctcccacaac caccaccaag 660

aagccacgag tatccacacc aaaaccagca acgccttcaa ctgactccac caccaccacg 720

actcgcaggc caaccaagaa accacggaca ccccggccag tgccccgggt caccaccaaa 780

gtttccatca ccagattgga aactgcctca ccgcctactc gtattcgcac caccaccagt 840

ggagtgcccc gtggcggaga acccaaccag cgcccagagc tcaagaacca tattgacagg 900

gtagatgcct gggttggcac ctactttgag gtgaagatcc cgtcagacac tttctatgac 960

catgaggaca ccaccactga caagctgaag ctgaccctga aactgcggga gcagcagctg 1020

gtgggcgaga agtcctgggt acagttcaac agcaacagcc agctcatgta tggccttccc 1080

gacagcagcc acgtgggcaa acacgagtat ttcatgcatg ccacagacaa ggggggcctg 1140

tcggctgtgg atgccttcga gatccacgtc cacaggcgcc cccaagggga tagggctcct 1200

gcaaggttca aggccaagtt tgtgggtgac ccggcactgg tgttgaatga catccacaag 1260

aagattgcct tggtaaagaa actggccttc gcctttggag accgaaactg tagcaccatc 1320

accctgcaga atatcacccg gggctaa 1347

<210> 10
<211> 6395
<212> DNA
<213> rAAV

<220>
<221> misc_feature
<222> (3590)..(3595)
<223> n is a, c, g, t or u

<220>
<221> misc_feature
<222> (6335)..(6340)
<223> n is a, c, g, t or u

<400> 10
gctcttccgc ttggtcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 60

tatcagctca ctcaaacccg gtaatacggt tatccacaga atcaggggat aacgcaggaa 120

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 180

cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga 240

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 300

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 360

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc 420

gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 480

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca 540

ctggtaacag gattagcaga gcgaggtatg tacgcggtgc tacagagttc ttgaagtggt 600

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 660

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 720

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 780

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 840

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 900

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 960

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg 1020

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac 1080

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 1140

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 1200

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta 1260

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 1320

gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc 1380

ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac 1440

tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 1500

caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa 1560

tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 1620

cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca 1680

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 1740

aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac 1800

tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 1860

gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc 1920

gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata 1980

ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 2040

acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 2100

cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 2160

cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa 2220

ggagaaaata ccgcatcagg aacttccaac atccaataaa tcatacaggc aaggcaaaga 2280

attagcaaaa ttaagcaata aagcctcaga gcataaagct aaatcggttg taccaaaaac 2340

attatgaccc tgtaatactt ttgcgggaga agcctttatt tcaacgcaag gataaaaatt 2400

tttagaaccc tcatatattt taaatgcaat gcctgagtaa tgtgtaggta aagattcaaa 2460

cgggtgagaa aggccggaga cagtcaaatc accatcaata tgatattcaa ccgttctagc 2520

tgataaattc atgccggaga gggtagctat ttttgagagg tctctacaaa ggctatcagg 2580

tcattgcctg agagtctgga gcaaacaaga gaatcgatga acggtaatcg taaaactagc 2640

atgtcaatca tatgtacccc ggttgataat cagaaaagcc ccaaaaacag gaagattgta 2700

taagcaaata tttaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt 2760

taaatcagct cattttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa 2820

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 2880

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 2940

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaatcact aaatcggaac 3000

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 3060

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 3120

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgta ctatggttgc 3180

tttgacgagc acgtataacg tgctttcctc gttagaatca gagcgggagc taaacaggag 3240

gccgattaaa gggattttag acaggaacgg tacgccagaa tcctgagaag tgtttttata 3300

atcagtgagg ccaccgagta aaagagtctg tccatcacgc aaattaaccg ttgtcgcaat 3360

acttctttga ttagtaataa catcacttgc ctgagtagaa gaactcaaac tatcggcctt 3420

gctggtaata tccagaacaa tattaccgcc agccattgca acaggaaaaa cgctcatgga 3480

aatacctaca ttttgacgct caatcgtctg gaacttccat tcgccattca ggctgcgcaa 3540

ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgn nnnnngcgcg 3600

ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc 3660

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 3720

ccttgtagtt aatgattaac ccgccatgct aattatctac gtagccatgt ctagggtcgt 3780

tacataactt acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac 3840

gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg 3900

ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaag 3960

tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtacat 4020

gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat 4080

ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt 4140

tccaagtctc caccccattg acgtcaatgg gagtttgttt tggcaccaaa atcaacggga 4200

ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg 4260

gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcgcct ggagacgcca 4320

tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc ggactctaga 4380

ggatccggta ctcgaggaac tgaaaaacca gaaagttaac tggtaagttt agtctttttg 4440

tcttttattt caggtcccgg atccggtggt ggtgcaaatc aaagaactgc tcctcagtgg 4500

atgttgcctt tacttctagg cctgtacgga agtgttactt ctgctctaaa agctgcggaa 4560

ttgtacccgc ggccgcacca tgaagctgct gccgtcggtg gtgctgaagc tctttctggc 4620

tgcagttctc tcggcactgg tgactggcga gagcctggag cggcttcgga gagggctagc 4680

tgctggaacc agcaacccgg accctcccac tgtatccacg gaccagctgc taccctagg 4740

aggcggccgg gaccggaaag tccgtgactt gcaagaggca gatctggacc ttttgagagt 4800

cactttatcc tccaagccac aagcactggc cacaccaaac aaggaggagc acgggaaaag 4860

aaagaagaaa ggcaaggggc tagggaagaa gagggaccca cagatccatg ctacacccac 4920

acctgtcact gccattgggc ccccaaccac ggctatccag gagcccccat ccaggatcgt 4980

gccaaccccc acatctccag ccattgctcc tccaacagag accatggctc ctccagtcag 5040

ggatcctgtt cctgggaaac ccacggtcac catccggact cgaggcgcca ttattcaaac 5100

cccaacccta ggccccatcc agcctactcg ggtgtcagaa gctggcacca cagttcctgg 5160

ccagattcgc ccaacgatga ccattcctgg ctatgtggag cctactgcag ttgctacccc 5220

tcccacaacc accaccaaga agccacgagt atccacacca aaaccagcaa cgccttcaac 5280

tgactccacc accaccacga ctcgcaggcc aaccaagaaa ccacggacac cccggccagt 5340

gccccgggtc accaccaaag tttccatcac cagattggaa actgcctcac cgcctactcg 5400

tattcgcacc accaccagtg gagtgccccg tggcggagaa cccaaccagc gcccagagct 5460

caagaaccat attgacaggg tagatgcctg ggttggcacc tactttgagg tgaagatccc 5520

gtcagacact ttctatgacc atgaggacac caccactgac aagctgaagc tgaccctgaa 5580

actgcgggag cagcagctgg tgggcgagaa gtcctgggta cagttcaaca gcaacagcca 5640

gctcatgtat ggccttcccg acagcagcca cgtgggcaaa cacgagtatt tcatgcatgc 5700

cacagacaag gggggcctgt cggctgtgga tgccttcgag atccacgtcc acaggcgccc 5760

ccaaggggat agggctcctg caaggttcaa ggccaagttt gtgggtgacc cggcactggt 5820

gttgaatgac atccacaaga agattgcctt ggtaaagaaa ctggccttcg cctttggaga 5880

ccgaaactgt agcaccatca ccctgcagaa tatcacccgg ggctaaacta gtgcggccgc 5940

ggggatccag acatgataag atacattgat gagtttggac aaaccacaac tagaatgcag 6000

tgaaaaaaat gctttatttg tgaaatttgt gatgctattg ctttatttgt aaccattata 6060

agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg 6120

gaggtgtggg aggttttttc ggatcctcta gagtcgacca catggctacg tagataatta 6180

gcatggcggg ttaatcatta actacaagga acccctagtg atggagttgg ccactccctc 6240

tctgcgcgct cgctcgctca ctgaggccgg gcgaccaaag gtcgcccgac gcccgggctt 6300

tgcccgggcg gcctcagtga gcgagcgagc gcgcnnnnnn cagctgcatt aatgaatcgg 6360

ccaacgcgcg gggagaggcg gtttgcgtat tgggc 6395

<210> 11
<211> 1473
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 11
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc atgtcttcgg aaatacaagg acttctgcat ccatggagaa 360

