ADENO ASSOCIATED VIRUS VECTORS FOR THE TREATMENT OF HUNTER DISEASE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This applications claims benefit of, and priority to, U.S. Serial Number 62/945,920 filed on December 10, 2019, the contents of which are incorporated herein.

BACKGROUND

[0002] Hunter syndrome, also known as mucopolysaccharidosis Type II (MPS II), is a lysosomal storage disease caused by deficiency or absence of iduronate-2-sulfatase (I2S) enzyme. Iduronate-2-sulfatase is involved in the break down and recycling of specific mucopolysaccharides, also known as glycosaminoglycans or GAG. As a result, in Hunter syndrome, GAG builds up in cells throughout the body, which in turn interferes with the normal function of various cells and organs in the body, resulting in a number of serious symptoms. In many cases of Hunter syndrome, there is a large buildup of GAGs in neurons and in the meninges of affected individuals, leading to various forms of central nervous system (CNS) symptoms, impaired cognitive performance and development delays.

[0003] Various treatment options have been used in the management of Hunter syndrome, including enzyme replacement therapy (ERT). Approved therapeutic ERT treatments include intravenous administration of recombinant I2S enzyme. However, intravenously administered I2S enzyme has various limitations, including poor distribution into the cells and tissues of the CNS and poor distribution into the cells of deep somatic tissues such as heart, lung, and bone. Treatment of Hunter syndrome remains a challenge.

[0004] The use of vectors that produce therapeutic proteins *in vivo* is desirable for the treatment of disease, but is limited by various factors including poor production of desired therapeutic proteins in vivo.

SUMMARY

[0005] The present invention provides efficient and robust recombinant adeno-associated virus (rAAV) vectors that encode I2S (referred to as I2S or IDS, throughout this application). The present invention is based in part on the surprising discovery that optimized rAAV vectors comprising I2S sequences result in robust I2S expression in vivo.

[0006] In some aspects, the present invention provides, a recombinant adeno-associated virus (rAAV) vector comprising an AAV8 capsid and a sequence encoding a human iduronate-2-sulfatase (I2S) enzyme.

[0007] In some aspects, the present invention provides, a recombinant adeno-associated virus (rAAV) vector comprising an AAV9 capsid and a sequence encoding human iduronate-2 sulfatase (I2S) enzyme.

[0008] In some embodiments, the rAAV encodes a codon-optimized human I2S enzyme. In some embodiments, the codon-optimized human I2S has a nucleotide sequence selected from SEQ ID NO: 11 or 12.

[0009] In some embodiments, the sequence encoding a human I2S enzyme comprises a sequence having at least about 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to SEQ ID NO:6. In some embodiments, the human I2S enzyme is encoded by a nucleotide sequence at least about 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to SEQ ID NO: 6. In some embodiments, the human I2S enzyme is encoded by a nucleotide sequence of SEQ ID NO: 6.

[0010] In some embodiments, the amino acid sequence of a human I2S enzyme comprises a sequence identical to SEQ ID NO: 1. In some embodiments, the amino acid sequence of a human I2S enzyme is the sequence identical to SEQ ID NO: 1.

[0011] In some embodiments, the amino acid sequence of a human I2S enzyme comprises a sequence having at least about 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to SEQ ID NO: 2.

[0012] In some embodiments, the sequence encoding a human I2S enzyme comprises a sequence identical to SEQ ID NO: 2. In some embodiments, the amino acid sequence of a human I2S enzyme is the sequence identical to SEQ ID NO: 2.

[0013] In some embodiments, the codon-optimized sequence encoding a human I2S enzyme comprises a sequence having at least about 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to SEQ ID NO: 11 or 12. In some embodiments, the codon-optimized sequence encoding a human I2S enzyme comprises a sequence identical to SEQ ID NO: 11 or 12.

[0014] In some embodiments, the vector further comprises a liver-specific promoter.

[0015] In some embodiments, the liver-specific promoter is transthyretin promoter (TTR).

[0016] In some embodiments, the vector further comprises a 5' and a 3' inverted terminal repeat (ITR), an intron upstream of the I2S sequence, and a cis-acting regulatory module (CRM).

[0017] In some embodiments, the vector further comprises a ubiquitous promoter.

[0018] In some embodiments, the vector further comprises a 5' and a 3' inverted terminal repeat, an intron upstream of the I2S sequence, and a cis-acting regulatory module (CRM).

[0019] In some embodiments, the rAAV vector comprises a sulfatase modifying factor 1 (SUMF1).

[0020] In some embodiments, the SUMF1 is preceded by an internal ribosome entry site (IRES).

[0021] In some embodiments, the vector further comprises a WPRE sequence. In some embodiments, the WPRE sequence is a variant WPRE sequence or an optimized WPRE sequence. In some embodiments, the WPRE sequence is encoded by a nucleotide sequence having at least about 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to SEQ ID NO: 7. In some embodiments, the WPRE sequence is encoded by a nucleotide sequence having a SEQ ID NO: 7.

[0022] In some embodiments, the intron is a minute virus of mice (MVM) or SV40 intron. In some embodiments, the intron is a β-globin/IgG chimeric intron.

[0023] In some embodiments, the CRM is liver-specific CRM.

[0024] In some embodiments, the CRM is a neuronal-specific CRM. In some embodiments, the CRM is a muscle-specific CRM.

[0025] In some embodiments, the CRM is CRM8.

[0026] In some embodiments, the vector comprises at least three CRMs.

[0027] In some aspects, the present invention provides, a recombinant adeno-associated virus (rAAV) comprising an AAV8 capsid and an rAAV vector, said vector comprising: a) a 5' inverted terminal repeat (ITR); b.) a cis-acting regulatory module (CRM); c.) a liver specific promoter; d) a minute virus of mice (MVM); e. a sequence encoding a

human iduronate-2-sulfatase (I2S) enzyme; f.) a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE); and g.) a 3' ITR.

[0028] In some embodiments, the sequence encoding a human I2S enzyme is a wild type sequence or a codon-optimized sequence.

[0029] In some embodiments, the nucleotide sequence encoding a human I2S enzyme is a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity with SEQ ID NO: 6. In some embodiments, the nucleotide sequence encoding human I2S is identical to SEQ ID NO 6. In some embodiments, the nucleotide sequence encoding a human I2S enzyme is a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity with SEQ ID NO: 11 or 12. In some embodiments, the nucleotide sequence encoding a human I2S enzyme is a sequence identical to SEQ ID NO: 11 or 12.

[0030] In some embodiments, comprising a sequence encoding a sulfatase modifying factor 1 (SUMF1) and an internal ribosome entry site (IRES).

[0031] In some aspects, the present invention provides a recombinant adeno-associated virus (rAAV) comprising an AAV9 capsid and an rAAV vector, said vector comprising: a.) 5' inverted terminal repeat (ITR); b.) a cis-acting regulatory module (CRM); c) a ubiquitous promoter; d.) a minute virus of mice (MVM); e.) a sequence encoding a human iduronate-2-sulfatase (I2S) enzyme; f.) a woodchuck hepatitis virus posttranscriptional regulatory element(WPRE); and g.) a 3' ITR.

[0032] In some embodiments, the sequence encoding a human I2S enzyme is a wild type sequence or a codon-optimized sequence.

[0033] In some embodiments, comprising a sequence encoding a sulfatase modifying factor 1 (SUMF1) and an internal ribosome entry site (IRES).

[0034] In some embodiments, the rAAV vector does not comprise ApoB.

[0035] A method of treating a subject having Hunter syndrome (MPS II), comprising administering to the subject in need thereof an rAAV of any one of the preceding claims.

[0036] A method of treating a subject having Hunter syndrome (MPS II), comprising administering to the subject in need thereof a recombinant adeno-associated virus (rAAV) vector comprising an AAV8 or AAV9 capsid, and a promoter operably linked to a nucleic

acid sequence that encodes iduronate-2-sulfatase (I2S), and wherein administering results in an increase in I2S enzymatic activity in the subject.

[0037] In some embodiments, the increase in I2S activity is detected in the serum of the subject.

[0038] In some embodiments, the increase in I2S activity is detected in the liver of the subject.

[0039] In some embodiments, the I2S activity is detected in the central nervous system (CNS).

[0040] In some embodiments, the increase in I2S activity is detected in the brain of the subject.

[0041] In some embodiments, the increase in I2S activity is detected in the hippocampus, thalamus, corpus callosum, cortex, cerebellum, or stratum of the brain.

[0042] In some embodiments, the increase of I2S activity is detected in the kidney etc. of the subject. In some embodiments, the increase in I2S activity is detected in the heart of the subject. In some embodiments, the increase in I2S activity is detected in the lung of the subject. In some embodiments, the increase in I2S activity is detected in the bone marrow of the subject. In some embodiments, the increase in I2S activity is detected in the kidney of the subject.

[0043] In some embodiments, the increase of I2S activity is maintained for at least 30, 60, 90, 120, 150, 180 days or more after a single administration.

[0044] In some embodiments, the level of I2S activity is measured by heparin sulfate assay.

[0045] In some embodiments, the level of I2S activity is measured by dermatan sulfate assay.

[0046] In some embodiments, the administering the AAV reduces the level of glycosaminoglycan (GAG) in the subject.

[0047] In some embodiments, the administering the AAV reduces the level of GAG in the serum of the subject.

[0048] In some embodiments, the administering the AAV reduces the level of GAG in the liver of the subject.

[0049] In some embodiments, the administering the AAV reduces the level of GAG in the kidney etc. of the subject. In some embodiments, the administering the AAV reduces the level of GAG in the heart of the subject. In some embodiments, the administering the AAV reduces the level of GAG in the lung of the subject. In some embodiments, the administering the AAV reduces the level of GAG in the bone marrow of the subject. In some embodiments, the administering the AAV reduces the level of GAG in the kidney of the subject.

[0050] In some embodiments, the administering the AAV reduces the level of GAG in the CNS of the subject.

[0051] In some embodiments, the administering the AAV reduces the level of GAG in the brain of the subject.

[0052] In some embodiments, the administering the AAV reduces the level of GAG in in the hippocampus, thalamus, corpus callosum, cortex, cerebellum, or stratum of the brain.