tgcaaatatg tgaaggagct ccgggctccc tcctgcatct gccacccggg ttaccatgga 420

gagaggtgtc atgggctgag cctcccacag atccatgcta cacccacacc tgtcactgcc 480

attgggcccc caaccacggc tatccaggag cccccatcca ggatcgtgcc aacccccaca 540

tctccagcca ttgctcctcc aacagagacc atggctcctc cagtcaggga tcctgttcct 600

gggaaaccca cggtcaccat ccggactcga ggcgccatta ttcaaacccc aaccctaggc 660

cccatccagc ctactcgggt gtcagaagct ggcaccacag ttcctggcca gattcgccca 720

acgatgacca ttcctggcta tgtggagcct actgcagttg ctacccctcc cacaaccacc 780

accaagaagc cacgagtatc cacaccaaaa ccagcaacgc cttcaactga ctccaccacc 840

accacgactc gcaggccaac caagaaacca cggacacccc ggccagtgcc ccgggtcacc 900

accaaagttt ccatcaccag attggaaact gcctcaccgc ctactcgtat tcgcaccacc 960

accagtggag tgccccgtgg cggagaaccc aaccagcgcc cagagctcaa gaaccatatt 1020

gacagggtag atgcctgggt tggcacctac tttgaggtga agatcccgtc agacactttc 1080

tatgaccatg aggacaccac cactgacaag ctgaagctga ccctgaaact gcgggagcag 1140

cagctggtgg gcgagaagtc ctgggtacag ttcaacagca acagccagct catgtatggc 1200

cttcccgaca gcagccacgt gggcaaacac gagtatttca tgcatgccac agacaagggg 1260

ggcctgtcgg ctgtggatgc cttcgagatc cacgtccaca ggcgccccca aggggatagg 1320

gctcctgcaa ggttcaaggc caagtttgtg ggtgacccgg cactggtgtt gaatgacatc 1380

cacaagaaga ttgccttggt aaagaaactg gccttcgcct ttggagaccg aaactgtagc 1440

accatcaccc tgcagaatat cacccggggc taa 1473

<210> 12
<211> 6521
<212> DNA
<213> rAAV

<220>
<221> misc_feature
<222> (3590)..(3595)
<223> n is a, c, g, t or u

<220>
<221> misc_feature

<222> (6461)..(6466)
<223> n is a, c, g, t or u

<400> 12
gctcttccgc ttggtcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 60

tatcagctca ctcaaacccg gtaatacggt tatccacaga atcaggggat aacgcaggaa 120

agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 180

cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga 240

ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 300

tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 360

gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc 420

gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 480

gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca 540

ctggtaacag gattagcaga gcgaggtatg tacgcggtgc tacagagttc ttgaagtggt 600

ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 660

ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 720

gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 780

ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 840

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 900

ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 960

gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg 1020

tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac 1080

cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 1140

ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 1200

gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta 1260

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 1320

gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc 1380

ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac 1440

tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 1500

caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa 1560

tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 1620

cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca 1680

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 1740

aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac 1800

tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 1860

gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc 1920

gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata 1980

ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 2040

acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 2100

cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 2160

cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa 2220

ggagaaaata ccgcatcagg aacttccaac atccaataaa tcatacaggc aaggcaaaga 2280

attagcaaaa ttaagcaata aagcctcaga gcataaagct aaatcggttg taccaaaaac 2340

attatgaccc tgtaatactt ttgcgggaga agcctttatt tcaacgcaag gataaaaatt 2400

tttagaaccc tcatatattt taaatgcaat gcctgagtaa tgtgtaggta aagattcaaa 2460

cgggtgagaa aggccggaga cagtcaaatc accatcaata tgatattcaa ccgttctagc 2520

tgataaattc atgccggaga gggtagctat ttttgagagg tctctacaaa ggctatcagg 2580

tcattgcctg agagtctgga gcaaacaaga gaatcgatga acggtaatcg taaaactagc 2640

atgtcaatca tatgtacccc ggttgataat cagaaaagcc ccaaaaacag gaagattgta 2700

taagcaaata tttaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt 2760

taaatcagct catttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa 2820

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 2880

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 2940

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaatcact aaatcggaac 3000

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 3060

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 3120

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgta ctatggttgc 3180

tttgacgagc acgtataacg tgctttcctc gttagaatca gagcgggagc taaacaggag 3240

gccgattaaa gggattttag acaggaacgg tacgccagaa tcctgagaag tgtttttata 3300

atcagtgagg ccaccgagta aaagagtctg tccatcacgc aaattaaccg ttgtcgcaat 3360

acttctttga ttagtaataa catcacttgc ctgagtagaa gaactcaaac tatcggcctt 3420

gctggtaata tccagaacaa tattaccgcc agccattgca acaggaaaaa cgctcatgga 3480

aatacctaca ttttgacgct caatcgtctg gaacttccat tcgccattca ggctgcgcaa 3540

ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgn nnnnngcgcg 3600

ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc 3660

cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt 3720

ccttgtagtt aatgattaac ccgccatgct aattatctac gtagccatgt ctagggtcgt 3780

tacataactt acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac 3840

gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg 3900

ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaag 3960

tacgcccccc attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtacat 4020

gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat 4080

ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt 4140

tccaagtctc caccccattg acgtcaatgg gagtttgttt tggcaccaaa atcaacggga 4200

ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg 4260

gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcgcct ggagacgcca 4320

tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc ggactctaga 4380

ggatccggta ctcgaggaac tgaaaaacca gaaagttaac tggtaagttt agtctttttg 4440

tcttttattt caggtcccgg atccggtggt ggtgcaaatc aaagaactgc tcctcagtgg 4500

atgttgcctt tacttctagg cctgtacgga agtgttactt ctgctctaaa agctgcggaa 4560

ttgtacccgc ggccgcacca tgaagctgct gccgtcggtg gtgctgaagc tctttctggc 4620

tgcagttctc tcggcactgg tgactggcga gagcctggag cggcttcgga gagggctagc 4680

tgctggaacc agcaacccgg accctcccac tgtatccacg gaccagctgc taccctagg 4740

aggcggccgg gaccggaaag tccgtgactt gcaagaggca gatctggacc ttttgagagt 4800

cactttatcc tccaagccac aagcactggc cacaccaaac aaggaggagc acgggaaaag 4860

aaagaagaaa ggcaaggggc tagggaagaa gagggaccca tgtcttcgga aatacaagga 4920

cttctgcatc catggagaat gcaaatatgt gaaggagctc cgggctccct cctgcatctg 4980

ccacccgggt taccatggag agaggtgtca tgggctgagc ctcccacaga tccatgctac 5040

acccacacct gtcactgcca ttgggccccc aaccacggct atccaggagc ccccatccag 5100

gatcgtgcca acccccacat ctccagccat tgctcctcca acagagacca tggctcctcc 5160

agtcagggat cctgttcctg ggaaacccac ggtcaccatc cggactcgag gcgccattat 5220

tcaaacccca accctaggcc ccatccagcc tactcgggtg tcagaagctg gcaccacagt 5280

tcctggccag attcgcccaa cgatgaccat tcctggctat gtggagccta ctgcagttgc 5340

taccctccc acaaccacca ccaagaagcc acgagtatcc acaccaaaac cagcaacgcc 5400

ttcaactgac tccaccacca ccacgactcg caggccaacc aagaaaccac ggacaccccg 5460

gccagtgccc cgggtcacca ccaaagtttc catcaccaga ttggaaactg cctcaccgcc 5520

tactcgtatt cgcaccacca ccagtggagt gccccgtggc ggagaaccca accagcgccc 5580

agagctcaag aaccatattg acagggtaga tgcctgggtt ggcacctact ttgaggtgaa 5640

gatcccgtca gacactttct atgaccatga ggacaccacc actgacaagc tgaagctgac 5700

cctgaaactg cgggagcagc agctggtggg cgagaagtcc tgggtacagt tcaacagcaa 5760

cagccagctc atgtatggcc ttcccgacag cagccacgtg ggcaaacacg agtatttcat 5820

gcatgccaca gacaaggggg gcctgtcggc tgtggatgcc ttcgagatcc acgtccacag 5880

gcgcccccaa ggggataggg ctcctgcaag gttcaaggcc aagtttgtgg gtgacccggc 5940

actggtgttg aatgacatcc acaagaagat tgccttggta aagaaactgg ccttcgcctt 6000

tggagaccga aactgtagca ccatcaccct gcagaatatc acccggggct aaactagtgc 6060

ggccgcgggg atccagacat gataagatac attgatgagt ttggacaaac cacaactaga 6120

atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt atttgtaacc 6180

attataagct gcaataaaca agttaacaac aacaattgca ttcattttat gtttcaggtt 6240

caggggggagg tgtgggaggt tttttcggat cctctagagt cgaccacatg gctacgtaga 6300

taattagcat ggcgggttaa tcattaacta caaggaaccc ctagtgatgg agttggccac 6360

tccctctctg cgcgctcgct cgctcactga ggccgggcga ccaaaggtcg cccgacgccc 6420

gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc nnnnnncagc tgcattaatg 6480

aatcggccaa cgcgcgggga gaggcggttt gcgtattggg c 6521

<210> 13
<211> 321
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 13
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc a 321

<210> 14
<211> 447
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 14
atgaagctgc tgccgtcggt ggtgctgaag ctctttctgg ctgcagttct ctcggcactg 60

gtgactggcg agagcctgga gcggcttcgg agagggctag ctgctggaac cagcaacccg 120

gaccctccca ctgtatccac ggaccagctg ctacccctag gaggcggccg ggaccggaaa 180

gtccgtgact tgcaagaggc agatctggac cttttgagag tcactttatc ctccaagcca 240

caagcactgg ccacaccaaa caaggaggag cacgggaaaa gaaagaagaa aggcaagggg 300

ctagggaaga agagggaccc atgtcttcgg aaatacaagg acttctgcat ccatggagaa 360

tgcaaatatg tgaaggagct ccgggctccc tcctgcatct gccacccggg ttaccatgga 420

gagaggtgtc atgggctgag cctccca 447

<210> 15
<211> 2901
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 15
aaagtatctg tgtcttcagg aggtgactgc attcgaacat acaaaccaga aatcaagaaa 60

ggaagttaca ataatattgt tgtcaacgta aagacagctg ttgctgataa cctcctcttt 120

tatcttggaa gtgccaaatt tattgacttt ctggctatag aaatgcgtaa aggcaaagtc 180

agcttcctct gggatgttgg atctggagtt ggacgtgtag agtacccaga tttgactatt 240

gatgactcat attggtaccg tatcgtagca tcaagaactg ggagaaatgg aactatttct 300

gtgagagccc tggatggacc caaagccagc attgtgccca gcacacacca ttcgacgtct 360

cctccagggt acacgattct agatgtggat gcaaatgcaa tgctgtttgt tggtggcctg 420

actgggaaat taaagaaggc tgatgctgta cgtgtgatta cattcactgg ctgcatggga 480

gaaacatact ttgacaacaa acctataggt ttgtggaatt tccgagaaaa agaaggtgac 540

tgcaaaggat gcactgtcag tcctcaggtg gaagatagtg aggggactat tcaatttgat 600

ggagaaggtt atgcattggt cagccgtccc attcgctggt accccaacat ctccactgtc 660

atgttcaagt tcagaacatt ttcttcgagt gctcttctga tgtatcttgc cacacgagac 720

ctgagagatt tcatgagtgt ggagctcact gatgggcaca taaaagtcag ttacgatctg 780

ggctcaggaa tggcttccgt tgtcagcaat caaaaccata atgatgggaa atggaaatca 840

ttcactctgt caagaattca aaaacaagcc aatatatcaa ttgtagatat agatactaat 900

caggaggaga atatagcaac ttcgtcttct ggaaacaact ttggtcttga cttgaaagca 960

gatgacaaaa tatattttgg tggcctgcca acgctgagaa acttgagtat gaaagcaagg 1020

ccagaagtaa atctgaagaa atattccggc tgcctcaaag atattgaaat ttcaagaact 1080

ccgtacaata tactcagtag tcccgattat gttggtgtta ccaaaggatg ttccctggag 1140

aatgtttaca cagttagctt tcctaagcct ggttttgtgg agctctcccc tgtgccaatt 1200

gatgtaggaa cagaaatcaa cctgtcattc agcaccaaga atgagtccgg catcattctt 1260

ttgggaagtg gagggacacc agcaccacct aggagaaaac gaaggcagac tggacaggcc 1320

tattatgtaa tactcctcaa caggggccgt ctggaagtgc atctctccac aggggcacga 1380

acaatgagga aaattgtcat cagaccagag ccgaatctgt ttcatgatgg aagagaacat 1440

tccgttcatg tagagcgaac tagaggcatc tttacagttc aagtggatga aaacagaaga 1500

tacatgcaaa acctgacagt tgaacagcct atcgaagtta aaaagctttt cgttgggggt 1560

gctccacctg aatttcaacc ttccccactc agaaatattc ctccttttga aggctgcata 1620

tggaatcttg ttattaactc tgtccccatg gactttgcaa ggcctgtgtc cttcaaaaat 1680

gctgacattg gtcgctgtgc ccatcagaaa ctccgtgaag atgaagatgg agcagctcca 1740

gctgaaatag ttatccagcc tgagccagtt cccaccccag cctttcctac gcccacccca 1800

gttctgacac atggtccttg tgctgcagaa tcagaaccag ctcttttgat agggagcaag 1860

cagttcgggc tttcaagaaa cagtcacatt gcaattgcat ttgatgacac caaagttaaa 1920

aaccgtctca caattgagtt ggaagtaaga accgaagctg aatccggctt gctttttac 1980

atggctcgca tcaatcatgc tgattttgca acagttcagc tgagaaatgg attgccctac 2040

ttcagctatg acttggggag tggggacacc cacaccatga tccccaccaa aatcaatgat 2100

ggccagtggc acaagattaa gataatgaga agtaagcaag aaggaattct ttatgtagat 2160

ggggcttcca acagaaccat cagtcccaaa aaagccgaca tcctggatgt cgtgggaatg 2220

ctgtatgttg gtgggttacc catcaactac actacccgaa gaattggtcc agtgacctat 2280

agcattgatg gctgcgtcag gaatctccac atggcagagg cccctgccga tctggaacaa 2340

cccacctcca gcttccatgt tgggacatgt tttgcaaatg ctcagagggg aacatatttt 2400

gacggaaccg gttttgccaa agcagttggt ggattcaaag tgggattgga ccttcttgta 2460

gaatttgaat tccgcacaac tacaacgact ggagttcttc tggggatcag tagtcaaaaa 2520

atggatggaa tgggtattga aatgattgat gaaaagttga tgtttcatgt ggacaatggt 2580

gcgggcagat tcactgctgt ctatgatgct ggggttccag ggcatttgtg tgatggacaa 2640

tggcataaag tcactgccaa caagatcaaa caccgcattg agctcacagt cgatgggaac 2700

caggtggaag cccaaagccc aaacccagca tctacatcag ctgacacaaa tgaccctgtg 2760

tttgttggag gcttcccaga tgacctcaag cagtttggcc taacaaccag tattccgttc 2820

cgaggttgca tcagatccct gaagctcacc aaaggcacag caagccactg gaggttaatt 2880

ttgccaaggc cctggaactg a 2901

<210> 16
<211> 1332
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 16
cagactggac aggcctatta tgtaatactc ctcaacaggg gccgtctgga agtgcatctc 60