[0053] In some embodiments, the AAV is administered intravenously.

[0054] In some embodiments, the AAV is administered intrathecally.

[0055] In some embodiments, the AAV is administered at dose of about 5×10^9 vg.

[0056] In some embodiments, the administering of the rAAV does not elicit immune response.

[0057] Various aspects of the invention are described in detail in the following sections. The use of sections is not meant to limit the invention. Each section can apply to any aspect of the invention. In this application, the use of "or" means "and/or" unless stated otherwise. As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

BRIEF DESCRIPTION OF THE DRAWINGS

[0058] Figure 1A is a schematic representation of the expression construct for hIDS expressing vector. Figure 1B is a series of schematics that show the optimized expression

constructs of hIDS-WPRE. ITR: inverted terminal repeat; pTBG: thyroid hormone-binding globulin promoter; hTTR: human transthyretin promoter; CRM: cis-acting regulatory module; MVM intron: minute virus of mice intron; WPRE: Woodchuck Hepatitis Virus (WHV) Posttranscriptional Regulatory Element; BGH pA: Bovine growth hormone terminator + polyA; COOLopt: codon-optimized by COOL (Codon Optimization Online) platform; AUSopt: codon-optimized by internal codon usage frequency table.

[0059] Figure 2 is a series of schematics that show the optimized expression constructs of hIDS-IRES-SUMF1. ITR: inverted terminal repeat; pTBG: thyroid hormone-binding globulin promoter; hTTR: human transthyretin promoter; CRM: cis-acting regulatory module; MVM intron: minute virus of mice intron; WPRE: Woodchuck Hepatitis Virus (HWP) Posttranscriptional Regulatory Element; BGH pA: Bovine growth hormone terminator + polyA; COOLopt: codon-optimized by COOL (Codon Optimization Online) platform; AUSopt: codon-optimized by internal codon usage frequency table.

[0060] Figure 3 is a graph that shows total I2S concentrations in mouse serum at day 0, day 2, day 7, day 21, week 8 and week 12 post injection of the identified rAAV vectors. The hI2S concentration was determined by ELISA.

[0061] Figure 4A is a graph that shows I2S activity in mouse serum at day 0, day 2, day 7, day 21, week 8 and week 12 post injection of the identified rAAV vectors. Figure 4B is a graph that shows I2S concentration in mouse tissues. Figure 4C is a graph that shows I2S activity in mouse tissues. Figure 4D is a graph that shows relative I2S activity in animals administered hIDS expressing vector relative to wild-type.

[0062] Figure 5A is a graph that shows the GAG level in mouse brain. Figure 5B is a graph that shows the GAG level in mouse liver. Figure 5C is a graph that shows the GAG level in mouse kidney at week 12 post injection of the identified rAAV vectors.

[0063] Figure 6A is a graph that shows the heparan sulfate (HS) and dermatan sulfate (DS) GAG levels in liver and kidney etc. of mice at week 12 post injection of the identified rAAV vectors. One-way ANOVA was carried out relative to untreated knockout animals with multiple comparisons using Dunnett method. **P<0.01, ***P<0.001.

[0064] Figure 6B is a graph that shows the percent reduction of heparan sulfate (HS) and dermatan sulfate (DS) levels in liver and kidney etc. relative to control animals at week 12 post injection of the identified rAAV vectors.

[0065] Figure 7A is a graph that shows the LAMP1 level in the liver determined by LAMP1 immunohistochemistry staining, which serves as a biomarker of detection of reduced lysosomal storage compartment. One-way ANOVA was carried out relative to untreated knockout animals with multiple comparisons using Dunnett method. *** indicated P<0.001. Figure 7B is a graph that shows the LAMP1 level in the brain determined by LAMP1 immunohistochemistry staining, which serves as a biomarker of detection of reduced lysosomal storage compartment. One-way ANOVA was carried out relative to untreated knockout animals with multiple comparisons using Dunnett method. *P<0.05, **P<0.01, ***P<0.001.

[0066] Figure 8 is a graph that shows total hI2S concentrations in mouse serum at day 2, day 7, and day 21 post injection with vectors expressing hI2S and SUMF1.

[0067] Figure 9 is a graph that shows hI2S activity in mouse serum at day 0, day 2, day 7, and day 21 post injection of the identified rAAV vectors.

[0068] Figure 10 is a graph that shows I2S concentration up to 364 days post-injection. Figure 11 is a graph that shows I2S concentration up to 364 days post-injection.

[0069] Figure 12A is a graph of I2S concentration in mice administered hI2S as measured by ELISA in serum and tissue over 12 months. Figure 12B is a graph of I2S activity in serum and mouse tissue. Figure 12C is a graph of GAG levels in mouse tissues. Figure 12D is a graph of percent reduction in GAG levels in mouse tissues. Figure 12E is a graph of LAMP1 staining, which is a biomarker of lysosomal storage compartment. Statistics were performed using one-way ANOVA relative to untreated knockout animals with multiple comparisons using Dunnett method. *P<0.05, **P<0.01, ***P<0.001, ns=not significant.

[0070] Figure 13A shows a graph of bone volume in the humerus, Figure 13B shows a graph of bone volume in the zygomatic arch as measured by microCT in mice administered I2S relative to control animals.

[0071] Figure 14A shows a graph of I2S concentration in mice administered low dose hI2S as measured by ELISA in serum over 90 days. Figure 14B is a graph of I2S activity in serum. Figure 14C is a graph of I2S concentration in mouse tissue. Figure 14D is a graph of I2S activity in mouse tissue. Figure 14E is a graph of relative I2S activity in animals administered a low dose of hI2S relative to wild-type animals.

[0072] Figure 14F is a graph that shows levels of LAMP1 staining in mice administered hI2S in somatic tissues including liver, spleen, heart and kidney. Figure 14G is a graph that shows levels of LAMP1 staining in the brain.

[0073] Figure 15A is a graph that shows I2S serum concentration in non-human primates that received a low dose of I2S. Figure 15B is a graph that shows I2S serum activity in non-human primates that received a low dose of I2S.

[0074] Figure 15C is a graph that shows I2S serum concentration in non-human primates that received a high dose of I2S. Figure 15D is a graph that shows I2S serum activity in non-human primates that received a high dose of I2S.

[0075] Figure 16 shows I2S concentration and enzyme activity profiles in individual non-human primates and anti-I2S ADA activity profiles in the same animals.

[0076] Figure 17A shows hI2S concentration in the liver of individual non-human primates administered a low dose of hI2S. Figure 17B shows hI2S concentration in the liver of individual non-human primates administered a high dose of hI2S.

[0077] Figure 18A shows hI2S concentration in the kidney of non-human primates administered high and low doses of I2S. Figure 18B shows hI2S concentration in the spleen of non-human primates administered high and low doses of I2S. Figure 18C shows hI2S concentration in the lung of non-human primates administered high and low doses of I2S. Figure 18D shows hI2S concentration in the heart of non-human primates administered high and low doses of I2S. Figure 18E shows hI2S concentration in the bone marrow of non-human primates administered high and low doses of I2S.

DEFINITIONS

[0078] *Adeno-associated virus (AAV)*: As used herein, the terms "adeno-associated virus" or "AAV" or recombinant AAV ("rAAV") includes, but is not limited to, AAV type 1, AAV type 2, AAV type 3 (including types 3A and 3B), AAV type 4, AAV type 5, AAV type 6, AAV type 7, AAV type 8, AAV type 9, AAV type 10, AAV type 11, avian AAV, bovine AAV, canine AAV, equine AAV, and ovine AAV (*see*, e.g., Fields et al., Virology, volume 2, chapter 69 (4th ed., Lippincott-Raven Publishers); Gao et al., J. Virology 78:6381-6388 (2004); Mori et al., Virology 330:375-383 (2004)). Typically, AAV can infect both dividing

and non-dividing cells and can be present in an extrachromosomal state without integrating into the genome of a host cell. AAV vectors are commonly used in gene therapy.

[0079] *Administering*: As used herein, the terms "administering," or "introducing" are used interchangeably in the context of delivering rAAV vectors encoding I2S into a subject, by a method or route which results in efficient delivery of the rAAV vector. Various methods are known in the art for administering rAAV vectors, including for example intravenously, subcutaneously or transdermally. Transdermal administration of rAAV vector can be performed by use of a "gene gun" or biolistic particle delivery system. In some embodiments, the rAAV vectors and/or the transgene expression cassette and/or the optimized *IDS* transgene sequences and/or any compositions of the gene expression cassette are administered via non-viral chemical particles such as lipid nanoparticles, non-viral biological molecules such as exosomes and/or extracellular vesicle.

[0080] *Animal:* As used herein, the term "animal" refers to any member of the animal kingdom. In some embodiments, "animal" refers to humans, at any stage of development. In some embodiments, "animal" refers to non-human animals, at any stage of development. In certain embodiments, the non-human animal is a mammal (*e.g.*, a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, and/or a pig). In some embodiments, animals include, but are not limited to, mammals, birds, reptiles, amphibians, fish, insects, and/or worms. In some embodiments, an animal may be a transgenic animal, genetically-engineered animal, and/or a clone.

[0081] The recognized immunoglobulin polypeptides include the kappa and lambda light chains and the alpha, gamma (IgG1, IgG2, IgG3, IgG4), delta, epsilon and mu heavy chains or equivalents in other species. Full-length immunoglobulin "light chains" (of about 25 kDa or about 214 amino acids) comprise a variable region of about 110 amino acids at the NH2-terminus and a kappa or lambda constant region at the COOH-terminus. Full-length immunoglobulin "heavy chains" (of about 50 kDa or about 446 amino acids), similarly comprise a variable region (of about 116 amino acids) and one of the aforementioned heavy chain constant regions, e.g., gamma (of about 330 amino acids).

[0082] *Approximately or about:* As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" or "about" refers

to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

[0083] *Biologically active*: As used herein, the phrase "biologically active" refers to a characteristic of any agent that has activity in a biological system, and particularly in an organism. For instance, an agent that, when administered to an organism, has a biological effect on that organism, is considered to be biologically active. In particular embodiments, where a peptide is biologically active, a portion of that peptide that shares at least one biological activity of the peptide is typically referred to as a "biologically active" portion.