tccacagggg cacgaacaat gaggaaaatt gtcatcagac cagagccgaa tctgtttcat 120

gatggaagag aacattccgt tcatgtagag cgaactagag gcatctttac agttcaagtg 180

gatgaaaaca gaagatacat gcaaaacctg acagttgaac agcctatcga agttaaaaag 240

cttttcgttg ggggtgctcc acctgaattt caaccttccc cactcagaaa tattcctcct 300

tttgaaggct gcatatggaa tcttgttatt aactctgtcc ccatggactt tgcaaggcct 360

gtgtccttca aaaatgctga cattggtcgc tgtgcccatc agaaactccg tgaagatgaa 420

gatggagcag ctccagctga aatagttatc cagcctgagc cagttcccac cccagccttt 480

cctacgccca ccccagttct gacacatggt ccttgtgctg cagaatcaga accagctctt 540

ttgataggga gcaagcagtt cgggctttca agaaacagtc acattgcaat tgcatttgat 600

gacaccaaag ttaaaaaccg tctcacaatt gagttggaag taagaaccga agctgaatcc 660

ggcttgcttt tttacatggc tcgcatcaat catgctgatt ttgcaacagt tcagctgaga 720

aatggattgc cctacttcag ctatgacttg gggagtgggg acacccacac catgatcccc 780

accaaaatca atgatggcca gtggcacaag attaagataa tgagaagtaa gcaagaagga 840

attctttatg tagatggggc ttccaacaga accatcagtc ccaaaaaagc cgacatcctg 900

gatgtcgtgg gaatgctgta tgttggtggg ttacccatca actacactac ccgaagaatt 960

ggtccagtga cctatagcat tgatggctgc gtcaggaatc tccacatggc agaggcccct 1020

gccgatctgg aacaacccac ctccagcttc catgttggga catgttttgc aaatgctcag 1080

aggggaacat attttgacgg aaccggtttt gccaaagcag ttggtggatt caaagtggga 1140

ttggaccttc ttgtagaatt tgaattccgc acaactacaa cgactggagt tcttctgggg 1200

atcagtagtc aaaaaatgga tggaatgggt attgaaatga ttgatgaaaa gttgatgttt 1260

catgtggaca atggtgcggg cagattcact gctgtctatg atgctggggt tccagggcat 1320

ttgtgtgatg ga 1332

<210> 17
<211> 1023
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 17
cagatccatg ctacacccac acctgtcact gccattgggc ccccaaccac ggctatccag 60

gagcccccat ccaggatcgt gccaaccccc acatctccag ccattgctcc tccaacagag 120

accatggctc ctccagtcag ggatcctgtt cctgggaaac ccacggtcac catccggact 180

cgaggcgcca ttattcaaac cccaacccta ggccccatcc agcctactcg ggtgtcagaa 240

gctggcacca cagttcctgg ccagattcgc ccaacgatga ccattcctgg ctatgtggag 300

cctactgcag ttgctacccc tcccacaacc accaccaaga agccacgagt atccacacca 360

aaaccagcaa cgccttcaac tgactccacc accaccacga ctcgcaggcc aaccaagaaa 420

ccacggacac cccggccagt gccccgggtc accaccaaag tttccatcac cagattggaa 480

actgcctcac cgcctactcg tattcgcacc accaccagtg gagtgccccg tggcggagaa 540

cccaaccagc gcccagagct caagaaccat attgacaggg tagatgcctg ggttggcacc 600

tactttgagg tgaagatccc gtcagacact ttctatgacc atgaggacac caccactgac 660

aagctgaagc tgaccctgaa actgcgggag cagcagctgg tgggcgagaa gtcctgggta 720

cagttcaaca gcaacagcca gctcatgtat ggccttcccg acagcagcca cgtgggcaaa 780

cacgagtatt tcatgcatgc cacagacaag gggggcctgt cggctgtgga tgccttcgag 840

atccacgtcc acaggcgccc ccaaggggat agggctcctg caaggttcaa ggccaagttt 900

gtgggtgacc cggcactggt gttgaatgac atccacaaga agattgcctt ggtaaagaaa 960

ctggccttcg cctttggaga ccgaaactgt agcaccatca ccctgcagaa tatcacccgg 1020

ggc 1023

<210> 18
<211> 529
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 18
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300

catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360

atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480

acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatc 529

<210> 19
<211> 1073
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Polypeptide

<400> 19

Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val

```
1               5                   10                  15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
            20                  25                  30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
        35                  40                  45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
    50                  55                  60

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65                  70                  75                  80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
                85                  90                  95

Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Lys Val Ser Val Ser
            100                 105                 110

Ser Gly Gly Asp Cys Ile Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly
        115                 120                 125

Ser Tyr Asn Asn Ile Val Val Asn Val Lys Thr Ala Val Ala Asp Asn
    130                 135                 140

Leu Leu Phe Tyr Leu Gly Ser Ala Lys Phe Ile Asp Phe Leu Ala Ile
145                 150                 155                 160

Glu Met Arg Lys Gly Lys Val Ser Phe Leu Trp Asp Val Gly Ser Gly
                165                 170                 175

Val Gly Arg Val Glu Tyr Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp
            180                 185                 190

Tyr Arg Ile Val Ala Ser Arg Thr Gly Arg Asn Gly Thr Ile Ser Val
        195                 200                 205

Arg Ala Leu Asp Gly Pro Lys Ala Ser Ile Val Pro Ser Thr His His
```

```
    210                 215                 220

Ser Thr Ser Pro Pro Gly Tyr Thr Ile Leu Asp Val Asp Ala Asn Ala
225                 230                 235                 240

Met Leu Phe Val Gly Gly Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala
                245                 250                 255

Val Arg Val Ile Thr Phe Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp
            260                 265                 270

Asn Lys Pro Ile Gly Leu Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys
        275                 280                 285

Lys Gly Cys Thr Val Ser Pro Gln Val Glu Asp Ser Glu Gly Thr Ile
    290                 295                 300

Gln Phe Asp Gly Glu Gly Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp
305                 310                 315                 320

Tyr Pro Asn Ile Ser Thr Val Met Phe Lys Phe Arg Thr Phe Ser Ser
                325                 330                 335

Ser Ala Leu Leu Met Tyr Leu Ala Thr Arg Asp Leu Arg Asp Phe Met
            340                 345                 350

Ser Val Glu Leu Thr Asp Gly His Ile Lys Val Ser Tyr Asp Leu Gly
        355                 360                 365

Ser Gly Met Ala Ser Val Val Ser Asn Gln Asn His Asn Asp Gly Lys
    370                 375                 380

Trp Lys Ser Phe Thr Leu Ser Arg Ile Gln Lys Gln Ala Asn Ile Ser
385                 390                 395                 400

Ile Val Asp Ile Asp Thr Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser
                405                 410                 415

Ser Gly Asn Asn Phe Gly Leu Asp Leu Lys Ala Asp Asp Lys Ile Tyr
```