[0084] *Functional equivalent or derivative*: As used herein, the term "functional equivalent" or "functional derivative" denotes, in the context of a functional derivative of an amino acid sequence, a molecule that retains a biological activity (either function or structural) that is substantially similar to that of the original sequence. A functional derivative or equivalent may be a natural derivative or is prepared synthetically. Exemplary functional derivatives include amino acid sequences having substitutions, deletions, or additions of one or more amino acids, provided that the biological activity of the protein is conserved. The substituting amino acid desirably has chemico-physical properties which are similar to that of the substituted amino acid. Desirable similar chemico-physical properties include, similarities in charge, bulkiness, hydrophobicity, hydrophilicity, and the like.

[0085] *In vitro*: As used herein, the term "*in vitro*" refers to events that occur in an artificial environment, *e.g.*, in a test tube or reaction vessel, in cell culture, *etc.*, rather than within a multi-cellular organism.

[0086] *In vivo*: As used herein, the term "*in vivo*" refers to events that occur within a multi-cellular organism, such as a human and a non-human animal. In the context of cell-based systems, the term may be used to refer to events that occur within a living cell (as opposed to, for example, *in vitro* systems).

[0087] *IRES*: As used herein, the term "IRES" refers to any suitable internal ribosome entry site sequence.

[0088] *Isolated*: As used herein, the term "isolated" refers to a substance and/or entity that has been (1) separated from at least some of the components with which it was

associated when initially produced (whether in nature and/or in an experimental setting), and/or (2) produced, prepared, and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, about 98%, about 99%, substantially 100%, or 100% of the other components with which they were initially associated. In some embodiments, isolated agents are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, substantially 100%, or 100% pure. As used herein, a substance is "pure" if it is substantially free of other components. As used herein, the term "isolated cell" refers to a cell not contained in a multi-cellular organism.

[0089] *Polypeptide*: The term, "polypeptide," as used herein refers a sequential chain of amino acids linked together via peptide bonds. The term is used to refer to an amino acid chain of any length, but one of ordinary skill in the art will understand that the term is not limited to lengthy chains and can refer to a minimal chain comprising two amino acids linked together via a peptide bond. As is known to those skilled in the art, polypeptides may be processed and/or modified.

[0090] *Protein*: The term "protein" as used herein refers to one or more polypeptides that function as a discrete unit. If a single polypeptide is the discrete functioning unit and does not require permanent or temporary physical association with other polypeptides in order to form the discrete functioning unit, the terms "polypeptide" and "protein" may be used interchangeably. If the discrete functional unit is comprised of more than one polypeptide that physically associate with one another, the term "protein" refers to the multiple polypeptides that are physically coupled and function together as the discrete unit.

[0091] *Regulatory element*: As used herein, the term "regulatory element" refers to transcriptional control elements, in particular non-coding cis-acting transcription control elements, capable of regulating and/or controlling transcription of a gene. Regulatory elements comprise at least one transcription factor binding site, for example at least one binding site for a tissue specific transcription factor. In embodiments described herein, regulatory elements have at least one binding site for a liver-specific transcription factor. Typically, regulatory elements increase or enhance promoter-driven gene expression when compared to the transcription of the gene from the promoter alone, without the regulatory elements. Thus, regulatory elements particularly comprise enhancer sequences, although it is

to be understood that the regulatory elements enhancing transcription are not limited to typical far upstream enhancer sequences, but may occur at any distance of the gene they regulate. As is understood in the art, sequences regulating transcription may be situated either upstream (e.g., in the promoter region) or downstream (e.g., in the 3'UTR) of the gene that is regulated in vivo, and may be located in the immediate vicinity of the gene or further away. Regulatory elements can comprise either naturally occurring sequences, combinations of (parts of) such regulatory elements or several copies of a regulatory element, e.g., non-naturally occurring sequences. Accordingly, regulatory elements include naturally occurring and optimized or engineered regulatory elements to achieve a desired expression level.

[0092] *Subject*: As used herein, the term "subject" refers to a human or any non-human animal (e.g., mouse, rat, rabbit, dog, cat, cattle, swine, sheep, horse or primate). A human includes pre- and post-natal forms. In many embodiments, a subject is a human being. A subject can be a patient, which refers to a human presenting to a medical provider for diagnosis or treatment of a disease. The term "subject" is used herein interchangeably with "individual" or "patient." A subject can be afflicted with or is susceptible to a disease or disorder but may or may not display symptoms of the disease or disorder.

[0093] *Substantially*: As used herein, the term "substantially" refers to the qualitative condition of exhibiting total or near-total extent or degree of a characteristic or property of interest. One of ordinary skill in the biological arts will understand that biological and chemical phenomena rarely, if ever, go to completion and/or proceed to completeness or achieve or avoid an absolute result. The term "substantially" is therefore used herein to capture the potential lack of completeness inherent in many biological and chemical phenomena.

[0094] *Substantial homology*: The phrase "substantial homology" is used herein to refer to a comparison between amino acid or nucleic acid sequences. As will be appreciated by those of ordinary skill in the art, two sequences are generally considered to be "substantially homologous" if they contain homologous residues in corresponding positions. Homologous residues may be identical residues. Alternatively, homologous residues may be non-identical residues will appropriately similar structural and/or functional characteristics. For example, as is well known by those of ordinary skill in the art, certain amino acids are typically classified as "hydrophobic" or "hydrophilic" amino acids, and/or as having "polar"

or "non-polar" side chains. Substitution of one amino acid for another of the same type may often be considered a "homologous" substitution.

[0095] As is well known in this art, amino acid or nucleic acid sequences may be compared using any of a variety of algorithms, including those available in commercial computer programs such as BLASTN for nucleotide sequences and BLASTP, gapped BLAST, and PSI-BLAST for amino acid sequences. Exemplary such programs are described in Altschul, et al., basic local alignment search tool, *J. Mol. Biol.*, 215(3): 403-410, 1990; Altschul, et al., *Methods in Enzymology*; Altschul, et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402, 1997; Baxevanis, et al., *Bioinformatics : A Practical Guide to the Analysis of Genes and Proteins*, Wiley, 1998; and Misener, et al., (eds.), *Bioinformatics Methods and Protocols* (Methods in Molecular Biology, Vol. 132), Humana Press, 1999. In addition to identifying homologous sequences, the programs mentioned above typically provide an indication of the degree of homology. In some embodiments, two sequences are considered to be substantially homologous if at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more of their corresponding residues are homologous over a relevant stretch of residues. In some embodiments, the relevant stretch is a complete sequence. In some embodiments, the relevant stretch is at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 or more residues.

[0096] *Substantial identity*: The phrase "substantial identity" is used herein to refer to a comparison between amino acid or nucleic acid sequences. As will be appreciated by those of ordinary skill in the art, two sequences are generally considered to be "substantially identical" if they contain identical residues in corresponding positions. As is well known in this art, amino acid or nucleic acid sequences may be compared using any of a variety of algorithms, including those available in commercial computer programs such as BLASTN for nucleotide sequences and BLASTP, gapped BLAST, and PSI-BLAST for amino acid sequences. Exemplary such programs are described in Altschul, et al., Basic local alignment search tool, *J. Mol. Biol.*, 215(3): 403-410, 1990; Altschul, *et al.*, *Methods in Enzymology*; Altschul *et al.*, *Nucleic Acids Res.* 25:3389-3402, 1997; Baxevanis *et al.*, Bioinformatics : A Practical Guide to the Analysis of Genes and Proteins, Wiley, 1998; and Misener, et al., (eds.), Bioinformatics Methods and Protocols (Methods in Molecular Biology, Vol. 132),

Humana Press, 1999. In addition to identifying identical sequences, the programs mentioned above typically provide an indication of the degree of identity. In some embodiments, two sequences are considered to be substantially identical if at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more of their corresponding residues are identical over a relevant stretch of residues. In some embodiments, the relevant stretch is a complete sequence. In some embodiments, the relevant stretch is at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 or more residues.

[0097] *Suffering from*: An individual who is "suffering from" a disease, disorder, and/or condition has been diagnosed with or displays one or more symptoms of the disease, disorder, and/or condition.

[0098] *Therapeutically effective amount*: As used herein, the term "therapeutically effective amount" of a therapeutic agent means an amount that is sufficient, when administered to a subject suffering from or susceptible to a disease, disorder, and/or condition, to treat, diagnose, prevent, and/or delay the onset of the symptom(s) of the disease, disorder, and/or condition. It will be appreciated by those of ordinary skill in the art that a therapeutically effective amount is typically administered via a dosing regimen comprising at least one unit dose.

[0099] *Treating*: As used herein, the term "treat," "treatment," or "treating" refers to any method used to partially or completely alleviate, ameliorate, relieve, inhibit, prevent, delay onset of, reduce severity of and/or reduce incidence of one or more symptoms or features of a particular disease, disorder, and/or condition. Treatment may be administered to a subject who does not exhibit signs of a disease and/or exhibits only early signs of the disease for the purpose of decreasing the risk of developing pathology associated with the disease.

[0100] The recitation of numerical ranges by endpoints herein includes all numbers and fractions subsumed within that range (e.g. 1 to 5 includes 1, 1.5, 2, 2.75, 3, 3.9, 4 and 5). It is also to be understood that all numbers and fractions thereof are presumed to be modified by the term "about."

[0101] Various aspects of the invention are described in detail in the following sections. The use of sections is not meant to limit the invention. Each section can apply to any aspect of the invention. In this application, the use of "or" means "and/or" unless stated otherwise. As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

[0102] Various aspects of the invention are described in detail in the following sections. The use of sections is not meant to limit the invention. Each section can apply to any aspect of the invention. In this application, the use of "or" means "and/or" unless stated otherwise.

DETAILED DESCRIPTION

[0103] The present disclosure describes efficient and robust recombinant adeno-associated virus (rAAV) vectors for the in vivo production of I2S for the treatment of diseases associated with an I2S deficiency, such as Hunter syndrome.