420 425 430

Phe Gly Gly Leu Pro Thr Leu Arg Asn Leu Ser Met Lys Ala Arg Pro
435 440 445

Glu Val Asn Leu Lys Lys Tyr Ser Gly Cys Leu Lys Asp Ile Glu Ile
450 455 460

Ser Arg Thr Pro Tyr Asn Ile Leu Ser Ser Pro Asp Tyr Val Gly Val
465 470 475 480

Thr Lys Gly Cys Ser Leu Glu Asn Val Tyr Thr Val Ser Phe Pro Lys
485 490 495

Pro Gly Phe Val Glu Leu Ser Pro Val Pro Ile Asp Val Gly Thr Glu
500 505 510

Ile Asn Leu Ser Phe Ser Thr Lys Asn Glu Ser Gly Ile Ile Leu Leu
515 520 525

Gly Ser Gly Gly Thr Pro Ala Pro Pro Arg Arg Lys Arg Arg Gln Thr
530 535 540

Gly Gln Ala Tyr Tyr Val Ile Leu Leu Asn Arg Gly Arg Leu Glu Val
545 550 555 560

His Leu Ser Thr Gly Ala Arg Thr Met Arg Lys Ile Val Ile Arg Pro
565 570 575

Glu Pro Asn Leu Phe His Asp Gly Arg Glu His Ser Val His Val Glu
580 585 590

Arg Thr Arg Gly Ile Phe Thr Val Gln Val Asp Glu Asn Arg Arg Tyr
595 600 605

Met Gln Asn Leu Thr Val Glu Gln Pro Ile Glu Val Lys Lys Leu Phe
610 615 620

Val Gly Gly Ala Pro Pro Glu Phe Gln Pro Ser Pro Leu Arg Asn Ile

625 630 635 640

Pro Pro Phe Glu Gly Cys Ile Trp Asn Leu Val Ile Asn Ser Val Pro
645 650 655

Met Asp Phe Ala Arg Pro Val Ser Phe Lys Asn Ala Asp Ile Gly Arg
660 665 670

Cys Ala His Gln Lys Leu Arg Glu Asp Glu Asp Gly Ala Ala Pro Ala
675 680 685

Glu Ile Val Ile Gln Pro Glu Pro Val Pro Thr Pro Ala Phe Pro Thr
690 695 700

Pro Thr Pro Val Leu Thr His Gly Pro Cys Ala Ala Glu Ser Glu Pro
705 710 715 720

Ala Leu Leu Ile Gly Ser Lys Gln Phe Gly Leu Ser Arg Asn Ser His
725 730 735

Ile Ala Ile Ala Phe Asp Asp Thr Lys Val Lys Asn Arg Leu Thr Ile
740 745 750

Glu Leu Glu Val Arg Thr Glu Ala Glu Ser Gly Leu Leu Phe Tyr Met
755 760 765

Ala Arg Ile Asn His Ala Asp Phe Ala Thr Val Gln Leu Arg Asn Gly
770 775 780

Leu Pro Tyr Phe Ser Tyr Asp Leu Gly Ser Gly Asp Thr His Thr Met
785 790 795 800

Ile Pro Thr Lys Ile Asn Asp Gly Gln Trp His Lys Ile Lys Ile Met
805 810 815

Arg Ser Lys Gln Glu Gly Ile Leu Tyr Val Asp Gly Ala Ser Asn Arg
820 825 830

Thr Ile Ser Pro Lys Lys Ala Asp Ile Leu Asp Val Val Gly Met Leu

```
        835                 840                 845

Tyr Val Gly Gly Leu Pro Ile Asn Tyr Thr Thr Arg Arg Ile Gly Pro
    850                 855                 860

Val Thr Tyr Ser Ile Asp Gly Cys Val Arg Asn Leu His Met Ala Glu
865                 870                 875                 880

Ala Pro Ala Asp Leu Glu Gln Pro Thr Ser Ser Phe His Val Gly Thr
                885                 890                 895

Cys Phe Ala Asn Ala Gln Arg Gly Thr Tyr Phe Asp Gly Thr Gly Phe
            900                 905                 910

Ala Lys Ala Val Gly Gly Phe Lys Val Gly Leu Asp Leu Leu Val Glu
        915                 920                 925

Phe Glu Phe Arg Thr Thr Thr Thr Thr Gly Val Leu Leu Gly Ile Ser
    930                 935                 940

Ser Gln Lys Met Asp Gly Met Gly Ile Glu Met Ile Asp Glu Lys Leu
945                 950                 955                 960

Met Phe His Val Asp Asn Gly Ala Gly Arg Phe Thr Ala Val Tyr Asp
                965                 970                 975

Ala Gly Val Pro Gly His Leu Cys Asp Gly Gln Trp His Lys Val Thr
            980                 985                 990

Ala Asn Lys Ile Lys His Arg Ile  Glu Leu Thr Val Asp  Gly Asn Gln
        995                 1000                1005

Val Glu  Ala Gln Ser Pro Asn  Pro Ala Ser Thr Ser  Ala Asp Thr
    1010                1015                1020

Asn Asp  Pro Val Phe Val Gly  Gly Phe Pro Asp Asp  Leu Lys Gln
    1025                1030                1035

Phe Gly  Leu Thr Thr Ser Ile  Pro Phe Arg Gly Cys  Ile Arg Ser
```

1040 1045 1050

Leu Lys Leu Thr Lys Gly Thr Ala Ser His Trp Arg Leu Ile Leu
1055 1060 1065

Pro Arg Pro Trp Asn
1070

<210> 20
<211> 1115
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Polypeptide

<400> 20

Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val
1 5 10 15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
20 25 30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
35 40 45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
50 55 60

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65 70 75 80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
85 90 95

Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Cys Leu Arg Lys Tyr
100 105 110