[0104] Mucopolysaccharidosis type II (MPS II, Hunter syndrome) is an X-chromosome-linked recessive lysosomal storage disorder that results from a deficiency in the enzyme iduronate-2-sulfatase (I2S). I2S cleaves the terminal 2-O-sulfate moieties from the glycosaminoglycans (GAG) dermatan sulfate and heparan sulfate. Due to the missing or defective I2S enzyme in patients with Hunter syndrome, GAG progressively accumulate in the lysosomes of a variety of cell types, leading to cellular engorgement, organomegaly, tissue destruction, and organ system dysfunction.

[0105] Generally, physical manifestations for people with Hunter syndrome include both somatic and neuronal symptoms. For example, in some cases of Hunter syndrome, central nervous system (CNS) involvement leads to developmental delays and nervous system problems. Symptoms such as neurodegeneration and mental retardation appear during childhood, and Hunter syndrome patients suffering from neuronal effects often die at an early age due to organ damage to the brain. Similarly, the accumulation of GAG can adversely affect the organ systems of the body. Manifesting initially as a thickening of the wall of the heart, lungs and airways, and abnormal enlargement of the liver, spleen and kidneys, these profound changes can ultimately lead to widespread catastrophic organ failure. As a result, Hunter syndrome is always severe, progressive, and life-limiting.

[0106] Enzyme replacement therapy (ERT) is an approved therapy for treating Hunter syndrome (MPS II), which involves administering exogenous replacement I2S enzyme to patients with Hunter syndrome. However, various drawbacks are associated with ERT, including, among these is for example limited distribution of I2S to various target organs.

[0107] The vectors described herein provide for robust expression of I2S in various organs, including for example liver, kidney, spleen, heart, lung, and the central nervous system. Accordingly, in some embodiments, the vectors described herein result in the expression of I2S in the liver. In some embodiments, the vectors described herein result in the expression of I2S in the kidney. In some embodiments, the vectors described herein result in the expression of I2S in the spleen. In some embodiments, the vectors described herein result in the expression of I2S in the heart. In some embodiments, the vectors described herein result in the expression of I2S in the lung. In some embodiments, the vectors described herein result in the expression of I2S in the central nervous system. In some embodiments, the vectors described herein result in the expression of I2S in the plasma.

rAAV I2S Vector Design

[0108] In some aspects, provided herewith is a recombinant adeno-associated virus (rAAV) vector encoding an iduronate-2-sulfatase (I2S) protein. A schematic that illustrates exemplary rAAV vectors of the present disclosure is illustrated in Figure 1B. As shown in Figure 1B, in some embodiments, an rAAV vector of the present disclosure comprises a liver specific promoter, a 5' and a 3' inverted terminal repeat (ITR), a cis-acting regulatory module (CRM), an intron, and a WPRE sequence.

[0109] In some embodiments, the vector also includes a sulfatase modifying factor 1 (SUMF) gene. In some embodiments, the vector comprises an internal ribosome entry site (IRES).

[0110] The iduronate-2-sulfatse (I2S) of the vector can be a wild-type or a codon-optimized variant. Accordingly, in some embodiments, the rAAV vector comprises a wild-type I2S nucleotide sequence. In some embodiments comprises a codon-optimized I2S sequence.

[0111] A suitable I2S for the present invention is any protein or a portion of a protein that can substitute for at least partial activity of naturally-occurring Iduronate-2-sulfatase (I2S) protein or rescue one or more phenotypes or symptoms associated with I2S-deficiency.

As used herein, the terms "an I2S enzyme" and "an I2S protein", and grammatical equivalents, are used inter-changeably.

[0112] Typically, the human I2S protein is produced as a precursor form. The precursor form of human I2S contains a signal peptide (amino acid residues 1-25 of the full length precursor), a pro-peptide (amino acid residues 26-33 of the full length precursor), and a chain (residues 34-550 of the full length precursor) that may be further processed into the 42 kDa chain (residues 34-455 of the full length precursor) and the 14 kDa chain (residues 446-550 of the full length precursor). Typically, the precursor form is also referred to as full-length precursor or full-length I2S protein, which contains 550 amino acids. The amino acid sequences of the mature form (SEQ ID NO: 1) having the signal peptide removed and full-length precursor (SEQ ID NO:2) of a typical wild-type or naturally-occurring human I2S protein are shown in Table 1. The signal peptide is underlined.

Table 1. Human Iduronate-2-sulfatase

Mature Form	SETQANSTTDALNVLLIIVDDLRPSLGCYGDKLVRSPNIDQL ASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNSY WRVHAGNFSTIPQYFKENGYVTMSVGKVFHPGISSNHTDD SPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVL DVPEGTLPDKQSTEQAIQLLEKMKTSASPFFLAVGYHKPHI PFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIR QREDVQALNISVPYGPIPVDFQRKIRQSYFASVSYLDTQVG RLLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFD VATHVPLIFYVPGRTASLPEAGEKLFPYLDPFDSASQLMEP GRQSMDLVELVSLFPTLAGLAGLQVPPRCPVPSFHVELCRE GKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQW NSDKPSLKDIKIMGYSIRTIDYRYTVWVGFNPDEFLANFSDI HAGELYFVDSDPLQDHNMYNDSQGGDLFQLLMP(SEQ ID NO:1)
Full-Length Precursor	MPPPRTGRGLLWLGLVLSSVCVALGSETQANSTTDALNVL LIIVDDLRPSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQ AVCAPSRVSFLTGRRPDTTRLYDFNSYWRVHAGNFSTIPQY FKENGYVTMSVGKVFHPGISSNHTDDSPYSWSFPPYHPSSE KYENTKTCRGPDGELHANLLCPVDVLDVPEGTLPDKQSTE QAIQLLEKMKTSASPFFLAVGYHKPHIPFRYPKEFQKLYPL ENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPY GPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANS TIIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRT ASLPEAGEKLFPYLDPFDSASQLMEPGRQSMDLVELVSLFP TLAGLAGLQVPPRCPVPSFHVELCREGKNLLKHFRFRDLEE DPYLPGNPRELIAYSQYPRPSDIPQWNSDKPSLKDIKIMGYS

	IRTIDYRYTVWVGFNPDEFLANFSDIHAGELYFVDSDPLQD HNMYNDSQGGDLFQLLMP(SEQ ID NO:2)

[0113] Various kinds of promoters can be used in the rAAV vector described herein. These include, for example, ubiquitous, tissue-specific, and regulatable (e.g. inducible or repressible) promoters. In some embodiments, the promoter is a liver-specific promoter. Examples of liver-specific promoters are known in the art and include, for example, human transthyrethin promoter (TTR), α-Antitrypsin promoter, human factor IX pro/liver transcription factor-responsive oligomers, LSP, and the basic albumin promoter. Liver specific promoters are described, for example, in Zhijian Wu *et al.*, *Molecular Therapy* vol 16, no 2, February 2008, the contents of which are incorporated herein by reference.

[0114] In some embodiments, the promotor is a ubiquitous promoter. In some embodiments, the promoter is a chicken beta actin promoter.

[0115] In some embodiments, the rAAV vector contains additional enhancer or regulatory elements to promote transcription and/or translation of the mRNA (e.g., enhancer sequences, Kozak sequences, polyadenylation sequences, transcriptional termination sequences, IRES and the like). In some embodiments, the vector comprises a 5' and a 3' inverted terminal repeat (ITR). In some embodiments, the vector comprises a one or more enhancer elements. In some embodiments, the vector comprises a poly(A) tail.

[0116] In some embodiments, the rAAV vector comprises one or more small elements, such as an intron. Various introns are known in the art. Suitable introns for the rAAV vector described herein include for example an MVM intron, a truncated F.IX intron, a chimeric β globin SD/immunoglobulin heavy chain SA intron, SV40 and/or an alpha globin 1st intron. In some embodiments, the rAAV vector comprises an MVM intron. In some embodiments, the rAAV vector comprises an SV40 intron.

[0117] In some embodiments, the rAAV vector comprises woodchuck hepatitis virus post-transcriptional control element (WPRE). Various optimized or variant forms of WPRE are known in the art, and include WPRE3, WPREmut6delATG among others. Other variant WPRE forms include, for example, WPRE2, WPRE_wt (GenBank accession no. J04514); WPRE_wt (GenBank accession no. J02442) and WPREmut6.

[0118] In some embodiments, the rAAV vector comprises a cis-actin regulatory module (CRM). Various kinds of CRM are suitable for use in the vectors described herein and include for example liver-specific CRM, neuronal-specific CRM. In some embodiments, the vectors described herein include a hepatocyte-specific CRM, for example, CRM8. In some embodiments, the vector includes more than one CRM. For example, in some embodiments, the vector comprises two, three, four, five or six CRM. In some embodiments, the vector comprises three CRM, for example three CRM8.

[0119] In some embodiments, the rAAV vector is sequence optimized to increase transcript stability, for more efficient translation, and to reduce immunogenicity. In some embodiments, the I2S is sequence optimized.

[0120] In some embodiments, the rAAV vector is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, or AAV11 vector. In some embodiments, the rAAV vector is AAV 1. In some embodiments, the rAAV vector is AAV 2. In some embodiments, the rAAV vector is AAV 3. In some embodiments, the rAAV vector is AAV 4. In some embodiments, the rAAV vector is AAV 5. In some embodiments, the rAAV vector is AAV 6. In some embodiments, the rAAV vector is AAV 7. In some embodiments, the rAAV vector is AAV 8. In some embodiments, the rAAV vector is AAV 9. In some embodiments, the rAAV vector is AAV 10. In some embodiments, the rAAV vector is AAV 11.

[0121] Exemplary element sequences are shown in Table 2 below. In some embodiments, the rAAV vector comprises a rAAV vector element comprising a nucleotide sequence having at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 99% identity with a vector element sequence shown in Table 2. In some embodiments, the rAAV vector comprises a vector element nucleotide sequence identical to a vector element nucleotide sequence shown in Table 2. In the table, Xn (60-100) means DNA titer tag comprising 60-100 nucleotides.