Lys Asp Phe Cys Ile His Gly Glu Cys Lys Tyr Val Lys Glu Leu Arg
115 120 125

```
Ala Pro Ser Cys Ile Cys His Pro Gly Tyr His Gly Glu Arg Cys His
    130                 135                 140

Gly Leu Ser Leu Pro Lys Val Ser Val Ser Ser Gly Gly Asp Cys Ile
145                 150                 155                 160

Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly Ser Tyr Asn Asn Ile Val
                165                 170                 175

Val Asn Val Lys Thr Ala Val Ala Asp Asn Leu Leu Phe Tyr Leu Gly
            180                 185                 190

Ser Ala Lys Phe Ile Asp Phe Leu Ala Ile Glu Met Arg Lys Gly Lys
        195                 200                 205

Val Ser Phe Leu Trp Asp Val Gly Ser Gly Val Gly Arg Val Glu Tyr
    210                 215                 220

Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp Tyr Arg Ile Val Ala Ser
225                 230                 235                 240

Arg Thr Gly Arg Asn Gly Thr Ile Ser Val Arg Ala Leu Asp Gly Pro
                245                 250                 255

Lys Ala Ser Ile Val Pro Ser Thr His His Ser Thr Ser Pro Pro Gly
            260                 265                 270

Tyr Thr Ile Leu Asp Val Asp Ala Asn Ala Met Leu Phe Val Gly Gly
        275                 280                 285

Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala Val Arg Val Ile Thr Phe
    290                 295                 300

Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp Asn Lys Pro Ile Gly Leu
305                 310                 315                 320

Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys Lys Gly Cys Thr Val Ser
                325                 330                 335
```

```
Pro Gln Val Glu Asp Ser Glu Gly Thr Ile Gln Phe Asp Gly Glu Gly
            340                 345                 350

Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp Tyr Pro Asn Ile Ser Thr
        355                 360                 365

Val Met Phe Lys Phe Arg Thr Phe Ser Ser Ser Ala Leu Leu Met Tyr
    370                 375                 380

Leu Ala Thr Arg Asp Leu Arg Asp Phe Met Ser Val Glu Leu Thr Asp
385                 390                 395                 400

Gly His Ile Lys Val Ser Tyr Asp Leu Gly Ser Gly Met Ala Ser Val
                405                 410                 415

Val Ser Asn Gln Asn His Asn Asp Gly Lys Trp Lys Ser Phe Thr Leu
            420                 425                 430

Ser Arg Ile Gln Lys Gln Ala Asn Ile Ser Ile Val Asp Ile Asp Thr
        435                 440                 445

Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser Ser Gly Asn Asn Phe Gly
    450                 455                 460

Leu Asp Leu Lys Ala Asp Asp Lys Ile Tyr Phe Gly Gly Leu Pro Thr
465                 470                 475                 480

Leu Arg Asn Leu Ser Met Lys Ala Arg Pro Glu Val Asn Leu Lys Lys
                485                 490                 495

Tyr Ser Gly Cys Leu Lys Asp Ile Glu Ile Ser Arg Thr Pro Tyr Asn
            500                 505                 510

Ile Leu Ser Ser Pro Asp Tyr Val Gly Val Thr Lys Gly Cys Ser Leu
        515                 520                 525

Glu Asn Val Tyr Thr Val Ser Phe Pro Lys Pro Gly Phe Val Glu Leu
    530                 535                 540
```

```
Ser Pro Val Pro Ile Asp Val Gly Thr Glu Ile Asn Leu Ser Phe Ser
545                 550                 555                 560

Thr Lys Asn Glu Ser Gly Ile Ile Leu Leu Gly Ser Gly Gly Thr Pro
                565                 570                 575

Ala Pro Pro Arg Arg Lys Arg Arg Gln Thr Gly Gln Ala Tyr Tyr Val
            580                 585                 590

Ile Leu Leu Asn Arg Gly Arg Leu Glu Val His Leu Ser Thr Gly Ala
        595                 600                 605

Arg Thr Met Arg Lys Ile Val Ile Arg Pro Glu Pro Asn Leu Phe His
    610                 615                 620

Asp Gly Arg Glu His Ser Val His Val Glu Arg Thr Arg Gly Ile Phe
625                 630                 635                 640

Thr Val Gln Val Asp Glu Asn Arg Arg Tyr Met Gln Asn Leu Thr Val
                645                 650                 655

Glu Gln Pro Ile Glu Val Lys Lys Leu Phe Val Gly Gly Ala Pro Pro
            660                 665                 670

Glu Phe Gln Pro Ser Pro Leu Arg Asn Ile Pro Pro Phe Glu Gly Cys
        675                 680                 685

Ile Trp Asn Leu Val Ile Asn Ser Val Pro Met Asp Phe Ala Arg Pro
    690                 695                 700

Val Ser Phe Lys Asn Ala Asp Ile Gly Arg Cys Ala His Gln Lys Leu
705                 710                 715                 720

Arg Glu Asp Glu Asp Gly Ala Ala Pro Ala Glu Ile Val Ile Gln Pro
                725                 730                 735

Glu Pro Val Pro Thr Pro Ala Phe Pro Thr Pro Thr Pro Val Leu Thr
            740                 745                 750
```

```
His Gly Pro Cys Ala Ala Glu Ser Glu Pro Ala Leu Leu Ile Gly Ser
        755                 760                 765

Lys Gln Phe Gly Leu Ser Arg Asn Ser His Ile Ala Ile Ala Phe Asp
    770                 775                 780

Asp Thr Lys Val Lys Asn Arg Leu Thr Ile Glu Leu Glu Val Arg Thr
785                 790                 795                 800

Glu Ala Glu Ser Gly Leu Leu Phe Tyr Met Ala Arg Ile Asn His Ala
                805                 810                 815

Asp Phe Ala Thr Val Gln Leu Arg Asn Gly Leu Pro Tyr Phe Ser Tyr
            820                 825                 830

Asp Leu Gly Ser Gly Asp Thr His Thr Met Ile Pro Thr Lys Ile Asn
        835                 840                 845

Asp Gly Gln Trp His Lys Ile Lys Ile Met Arg Ser Lys Gln Glu Gly
    850                 855                 860

Ile Leu Tyr Val Asp Gly Ala Ser Asn Arg Thr Ile Ser Pro Lys Lys
865                 870                 875                 880

Ala Asp Ile Leu Asp Val Val Gly Met Leu Tyr Val Gly Gly Leu Pro
                885                 890                 895

Ile Asn Tyr Thr Thr Arg Arg Ile Gly Pro Val Thr Tyr Ser Ile Asp
            900                 905                 910

Gly Cys Val Arg Asn Leu His Met Ala Glu Ala Pro Ala Asp Leu Glu
        915                 920                 925

Gln Pro Thr Ser Ser Phe His Val Gly Thr Cys Phe Ala Asn Ala Gln
    930                 935                 940

Arg Gly Thr Tyr Phe Asp Gly Thr Gly Phe Ala Lys Ala Val Gly Gly
945                 950                 955                 960
```

```
Phe Lys Val Gly Leu Asp Leu Leu Val Glu Phe Glu Phe Arg Thr Thr
                965                 970                 975


Thr Thr Thr Gly Val Leu Leu Gly Ile Ser Ser Gln Lys Met Asp Gly
            980                 985                 990


Met Gly Ile Glu Met Ile Asp Glu  Lys Leu Met Phe His  Val Asp Asn
        995                 1000                1005


Gly Ala  Gly Arg Phe Thr Ala  Val Tyr Asp Ala Gly  Val Pro Gly
    1010                1015                1020


His Leu  Cys Asp Gly Gln Trp  His Lys Val Thr Ala  Asn Lys Ile
    1025                1030                1035


Lys His  Arg Ile Glu Leu Thr  Val Asp Gly Asn Gln  Val Glu Ala
    1040                1045                1050


Gln Ser  Pro Asn Pro Ala Ser  Thr Ser Ala Asp Thr  Asn Asp Pro
    1055                1060                1065


Val Phe  Val Gly Gly Phe Pro  Asp Asp Leu Lys Gln  Phe Gly Leu
    1070                1075                1080


Thr Thr  Ser Ile Pro Phe Arg  Gly Cys Ile Arg Ser  Leu Lys Leu
    1085                1090                1095


Thr Lys  Gly Thr Ala Ser His  Trp Arg Leu Ile Leu  Pro Arg Pro
    1100                1105                1110


Trp Asn
    1115
```

<210> 21
<211> 680
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Polypeptide

<400> 21

```
Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val
1               5                  10                  15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
            20                  25                  30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
        35                  40                  45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
    50                  55                  60

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65                  70                  75                  80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
                85                  90                  95

Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Cys Leu Arg Lys Tyr
            100                 105                 110

Lys Asp Phe Cys Ile His Gly Glu Cys Lys Tyr Val Lys Glu Leu Arg
        115                 120                 125

Ala Pro Ser Cys Ile Cys His Pro Gly Tyr His Gly Glu Arg Cys His
    130                 135                 140

Gly Leu Ser Leu Pro Gln Thr Gly Gln Ala Tyr Tyr Val Ile Leu Leu
145                 150                 155                 160

Asn Arg Gly Arg Leu Glu Val His Leu Ser Thr Gly Ala Arg Thr Met
                165                 170                 175

Arg Lys Ile Val Ile Arg Pro Glu Pro Asn Leu Phe His Asp Gly Arg
            180                 185                 190

Glu His Ser Val His Val Glu Arg Thr Arg Gly Ile Phe Thr Val Gln
```