Table 2. Exemplary I2S rAAV Element Sequences

3xCRM8
ggggaggctgctggtgaatattaaccaaggtcaccccagttatcggaggagcaaacaggg gctaagtccaccgggggaggctgctggtgaatattaaccaaggtcaccccagttatcggag gagcaaacaggggctaagtccaccgggggaggctgctggtgaatattaaccaaggtcaccc cagttatcggaggagcaaacaggggctaagtccac

(SEQ ID NO: 3)

hTTR promoter

aaatgacctattaagaatatttcatagaacgaatgttccgatgctctaatctctctagaca
aggttcatatttgtatgggttacttattctctctttgttgactaagtcaataatcagaatc
agcaggtttgcagtcagattggcagggataagcagcctagctcaggagaagtgagtataaa
agccccaggctgggagcagccatcacagaagtccactcattcttggcagg

(SEQ ID NO: 4)

MVM intron

ctaaggtaagttggcgccgtttaagggatggttggttggtggggtattaatgtttaattac
ctttttacaggcctg

(SEQ ID NO: 5)

hIDS wt

atgccgccaccccggaccggccgaggccttctctggctgggtctggttctgagctccgtct
gcgtcgccctcggatccgaaacgcaggccaactcgaccacagatgctctgaacgttcttct
catcatcgtggatgacctgcgcccctccctgggctgttatggggataagctggtgaggtcc
ccaaatattgaccaactggcatcccacagcctcctcttccagaatgcctttgcgcagcaag
cagtgtgcgccccgagccgcgtttctttcctcactggcaggagacctgacaccacccgcct
gtacgacttcaactcctactggagggtgcacgctggaaacttctccaccatcccccagtac
ttcaaggagaatggctatgtgaccatgtcggtgggaaaagtctttcaccctgggatatctt
ctaaccataccgatgattctccgtatagctggtcttttccaccttatcatccttcctctga
gaagtatgaaaacactaagacatgtcgagggccagatggagaactccatgccaacctgctt
tgccctgtggatgtgctggatgttcccgagggcaccttgcctgacaaacagagcactgagc
aagccatacagttgttggaaaagatgaaaacgtcagccagtcctttcttcctggccgttgg
gtatcataagccacacatccccttcagataccccaaggaatttcagaagttgtatcccttg
gagaacatcaccctggcccccgatcccgaggtccctgatggcctacccccctgtggcctaca
acccctggatggacatcaggcaacgggaagacgtccaagccttaaacatcagtgtgccgta
tggtccaattcctgtggactttcagcggaaaatccgccagagctactttgcctctgtgtca
tatttggatacacaggtcggccgcctcttgagtgctttggacgatcttcagctggccaaca
gcaccatcattgcatttacctcggatcatgggtgggctctaggtgaacatggagaatgggc
caaatacagcaattttgatgttgctacccatgttcccctgatattctatgttcctggaagg
acggcttcacttccggaggcaggcgagaagcttttcccttacctcgacccttttgattccg
cctcacagttgatggagccaggcaggcaatccatggaccttgtggaacttgtgtctctttt
tcccacgctggctggacttgcaggactgcaggttccacctcgctgccccgttccttcattt
cacgttgagctgtgcagagaaggcaagaaccttctgaagcattttcgattccgtgacttgg
aagaggatccgtacctccctggtaatccccgtgaactgattgcctatagccagtatccccg
gccttcagacatccctcagtggaattctgacaagccgagtttaaaagatataaagatcatg
ggctattccatacgcaccatagactataggtatactgtgtgggttggcttcaatcctgatg
aatttctagctaacttttctgacatccatgcaggggaactgtattttgtggattctgaccc
attgcaggatcacaatatgtataatgattcccaaggtggagatcttttccagttgttgatg
ccttga

(SEQ ID NO: 6)

WPREmut6delATG
aatcaacctctggattacaaaatttgtgaaagattgactggtattcttaactttgttgctc cttttacgctttgtggatacgctgctttattgcctttgtatcttgctattgcttcccgttt ggctttcattttctcctccttgtataaatcctggttgctgtctctttttgaggagttgtgg cccgttgtcaggcaacgtggcgtggtgtgcactgtgtttgctgacgcaacccccactggtt ggggcattgccaccacctgtcagctcctttccgggactttcgctttcccctccctattgc cacggcggaactcatcgccgcctgccttgcccgctgctggacaggggctcggctgttgggc actgacaattccgtggtgttgtcggggaaatcatcgtcctttccttggctgctcgcctgtg ttgccacctggattctgcgcgggacgtccttctgctacgtcccttcggccctcaatccagc ggaccttccttcccgcggcctgctgccggctctgcggcctcttccgcgtcttcgccttcgc cctcagacgagtcggatctccctttgggccgcctccccgcatc (SEQ ID NO: 7)
BGH pA
cctagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcc cctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaa tgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtgggg caggacagcaaggggaggattgggaagacaatagcaggcatgctggggaa (SEQ ID NO: 8)
3' ITR
aggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggc cgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcga gcgcgcagaga (SEQ ID NO: 9)
5' ITR
ctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttg gtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaactccatcactag gggttcct (SEQ ID NO: 10)
Codon Optimized human *IDS* (COOL opt)
atgccaccacctaggacaggcaggggcctgctttggcttggactggtgctgagctctgtct gtgttgccctgggctccgagacccaagccaactctacaaccgatgctctcaatgttctgct catcatagtggatgacctgcggccctctctaggctgctatggagacaagttggtgcggagc cccaacatagaccagctagcctctcactccctgctgttccagaatgccttcgcccagcaag ctgtgtgcgcccccctctagagtgtctttcctgaccgggagaaggcctgatacaacaaggct gtatgactttaacagctactggagggtgcacgcaggcaacttctccactatcccccaatac ttcaaggagaatggctatgtgaccatgagcgtgggcaaggtcttccaccctggaatctcct ccaaccacactgatgatagtccctactcttggtcttttcctccctatcaccctagcagtga gaagtatgagaacaccaaaacctgcagaggccctgatggggagctgcatgctaacctcctg tgtcctgtagatgtgctggacgtcccagagggcaccttgccagataagcagtctactgagc aggctatccagctgcttgagaaaatgaagacttctgcatctcccttctttctggctgttgg ctaccacaagcctcacatcccccttcaggtaccctaaggagttccaaaagctctatcctctg gaaaacatcacacttgcccccgatcctgaggtccctgacggcctcccaccagtagcctaca atccttggatggacattaggcagagagaggatgtccaggctctgaatatttctgtgcccta tgggcccatcccggtggacttccagcgcaaaatcagacagtcctactttgcctctgtgagc tatctggacacccaggttgggaggctcctctccgcccttgacgacctccagttggccaaca

gcaccattatagccttcacctctgaccacggctgggcactgggggaacacggggagtgggc
taagtactctaactttgatgtggccacccacgtgcccctcatcttttatgtgcctggcagg
actgccagcctgcccgaagctggggaaaaactgtttccatacctggacccttttgacagtg
cttctcagctcatggaacctggccgtcagagcatggatctggtggagctagtgtccctctt
cccaaccttggctggccttgctggtctccaggtgcctcctagatgcccagtcccctccttc
catgttgaactctgccgtgaggggaagaatctgctgaagcacttcagattcagagacttgg
aggaggacccctaccttcctgggaaccccagggagttgattgcatactcccagtatcccag
gccaagtgacattccccagtggaactccgacaaaccaagtctgaaggacatcaagatcatg
gggtacagcatcaggaccattgactacagatacacagtgtgggttggatttaacccagatg
agttcttggcaaacttttctgacatccatgcaagtcagttgtattttgtggacagcgaccc
tctgcaggatcacaacatgtacaatgacagccagggtggggacctctttcaactcctcatg
ccatag

(SEQ ID NO: 11)

Codon Optimized human *IDS* (AUS optimized)

atgccacccccccggaccgggagaggcctcttgtggttgggcctggtgctgagcagcgtgt
gcgtggccctgggcagtgagacccaggctaactctacaacagatgccttgaatgtgctgct
gatcattgtggatgacctgaggccaagtctgggctgctatggggacaaattggtgaggtcc
cccaacatcgaccagttggcctcccactctctcctattccaaaatgctttcgcccagcagg
cagtttgtgccccctctagggtgagcttcctcactggcaggcgccctgacaccactagact
gtatgactttaacagctattggagggtgcacgcaggaaacttctccacaatccctcaatac
ttcaaggagaatggttatgtgacaatgtctgtgggcaaggtgttccaccctggcatcagca
gcaaccacaccgatgactcaccctatagttggtcttttcccccctaccatccttcatctga
gaaatatgaaaacacaaaaacctgccgaggcccagacggggaactgcatgccaacctactc
tgtcctgttgatgtactggacgtgcccgagggcaccctccctgataagcagtccacagaac
aggccattcagctgcttgaaaagatgaagacctccgcatccccttcttcttggctgtcgg
ctaccacaagccccatatccccttagataccccaaggaattccagaaactgtacccactg
gagaacatcacacttgctcctgaccctgaagtgcctgacggactgcctccagtggcctata
acccttggatggacatccggcagcgcgaggatgtgcaggctctgaacattagtgtgcctta
tgggcccatccctgtggactttcagaggaagattcgccagtcctactttgcctctgtatcc
tacctggacacacaggtgggacgcctgctgtctgcccttgatgatctgcaactggccaaca
gcaccattatagctttcacatcagaccatgggtgggctcttggggagcatggtgaatgggc
taagtactccaacttcgatgtggcaacccatgtccctctgatcttctatgtgccaggaagg
accgcctctctgccagaggcaggtgagaagctgttcccctatctggacccttttgactccg
ccagccagctgatggagcctggccgacagtctatggacctggttgagctggtcagcctgtt
tcccacactcgctggactggctggcctgcaagtacccccacgctgcccagtgccctccttc
catgtggagctttgcagggaggggaagaacctcctcaagcacttcaggttcagggacctag
aggaggatccttatctgcctggaaaccccagagagcttattgcttactcccagtatccaag
gcctagtgacattccccaatggaactcagacaaaccaagcctgaaagacatcaagatcatg
ggatactctatcaggaccattgactacaggtacactgtgtgggttggcttcaacccggatg
agttcctggctaatttctctgacatacatgctggcgagctgtacttcgtggacagtgaccc
cctgcaggatcacaacatgtacaatgattcccaggggggtgacctcttccagcttctgatg
ccctaa

(SEQ ID NO: 12)