195 200 205

Val Asp Glu Asn Arg Arg Tyr Met Gln Asn Leu Thr Val Glu Gln Pro
210 215 220

Ile Glu Val Lys Lys Leu Phe Val Gly Gly Ala Pro Pro Glu Phe Gln
225 230 235 240

Pro Ser Pro Leu Arg Asn Ile Pro Pro Phe Glu Gly Cys Ile Trp Asn
245 250 255

Leu Val Ile Asn Ser Val Pro Met Asp Phe Ala Arg Pro Val Ser Phe
260 265 270

Lys Asn Ala Asp Ile Gly Arg Cys Ala His Gln Lys Leu Arg Glu Asp
275 280 285

Glu Asp Gly Ala Ala Pro Ala Glu Ile Val Ile Gln Pro Glu Pro Val
290 295 300

Pro Thr Pro Ala Phe Pro Thr Pro Thr Pro Val Leu Thr His Gly Pro
305 310 315 320

Cys Ala Ala Glu Ser Glu Pro Ala Leu Leu Ile Gly Ser Lys Gln Phe
325 330 335

Gly Leu Ser Arg Asn Ser His Ile Ala Ile Ala Phe Asp Asp Thr Lys
340 345 350

Val Lys Asn Arg Leu Thr Ile Glu Leu Glu Val Arg Thr Glu Ala Glu
355 360 365

Ser Gly Leu Leu Phe Tyr Met Ala Arg Ile Asn His Ala Asp Phe Ala
370 375 380

Thr Val Gln Leu Arg Asn Gly Leu Pro Tyr Phe Ser Tyr Asp Leu Gly
385 390 395 400

Ser Gly Asp Thr His Thr Met Ile Pro Thr Lys Ile Asn Asp Gly Gln

405 410 415

Trp His Lys Ile Lys Ile Met Arg Ser Lys Gln Glu Gly Ile Leu Tyr
420 425 430

Val Asp Gly Ala Ser Asn Arg Thr Ile Ser Pro Lys Lys Ala Asp Ile
435 440 445

Leu Asp Val Val Gly Met Leu Tyr Val Gly Gly Leu Pro Ile Asn Tyr
450 455 460

Thr Thr Arg Arg Ile Gly Pro Val Thr Tyr Ser Ile Asp Gly Cys Val
465 470 475 480

Arg Asn Leu His Met Ala Glu Ala Pro Ala Asp Leu Glu Gln Pro Thr
485 490 495

Ser Ser Phe His Val Gly Thr Cys Phe Ala Asn Ala Gln Arg Gly Thr
500 505 510

Tyr Phe Asp Gly Thr Gly Phe Ala Lys Ala Val Gly Gly Phe Lys Val
515 520 525

Gly Leu Asp Leu Leu Val Glu Phe Glu Phe Arg Thr Thr Thr Thr Thr
530 535 540

Gly Val Leu Leu Gly Ile Ser Ser Gln Lys Met Asp Gly Met Gly Ile
545 550 555 560

Glu Met Ile Asp Glu Lys Leu Met Phe His Val Asp Asn Gly Ala Gly
565 570 575

Arg Phe Thr Ala Val Tyr Asp Ala Gly Val Pro Gly His Leu Cys Asp
580 585 590

Gly Gln Trp His Lys Val Thr Ala Asn Lys Ile Lys His Arg Ile Glu
595 600 605

Leu Thr Val Asp Gly Asn Gln Val Glu Ala Gln Ser Pro Asn Pro Ala

610 615 620

Ser Thr Ser Ala Asp Thr Asn Asp Pro Val Phe Val Gly Gly Phe Pro
625 630 635 640

Asp Asp Leu Lys Gln Phe Gly Leu Thr Thr Ser Ile Pro Phe Arg Gly
645 650 655

Cys Ile Arg Ser Leu Lys Leu Thr Lys Gly Thr Ala Ser His Trp Arg
660 665 670

Leu Ile Leu Pro Arg Pro Trp Asn
675 680

<210> 22
<211> 680
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Polypeptide

<400> 22

Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val
1 5 10 15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
20 25 30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
35 40 45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
50 55 60

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65 70 75 80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
85 90 95

```
Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Cys Leu Arg Lys Tyr
            100                 105                 110

Lys Asp Phe Cys Ile His Gly Glu Cys Lys Tyr Val Lys Glu Leu Arg
        115                 120                 125

Ala Pro Ser Cys Ile Cys His Pro Gly Tyr His Gly Glu Arg Cys His
    130                 135                 140

Gly Leu Ser Leu Pro Gln Thr Gly Gln Ala Tyr Tyr Val Ile Leu Leu
145                 150                 155                 160

Asn Arg Gly Arg Leu Glu Val His Leu Ser Thr Gly Ala Arg Thr Met
                165                 170                 175

Arg Lys Ile Val Ile Arg Pro Glu Pro Asn Leu Phe His Asp Gly Arg
            180                 185                 190

Glu His Ser Val His Val Glu Arg Thr Arg Gly Ile Phe Thr Val Gln
        195                 200                 205

Val Asp Glu Asn Arg Arg Tyr Met Gln Asn Leu Thr Val Glu Gln Pro
    210                 215                 220

Ile Glu Val Lys Lys Leu Phe Val Gly Gly Ala Pro Pro Glu Phe Gln
225                 230                 235                 240

Pro Ser Pro Leu Arg Asn Ile Pro Pro Phe Glu Gly Cys Ile Trp Asn
                245                 250                 255

Leu Val Ile Asn Ser Val Pro Met Asp Phe Ala Arg Pro Val Ser Phe
            260                 265                 270

Lys Asn Ala Asp Ile Gly Arg Cys Ala His Gln Lys Leu Arg Glu Asp
        275                 280                 285

Glu Asp Gly Ala Ala Pro Ala Glu Ile Val Ile Gln Pro Glu Pro Val
    290                 295                 300
```

```
Pro Thr Pro Ala Phe Pro Thr Pro Thr Pro Val Leu Thr His Gly Pro
305                 310                 315                 320

Cys Ala Ala Glu Ser Glu Pro Ala Leu Leu Ile Gly Ser Lys Gln Phe
                325                 330                 335

Gly Leu Ser Arg Asn Ser His Ile Ala Ile Ala Phe Asp Asp Thr Lys
            340                 345                 350

Val Lys Asn Arg Leu Thr Ile Glu Leu Glu Val Arg Thr Glu Ala Glu
        355                 360                 365

Ser Gly Leu Leu Phe Tyr Met Ala Arg Ile Asn His Ala Asp Phe Ala
    370                 375                 380

Thr Val Gln Leu Arg Asn Gly Leu Pro Tyr Phe Ser Tyr Asp Leu Gly
385                 390                 395                 400

Ser Gly Asp Thr His Thr Met Ile Pro Thr Lys Ile Asn Asp Gly Gln
                405                 410                 415

Trp His Lys Ile Lys Ile Met Arg Ser Lys Gln Glu Gly Ile Leu Tyr
            420                 425                 430

Val Asp Gly Ala Ser Asn Arg Thr Ile Ser Pro Lys Lys Ala Asp Ile
        435                 440                 445

Leu Asp Val Val Gly Met Leu Tyr Val Gly Gly Leu Pro Ile Asn Tyr
    450                 455                 460

Thr Thr Arg Arg Ile Gly Pro Val Thr Tyr Ser Ile Asp Gly Cys Val
465                 470                 475                 480

Arg Asn Leu His Met Ala Glu Ala Pro Ala Asp Leu Glu Gln Pro Thr
                485                 490                 495

Ser Ser Phe His Val Gly Thr Cys Phe Ala Asn Ala Gln Arg Gly Thr
            500                 505                 510
```

```
Tyr Phe Asp Gly Thr Gly Phe Ala Lys Ala Val Gly Gly Phe Lys Val
        515                 520                 525


Gly Leu Asp Leu Leu Val Glu Phe Glu Phe Arg Thr Thr Thr Thr Thr
    530                 535                 540


Gly Val Leu Leu Gly Ile Ser Ser Gln Lys Met Asp Gly Met Gly Ile
545                 550                 555                 560


Glu Met Ile Asp Glu Lys Leu Met Phe His Val Asp Asn Gly Ala Gly
                565                 570                 575


Arg Phe Thr Ala Val Tyr Asp Ala Gly Val Pro Gly His Leu Cys Asp
            580                 585                 590


Gly Gln Trp His Lys Val Thr Ala Asn Lys Ile Lys His Arg Ile Glu
        595                 600                 605


Leu Thr Val Asp Gly Asn Gln Val Glu Ala Gln Ser Pro Asn Pro Ala
    610                 615                 620


Ser Thr Ser Ala Asp Thr Asn Asp Pro Val Phe Val Gly Gly Phe Pro
625                 630                 635                 640


Asp Asp Leu Lys Gln Phe Gly Leu Thr Thr Ser Ile Pro Phe Arg Gly
                645                 650                 655


Cys Ile Arg Ser Leu Lys Leu Thr Lys Gly Thr Ala Ser His Trp Arg
            660                 665                 670


Leu Ile Leu Pro Arg Pro Trp Asn
        675                 680


<210>  23
<211>  448
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic Polypeptide
```