IRES

gcccctctccctcccccccccctaacgttactggccgaagccgcttggaataaggccggtg
tgcgtttgtctatatgttattttccaccatattgccgtcttttggcaatgtgagggcccgg
aaacctggccctgtcttcttgacgagcattcctaggggtctttcccctctcgccaaaggaa

tgcaaggtctgttgaatgtcgtgaaggaagcagttcctctggaagcttcttgaagacaaac
aacgtctgtagcgacccttgcaggcagcggaacccccacctggcgacaggtgcctctgc
ggccaaaagccacgtgtataagatacacctgcaaaggcggcacaaccccagtgccacgttg
tgagttggatagttgtggaaagagtcaaatggctcacctcaagcgtattcaacaaggggct
gaaggatgcccagaaggtaccccattgtatgggatctgatctggggcctcggtgcacatgc
tttacatgtgtttagtcgaggttaaaaaacgtctaggccccccgaaccacggggacgtggt
ttttcctttgaaaaacacgatgataat

(SEQ ID NO: 13)

SUMF1 wild type

atggctgcgcccgcactagggctggtgtgtggacgttgccctgagctgggtctcgtcctct
tgctgctgctgctctcgctgctgtgtggagcggcagggagccaggaggccgggaccggtgc
gggcgcggggtcccttgcgggttcttgcggctgcggcacgccccagcggcctggcgcccat
ggcagttcggcagccgctcaccgatactcgcgggaggctaacgctccgggccccgtacccg
gagagcggcaactcgcgcactcaaagatggtccccatccctgctggagtatttacaatggg
cacagatgatcctcagataaagcaggatggggaagcacctgcgaggagagttactattgat
gcctttacatggatgcctatgaagtcagtaatactgaatttgagaagtttgtgaactcaa
ctggctatttgacagaggctgagaagtttggcgactcctttgtctttgaaggcatgttgag
tgagcaagtgaagaccaatattcaacaggcagttgcagctgctccctggtggttacctgtg
aaaggcgctaactggagacacccagaagggcctgactctactattctgcacaggccggatc
atccagttctccatgtgtcctggaatgatgcggttgcctactgcacttgggcagggaagcg
gctgcccacggaagctgagtgggaatacagctgtcgaggaggcctgcataatagacttttc
ccctggggcaacaaactgcagcccaaaggccagcattatgccaacatttggcagggcgagt
ttccggtgaccaacactggtgaggatggcttccaaggaactgcgcctgttgatgccttccc
tcccaatggttatggcttatacaacatagtggggaacgcatgggaatggacttcagactgg
tggactgttcatcattctgttgaagaaacgcttaacccaaaaggtccccttctgggaaag
accgagtgaagaaaggtggatcctacatgtgccataggtcttattgttacaggtatcgctg
tgctgctcggagccagaacacacctgatagctctgcttcgaatctgggattccgctgtgca
gccgaccgcctgcccactatggactga

(SEQ ID NO: 14)

Codon Optimized _SUMF-1_ (COOL optimized)

atggctgctcctgccctggggctggtgtgtggaagatgtcctgaactgggcctggttctgt
tactgcttctgctcagcctgctctgtggtgctgccggcagccaagaggcaggcactggcgc
tggagctggaagcctggctgggtcttgtggatgtggcacaccacagaggccaggggctcat
ggctcctctgctgcagctcataggtacagcagagaagccaatgctccaggcccagtgcctg
gagagagacagctggctcacagcaagatggtgcccatccctgctggggtgttcacaatggg
aacagatgatccccagatcaagcaggatggggaggcgcctgccaggagggtgaccattgat
gcattctatatggatgcctatgaggtgagcaatacagaatttgagaagtttgtgaactcta
ctggctacctgactgaggctgaaaaatttggagactcttttgtgtttgaaggaatgcttag
tgaacaggttaagaccaacatccagcaggctgttgcagcagccccctggtggttgcctgtc
aagggagctaactggaggcaccctgagggaccagattctacaatcctgcatagacctgatc
atcctgttctgcatgtgtcttggaatgatgctgtggcttactgtacctgggcaggaaaaag
gctgccaacagaagctgagtgggaatactcttgcagaggaggcctgcacaatagactgttc
ccatggggcaacaagctgcaacccaagggccagcactatgctaacatctggcagggagaat
tccctgtgacaaacacaggagaggacggcttccagggaactgcccctgtagatgctttccc
tcctaatggctatggcctgtataacattgttggcaacgcctgggagtggacttctgattgg
tggacagtgcaccactctgttgaggagacactgaatcctaaggggccaccttctggaaagg
atagagtgaagaaggggggaagctacatgtgccacaggtcttattgttacagatacaggtg

cgctgctaggtctcagaacacccctgatagcagtgctagcaatctgggcttcaggtgtgcc gctgacagactgcctaccatggattaa (SEQ ID NO: 15)
Codon Optimized *SUMF-1* (AUS optimized)
atggctgcccctgctctgggattggtttgtggcagatgtcctgagcttggtctggtgctgt tgctccttctgttgtctctgctgtggagcagctgggtctcaggaagctggcacaggcgc tggggctggctctctggccgggtcatgtggctgtggaactccccagcggcctggagcccat ggcagctctgccgcagcacacaggtattctagggaagccaatgccccaggccctgtgcctg gggagagacagctagctcattctaagatggtgcctatcccagccggggtttttacaatggg cactgatgatcctcagattaagcaggatggagaggcccccgccagaagagtgaccattgat gctttctacatggatgcatatgaagtgtccaacacagagtttgagaaatttgtgaactcta ctggatacttgaccgaggctgagaagtttggagattcctttgtctttgaaggcatgctgtc tgagcaggtcaagaccaacattcagcaagcagtggccgctgcaccttggtggcttcctgtg aagggcgccaactggagacatccagaggggccagatagtaccatcctccacagacctgatc acccagtccttcatgtttcctggaatgatgcagttgcttactgcacttgggccggcaagag gctccctactgaggcagagtgggaatactcctgcagaggaggcctgcacaacagactgttc ccttgggggaacaagcttcagcccaaaggccagcactatgctaacatctggcagggtgagt ttccagtcaccaatacaggggaggacggattccagggaaccgcaccagtagatgccttccc tcctaatggctatggcctgtataatattgtgggcaatgcatgggagtggacctctgactgg tggactgtgcaccactcagtggaggaaaccctgaaccctaagggacccccttcaggcaaag atagagtcaaaaagggagggagctatatgtgtcacagatcctattgctacagatatagatg tgcagccaggtcccagaacacccctgactcttctgctagcaacctgggctttcggtgtgct gctgatagactgcccaccatggactaa (SEQ ID NO: 16)

[0122] In some embodiments, the rAAV I2S vector comprises a nucleotide sequence having at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 99% identity with a nucleotide sequence shown in Table 3 below. In some embodiment, the rAAV I2S vector comprises a sequence identical to a nucleotide sequence shown in Table 3 below.

Table 3: Exemplary rAAV I2S vector nucleotide sequences

pXL024
CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA

AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGCCGCCACCCCGGACCGGCCGAGGCCTTCTCT
GGCTGGGTCTGGTTCTGAGCTCCGTCTGCGTCGCCCTCGGATCCGAAACGCAGGCCAACTC
GACCACAGATGCTCTGAACGTTCTTCTCATCATCGTGGATGACCTGCGCCCCTCCCTGGGC
TGTTATGGGGATAAGCTGGTGAGGTCCCCAAATATTGACCAACTGGCATCCCACAGCCTCC
TCTTCCAGAATGCCTTTGCGCAGCAAGCAGTGTGCGCCCCGAGCCGCGTTTCTTTCCTCAC
TGGCAGGAGACCTGACACCACCCGCCTGTACGACTTCAACTCCTACTGGAGGGTGCACGCT
GGAAACTTCTCCACCATCCCCCAGTACTTCAAGGAGAATGGCTATGTGACCATGTCGGTGG
GAAAAGTCTTTCACCCTGGGATATCTTCTAACCATACCGATGATTCTCCGTATAGCTGGTC
TTTTCCACCTTATCATCCTTCCTCTGAGAAGTATGAAAACACTAAGACATGTCGAGGGCCA
GATGGAGAACTCCATGCCAACCTGCTTTGCCCTGTGGATGTGCTGGATGTTCCCGAGGGCA
CCTTGCCTGACAAACAGAGCACTGAGCAAGCCATACAGTTGTTGGAAAAGATGAAAACGTC
AGCCAGTCCTTTCTTCCTGGCCGTTGGGTATCATAAGCCACACATCCCCTTCAGATACCCC
AAGGAATTTCAGAAGTTGTATCCCTTGGAGAACATCACCCTGGCCCCCGATCCCGAGGTCC
CTGATGGCCTACCCCCTGTGGCCTACAACCCCTGGATGGACATCAGGCAACGGGAAGACGT
CCAAGCCTTAAACATCAGTGTGCCGTATGGTCCAATTCCTGTGGACTTTCAGCGGAAAATC
CGCCAGAGCTACTTTGCCTCTGTGTCATATTTGGATACACAGGTCGGCCGCCTCTTGAGTG
CTTTGGACGATCTTCAGCTGGCCAACAGCACCATCATTGCATTTACCTCGGATCATGGGTG
GGCTCTAGGTGAACATGGAGAATGGGCCAAATACAGCAATTTTGATGTTGCTACCCATGTT
CCCCTGATATTCTATGTTCCTGGAAGGACGGCTTCACTTCCGGAGGCAGGCGAGAAGCTTT
TCCCTTACCTCGACCCTTTTGATTCCGCCTCACAGTTGATGGAGCCAGGCAGGCAATCCAT
GGACCTTGTGGAACTTGTGTCTCTTTTTCCCACGCTGGCTGGACTTGCAGGACTGCAGGTT
CCACCTCGCTGCCCCGTTCCTTCATTTCACGTTGAGCTGTGCAGAGAAGGCAAGAACCTTC
TGAAGCATTTTCGATTCCGTGACTTGGAAGAGGATCCGTACCTCCCTGGTAATCCCCGTGA
ACTGATTGCCTATAGCCAGTATCCCCGGCCTTCAGACATCCCTCAGTGGAATTCTGACAAG
CCGAGTTTAAAAGATATAAAGATCATGGGCTATTCCATACGCACCATAGACTATAGGTATA
CTGTGTGGGTTGGCTTCAATCCTGATGAATTTCTAGCTAACTTTTCTGACATCCATGCAGG
GGAACTGTATTTTGTGGATTCTGACCCATTGCAGGATCACAATATGTATAATGATTCCCAA
GGTGGAGATCTTTTCCAGTTGTTGATGCCTTGAGGTACCAATCAACCTCTGGATTACAAAA
TTTGTGAAAGATTGACTGGTATTCTTAACTTTGTTGCTCCTTTTACGCTTTGTGGATACGC
TGCTTTATTGCCTTTGTATCTTGCTATTGCTTCCCGTTTGGCTTTCATTTTCTCCTCCTTG
TATAAATCCTGGTTGCTGTCTCTTTTTGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCG
TGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCA
GCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCC
TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGT
CGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGG
GACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTG
CTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCC
TTTGGGCCGCCTCCCCGCATCGGTACCGTCGACCCTAGAGCTCGCTGATCAGCCTCGACTG
TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGA
AGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGT
AGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAG
ACAATAGCAGGCATGCTGGGGAATCTAGA Xn(60-100)
GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGC
CTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCTGGGG
GGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTCTTTG
CAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTCGGAC
CTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGCCATC

TGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGAGATC
CCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCACGTG
TCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTGCGCA
GCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAACCGG
GAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCCTGCC
TCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCCGATT
TCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTTTATG
GAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAGCCGG
GAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTCTGGT
GGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGCCTAG
CGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTGGGCG
CCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCGCACC
CTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGGGCTC
TCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTCTGGA
AAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTAAAAT
ACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTAGAGA
AGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTTTGAT
GATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGCAAGC
ATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGACTGTG
GGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTGTTGG
GCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGAAGCT
CAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATTCGCC
CTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCCTCAT
GCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAAATTT
ATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATC
GTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTAT
TTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGA
TGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCG
TTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGC
CCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAG
CGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCC
GAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCC
ATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACT
CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA
GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACTA
GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGC
CGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCAAAGT
GCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTA

ATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTG
ATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAGATAT
AGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAAC
CGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGG
TAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCGAAAA
GATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGAT
GCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTG
CTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAG
GCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACC
ACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTCCTCG
GCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCC
GACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTG
TCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGG
CCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACATATT
CTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAA
TATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACAAGGG
TGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGATGAG
CATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTT
TACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCA
ACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTAT
ATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCATCAAA
TGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTC
TGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGG
TTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTAT
GCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGC
ATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATCGCTG
TTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGCGCAT
CAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTCCGGG
GATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGA
AGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAA
CGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATA
GATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCA
TCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTTTTTT
CCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAGCAAT
AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC
CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA
ACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATT
CAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC
CACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG
CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATA
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG

AAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGA
GCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGG
TGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGT
GGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTC
GAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGGGGCT
CGAG
(SEQ ID NO: 17)

pXL025

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA
GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC
AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG
TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA
GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA
CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA
TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA
AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA
AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGCCACCACCTAGGACAGGCAGGGGCCTGCTTT
GGCTTGGACTGGTGCTGAGCTCTGTCTGTGTTGCCCTGGGCTCCGAGACCCAAGCCAACTC
TACAACCGATGCTCTCAATGTTCTGCTCATCATAGTGGATGACCTGCGGCCCTCTCTAGGC
TGCTATGGAGACAAGTTGGTGCGGAGCCCCAACATAGACCAGCTAGCCTCTCACTCCCTGC
TGTTCCAGAATGCCTTCGCCCAGCAAGCTGTGTGCGCCCCCTCTAGAGTGTCTTTCCTGAC
CGGGAGAAGGCCTGATACAACAAGGCTGTATGACTTTAACAGCTACTGGAGGGTGCACGCA
GGCAACTTCTCCACTATCCCCCAATACTTCAAGGAGAATGGCTATGTGACCATGAGCGTGG
GCAAGGTCTTCCACCCTGGAATCTCCTCCAACCACACTGATGATAGTCCCTACTCTTGGTC
TTTTCCTCCCTATCACCCTAGCAGTGAGAAGTATGAGAACACCAAAACCTGCAGAGGCCCT
GATGGGGAGCTGCATGCTAACCTCCTGTGTCCTGTAGATGTGCTGGACGTCCCAGAGGGCA
CCTTGCCAGATAAGCAGTCTACTGAGCAGGCTATCCAGCTGCTTGAGAAAATGAAGACTTC
TGCATCTCCCTTCTTTCTGGCTGTTGGCTACCACAAGCCTCACATCCCCTTCAGGTACCCT
AAGGAGTTCCAAAAGCTCTATCCTCTGGAAAACATCACACTTGCCCCCGATCCTGAGGTCC
CTGACGGCCTCCCACCAGTAGCCTACAATCCTTGGATGGACATTAGGCAGAGAGAGGATGT
CCAGGCTCTGAATATTTCTGTGCCCTATGGGCCCATCCCGGTGGACTTCCAGCGCAAAATC
AGACAGTCCTACTTTGCCTCTGTGAGCTATCTGGACACCCAGGTTGGGAGGCTCCTCTCCG
CCCTTGACGACCTCCAGTTGGCCAACAGCACCATTATAGCCTTCACCTCTGACCACGGCTG
GGCACTGGGGGAACACGGGGAGTGGGCTAAGTACTCTAACTTTGATGTGGCCACCCACGTG
CCCCTCATCTTTTATGTGCCTGGCAGGACTGCCAGCCTGCCCGAAGCTGGGGAAAAACTGT
TTCCATACCTGGACCCTTTTGACAGTGCTTCTCAGCTCATGGAACCTGGCCGTCAGAGCAT
GGATCTGGTGGAGCTAGTGTCCCTCTTCCCAACCTTGGCTGGCCTTGCTGGTCTCCAGGTG
CCTCCTAGATGCCCAGTCCCCTCCTTCCATGTTGAACTCTGCCGTGAGGGGAAGAATCTGC
TGAAGCACTTCAGATTCAGAGACTTGGAGGAGGACCCCTACCTTCCTGGGAACCCCAGGGA
GTTGATTGCATACTCCCAGTATCCCAGGCCAAGTGACATTCCCCAGTGGAACTCCGACAAA
CCAAGTCTGAAGGACATCAAGATCATGGGGTACAGCATCAGGACCATTGACTACAGATACA
CAGTGTGGGTTGGATTTAACCCAGATGAGTTCTTGGCAAACTTTTCTGACATCCATGCAAG
TCAGTTGTATTTTGTGGACAGCGACCCTCTGCAGGATCACAACATGTACAATGACAGCCAG
GGTGGGGACCTCTTTCAACTCCTCATGCCATAGCAATTGAATCAACCTCTGGATTACAAAA
TTTGTGAAAGATTGACTGGTATTCTTAACTTTGTTGCTCCTTTTACGCTTTGTGGATACGC

TGCTTTATTGCCTTTGTATCTTGCTATTGCTTCCCGTTTGGCTTTCATTTTCTCCTCCTTG
TATAAATCCTGGTTGCTGTCTCTTTTTGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCG
TGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCA
GCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCC
TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGT
CGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGG
GACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTG
CTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCC
TTTGGGCCGCCTCCCCGCATCCAATTGGTCGACCCTAGAGCTCGCTGATCAGCCTCGACTG
TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGA
AGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGT
AGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAG
ACAATAGCAGGCATGCTGGGGAATCTAGA Xn(60-100)
GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGC
CTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCTGGGG
GGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTCTTTG
CAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTCGGAC
CTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGCCATC
TGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGAGATC
CCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCACGTG
TCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTGCGCA
GCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAACCGG
GAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCCTGCC
TCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCCGATT
TCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTTTATG
GAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAGCCGG
GAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTCTGGT
GGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGCCTAG
CGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTGGGCG
CCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCGCACC
CTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGGGCTC
TCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTCTGGA
AAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTAAAAT
ACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTAGAGA
AGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTTTGAT
GATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGCAAGC
ATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGACTGTG
GGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTGTTGG
GCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGAAGCT
CAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATTCGCC
CTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCCTCAT
GCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAAATTT
ATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATC
GTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTAT
TTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGA
TGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCG
TTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGC
CCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAG

CGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCC
GAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCC
ATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACT
CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA
GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACTA
GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGC
CGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCAAAGT
GCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTA
ATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTG
ATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAGATAT
AGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAAC
CGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGG
TAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCGAAAA
GATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGAT
GCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTG
CTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAG
GCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACC
ACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTCCTCG
GCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCC
GACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTG
TCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGG
CCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACATATT
CTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAA
TATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACAAGGG
TGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGATGAG
CATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTT
TACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCA
ACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTAT
ATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCATCAAA
TGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTC
TGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGG
TTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTAT
GCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGC
ATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATCGCTG
TTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGCGCAT
CAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTCCGGG
GATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGA

AGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAA
CGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATA
GATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCA
TCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTTTTTT
CCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAGCAAT
AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC
CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA
ACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATT
CAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC
CACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG
CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATA
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG
AAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGA
GCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGG
TGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGT
GGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTC
GAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGGGGCT
CGAG

pXL026

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA
GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC
AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG
TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA
GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA
CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA
TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA
AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA
AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGCCACCCCCCCGGACCGGGAGAGGCCTCTTGT
GGTTGGGCCTGGTGCTGAGCAGCGTGTGCGTGGCCCTGGGCAGTGAGACCCAGGCTAACTC
TACAACAGATGCCTTGAATGTGCTGCTGATCATTGTGGATGACCTGAGGCCAAGTCTGGGC
TGCTATGGGGACAAATTGGTGAGGTCCCCCAACATCGACCAGTTGGCCTCCCACTCTCTCC
TATTCCAAAATGCTTTCGCCCAGCAGGCAGTTTGTGCCCCCTCTAGGGTGAGCTTCCTCAC
TGGCAGGCGCCCTGACACCACTAGACTGTATGACTTTAACAGCTATTGGAGGGTGCACGCA
GGAAACTTCTCCACAATCCCTCAATACTTCAAGGAGAATGGTTATGTGACAATGTCTGTGG
GCAAGGTGTTCCACCCTGGCATCAGCAGCAACCACACCGATGACTCACCCTATAGTTGGTC
TTTTCCCCCCTACCATCCTTCATCTGAGAAATATGAAAACACAAAAACCTGCCGAGGCCCA
GACGGGGAACTGCATGCCAACCTACTCTGTCCTGTTGATGTACTGGACGTGCCCGAGGGCA
CCCTCCCTGATAAGCAGTCCACAGAACAGGCCATTCAGCTGCTTGAAAAGATGAAGACCTC