<400> 23

```
Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val
1               5                  10                  15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
            20                  25                  30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
        35                  40                  45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
    50                  55                  60

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65                  70                  75                  80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
                85                  90                  95

Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Gln Ile His Ala Thr
            100                 105                 110

Pro Thr Pro Val Thr Ala Ile Gly Pro Pro Thr Thr Ala Ile Gln Glu
        115                 120                 125

Pro Pro Ser Arg Ile Val Pro Thr Pro Thr Ser Pro Ala Ile Ala Pro
    130                 135                 140

Pro Thr Glu Thr Met Ala Pro Pro Val Arg Asp Pro Val Pro Gly Lys
145                 150                 155                 160

Pro Thr Val Thr Ile Arg Thr Arg Gly Ala Ile Ile Gln Thr Pro Thr
                165                 170                 175

Leu Gly Pro Ile Gln Pro Thr Arg Val Ser Glu Ala Gly Thr Thr Val
            180                 185                 190

Pro Gly Gln Ile Arg Pro Thr Met Thr Ile Pro Gly Tyr Val Glu Pro
```

195 200 205

Thr Ala Val Ala Thr Pro Pro Thr Thr Thr Thr Lys Lys Pro Arg Val
210 215 220

Ser Thr Pro Lys Pro Ala Thr Pro Ser Thr Asp Ser Thr Thr Thr Thr
225 230 235 240

Thr Arg Arg Pro Thr Lys Lys Pro Arg Thr Pro Arg Pro Val Pro Arg
245 250 255

Val Thr Thr Lys Val Ser Ile Thr Arg Leu Glu Thr Ala Ser Pro Pro
260 265 270

Thr Arg Ile Arg Thr Thr Thr Ser Gly Val Pro Arg Gly Gly Glu Pro
275 280 285

Asn Gln Arg Pro Glu Leu Lys Asn His Ile Asp Arg Val Asp Ala Trp
290 295 300

Val Gly Thr Tyr Phe Glu Val Lys Ile Pro Ser Asp Thr Phe Tyr Asp
305 310 315 320

His Glu Asp Thr Thr Thr Asp Lys Leu Lys Leu Thr Leu Lys Leu Arg
325 330 335

Glu Gln Gln Leu Val Gly Glu Lys Ser Trp Val Gln Phe Asn Ser Asn
340 345 350

Ser Gln Leu Met Tyr Gly Leu Pro Asp Ser Ser His Val Gly Lys His
355 360 365

Glu Tyr Phe Met His Ala Thr Asp Lys Gly Gly Leu Ser Ala Val Asp
370 375 380

Ala Phe Glu Ile His Val His Arg Arg Pro Gln Gly Asp Arg Ala Pro
385 390 395 400

Ala Arg Phe Lys Ala Lys Phe Val Gly Asp Pro Ala Leu Val Leu Asn

405 410 415

Asp Ile His Lys Lys Ile Ala Leu Val Lys Lys Leu Ala Phe Ala Phe
420 425 430

Gly Asp Arg Asn Cys Ser Thr Ile Thr Leu Gln Asn Ile Thr Arg Gly
435 440 445

<210> 24
<211> 490
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Polypeptide

<400> 24

Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val
1 5 10 15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
20 25 30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
35 40 45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
50 55 60

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65 70 75 80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
85 90 95

Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Cys Leu Arg Lys Tyr
100 105 110

Lys Asp Phe Cys Ile His Gly Glu Cys Lys Tyr Val Lys Glu Leu Arg
115 120 125

```
Ala Pro Ser Cys Ile Cys His Pro Gly Tyr His Gly Glu Arg Cys His
    130                 135                 140

Gly Leu Ser Leu Pro Gln Ile His Ala Thr Pro Thr Pro Val Thr Ala
145                 150                 155                 160

Ile Gly Pro Pro Thr Thr Ala Ile Gln Glu Pro Pro Ser Arg Ile Val
                165                 170                 175

Pro Thr Pro Thr Ser Pro Ala Ile Ala Pro Pro Thr Glu Thr Met Ala
            180                 185                 190

Pro Pro Val Arg Asp Pro Val Pro Gly Lys Pro Thr Val Thr Ile Arg
        195                 200                 205

Thr Arg Gly Ala Ile Ile Gln Thr Pro Thr Leu Gly Pro Ile Gln Pro
    210                 215                 220

Thr Arg Val Ser Glu Ala Gly Thr Thr Val Pro Gly Gln Ile Arg Pro
225                 230                 235                 240

Thr Met Thr Ile Pro Gly Tyr Val Glu Pro Thr Ala Val Ala Thr Pro
                245                 250                 255

Pro Thr Thr Thr Thr Lys Lys Pro Arg Val Ser Thr Pro Lys Pro Ala
            260                 265                 270

Thr Pro Ser Thr Asp Ser Thr Thr Thr Thr Thr Arg Arg Pro Thr Lys
        275                 280                 285

Lys Pro Arg Thr Pro Arg Pro Val Pro Arg Val Thr Thr Lys Val Ser
    290                 295                 300

Ile Thr Arg Leu Glu Thr Ala Ser Pro Pro Thr Arg Ile Arg Thr Thr
305                 310                 315                 320

Thr Ser Gly Val Pro Arg Gly Gly Glu Pro Asn Gln Arg Pro Glu Leu
                325                 330                 335
```

```
Lys Asn His Ile Asp Arg Val Asp Ala Trp Val Gly Thr Tyr Phe Glu
            340                 345                 350

Val Lys Ile Pro Ser Asp Thr Phe Tyr Asp His Glu Asp Thr Thr Thr
        355                 360                 365

Asp Lys Leu Lys Leu Thr Leu Lys Leu Arg Glu Gln Gln Leu Val Gly
    370                 375                 380

Glu Lys Ser Trp Val Gln Phe Asn Ser Asn Ser Gln Leu Met Tyr Gly
385                 390                 395                 400

Leu Pro Asp Ser Ser His Val Gly Lys His Glu Tyr Phe Met His Ala
                405                 410                 415

Thr Asp Lys Gly Gly Leu Ser Ala Val Asp Ala Phe Glu Ile His Val
            420                 425                 430

His Arg Arg Pro Gln Gly Asp Arg Ala Pro Ala Arg Phe Lys Ala Lys
        435                 440                 445

Phe Val Gly Asp Pro Ala Leu Val Leu Asn Asp Ile His Lys Lys Ile
    450                 455                 460

Ala Leu Val Lys Lys Leu Ala Phe Ala Phe Gly Asp Arg Asn Cys Ser
465                 470                 475                 480

Thr Ile Thr Leu Gln Asn Ile Thr Arg Gly
                485                 490
```