CGCATCCCCCTTCTTCTTGGCTGTCGGCTACCACAAGCCCCATATCCCCTTTAGATACCCC
AAGGAATTCCAGAAACTGTACCCACTGGAGAACATCACACTTGCTCCTGACCCTGAAGTGC
CTGACGGACTGCCTCCAGTGGCCTATAACCCTTGGATGGACATCCGGCAGCGCGAGGATGT
GCAGGCTCTGAACATTAGTGTGCCTTATGGGCCCATCCCTGTGGACTTTCAGAGGAAGATT
CGCCAGTCCTACTTTGCCTCTGTATCCTACCTGGACACACAGGTGGGACGCCTGCTGTCTG
CCCTTGATGATCTGCAACTGGCCAACAGCACCATTATAGCTTTCACATCAGACCATGGGTG
GGCTCTTGGGGAGCATGGTGAATGGGCTAAGTACTCCAACTTCGATGTGGCAACCCATGTC
CCTCTGATCTTCTATGTGCCAGGAAGGACCGCCTCTCTGCCAGAGGCAGGTGAGAAGCTGT
TCCCCTATCTGGACCCTTTTGACTCCGCCAGCCAGCTGATGGAGCCTGGCCGACAGTCTAT
GGACCTGGTTGAGCTGGTCAGCCTGTTTCCCACACTCGCTGGACTGGCTGGCCTGCAAGTA
CCCCCACGCTGCCCAGTGCCCTCCTTCCATGTGGAGCTTTGCAGGGAGGGGAAGAACCTCC
TCAAGCACTTCAGGTTCAGGGACCTAGAGGAGGATCCTTATCTGCCTGGAAACCCCAGAGA
GCTTATTGCTTACTCCCAGTATCCAAGGCCTAGTGACATTCCCCAATGGAACTCAGACAAA
CCAAGCCTGAAAGACATCAAGATCATGGGATACTCTATCAGGACCATTGACTACAGGTACA
CTGTGTGGGTTGGCTTCAACCCGGATGAGTTCCTGGCTAATTTCTCTGACATACATGCTGG
CGAGCTGTACTTCGTGGACAGTGACCCCCTGCAGGATCACAACATGTACAATGATTCCCAG
GGGGGTGACCTCTTCCAGCTTCTGATGCCCTAAGGTACCAATCAACCTCTGGATTACAAAA
TTTGTGAAAGATTGACTGGTATTCTTAACTTTGTTGCTCCTTTTACGCTTTGTGGATACGC
TGCTTTATTGCCTTTGTATCTTGCTATTGCTTCCCGTTTGGCTTTCATTTTCTCCTCCTTG
TATAAATCCTGGTTGCTGTCTCTTTTTGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCG
TGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCA
GCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCC
TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGT
CGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGG
GACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTG
CTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCC
TTTGGGCCGCCTCCCCGCATCGGTACCGTCGACCCTAGAGCTCGCTGATCAGCCTCGACTG
TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGA
AGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGT
AGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAG
ACAATAGCAGGCATGCTGGGGAATCTAGA Xn(60-100)
GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGC
CTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCTGGGG
GGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTCTTTG
CAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTCGGAC
CTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGCCATC
TGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGAGATC
CCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCACGTG
TCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTGCGCA
GCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAACCGG
GAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCCTGCC
TCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCCGATT
TCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTTTATG
GAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAGCCGG
GAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTCTGGT
GGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGCCTAG
CGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTGGGCG
CCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCGCACC

CTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGGGCTC
TCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTCTGGA
AAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTAAAAT
ACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTAGAGA
AGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTTTGAT
GATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGCAAGC
ATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGACTGTG
GGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTGTTGG
GCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGAAGCT
CAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATTCGCC
CTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCCTCAT
GCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAAATTT
ATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATC
GTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTAT
TTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGA
TGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCG
TTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGC
CCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAG
CGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCC
GAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCC
ATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACT
CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA
GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACTA
GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGC
CGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCAAAGT
GCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTA
ATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTG
ATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAGATAT
AGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAAC
CGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGG
TAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCGAAAA
GATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGAT
GCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTG
CTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAG
GCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACC
ACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTCCTCG
GCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCC
GACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTG

TCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGG
CCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACATATT
CTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAA
TATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACAAGGG
TGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGATGAG
CATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTT
TACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCA
ACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTAT
ATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCATCAAA
TGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTC
TGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGG
TTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTAT
GCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGC
ATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATCGCTG
TTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGCGCAT
CAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTCCGGG
GATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGA
AGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAA
CGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATA
GATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCA
TCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTTTTTT
CCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAGCAAT
AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC
CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA
ACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATT
CAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC
CACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG
CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATA
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG
AAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGA
GCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGG
TGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGT
GGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTC
GAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGGGGCT
CGAG
(SEQ ID NO: 18)

pXL027

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA

GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC
AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG
TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA
GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA
CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA
TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA
AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA
AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGCCGCCACCCCGGACCGGCCGAGGCCTTCTCT
GGCTGGGTCTGGTTCTGAGCTCCGTCTGCGTCGCCCTCGGATCCGAAACGCAGGCCAACTC
GACCACAGATGCTCTGAACGTTCTTCTCATCATCGTGGATGACCTGCGCCCCTCCCTGGGC
TGTTATGGGGATAAGCTGGTGAGGTCCCCAAATATTGACCAACTGGCATCCCACAGCCTCC
TCTTCCAGAATGCCTTTGCGCAGCAAGCAGTGTGCGCCCCGAGCCGCGTTTCTTTCCTCAC
TGGCAGGAGACCTGACACCACCCGCCTGTACGACTTCAACTCCTACTGGAGGGTGCACGCT
GGAAACTTCTCCACCATCCCCCAGTACTTCAAGGAGAATGGCTATGTGACCATGTCGGTGG
GAAAAGTCTTTCACCCTGGGATATCTTCTAACCATACCGATGATTCTCCGTATAGCTGGTC
TTTTCCACCTTATCATCCTTCCTCTGAGAAGTATGAAAACACTAAGACATGTCGAGGGCCA
GATGGAGAACTCCATGCCAACCTGCTTTGCCCTGTGGATGTGCTGGATGTTCCCGAGGGCA
CCTTGCCTGACAAACAGAGCACTGAGCAAGCCATACAGTTGTTGGAAAAGATGAAAACGTC
AGCCAGTCCTTTCTTCCTGGCCGTTGGGTATCATAAGCCACACATCCCCTTCAGATACCCC
AAGGAATTTCAGAAGTTGTATCCCTTGGAGAACATCACCCTGGCCCCCGATCCCGAGGTCC
CTGATGGCCTACCCCCTGTGGCCTACAACCCCTGGATGGACATCAGGCAACGGGAAGACGT
CCAAGCCTTAAACATCAGTGTGCCGTATGGTCCAATTCCTGTGGACTTTCAGCGGAAAATC
CGCCAGAGCTACTTTGCCTCTGTGTCATATTTGGATACACAGGTCGGCCGCCTCTTGAGTG
CTTTGGACGATCTTCAGCTGGCCAACAGCACCATCATTGCATTTACCTCGGATCATGGGTG
GGCTCTAGGTGAACATGGAGAATGGGCCAAATACAGCAATTTTGATGTTGCTACCCATGTT
CCCCTGATATTCTATGTTCCTGGAAGGACGGCTTCACTTCCGGAGGCAGGCGAGAAGCTTT
TCCCTTACCTCGACCCTTTTGATTCCGCCTCACAGTTGATGGAGCCAGGCAGGCAATCCAT
GGACCTTGTGGAACTTGTGTCTCTTTTTCCCACGCTGGCTGGACTTGCAGGACTGCAGGTT
CCACCTCGCTGCCCCGTTCCTTCATTTCACGTTGAGCTGTGCAGAGAAGGCAAGAACCTTC
TGAAGCATTTTCGATTCCGTGACTTGGAAGAGGATCCGTACCTCCCTGGTAATCCCCGTGA
ACTGATTGCCTATAGCCAGTATCCCCGGCCTTCAGACATCCCTCAGTGGAATTCTGACAAG
CCGAGTTTAAAAGATATAAAGATCATGGGCTATTCCATACGCACCATAGACTATAGGTATA
CTGTGTGGGTTGGCTTCAATCCTGATGAATTTCTAGCTAACTTTTCTGACATCCATGCAGG
GGAACTGTATTTTGTGGATTCTGACCCATTGCAGGATCACAATATGTATAATGATTCCCAA
GGTGGAGATCTTTTCCAGTTGTTGATGCCTTGACAATTGGCCCCTCTCCCTCCCCCCCCCC
TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTT
TCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGA
CGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGT
GAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGC
AGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAG
ATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAG
AGTCAAATGGCTCACCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCC
CATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGT
TAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGA
TAATCATATGGCCACCATGGCTGCGCCCGCACTAGGGCTGGTGTGTGGACGTTGCCCTGAG
CTGGGTCTCGTCCTCTTGCTGCTGCTGCTCTCGCTGCTGTGTGGAGCGGCAGGGAGCCAGG
AGGCCGGGACCGGTGCGGGCGCGGGGTCCCTTGCGGGTTCTTGCGGCTGCGGCACGCCCCA
GCGGCCTGGCGCCCATGGCAGTTCGGCAGCCGCTCACCGATACTCGCGGGAGGCTAACGCT

CCGGGCCCCGTACCCGGAGAGCGGCAACTCGCGCACTCAAAGATGGTCCCCATCCCTGCTG
GAGTATTTACAATGGGCACAGATGATCCTCAGATAAAGCAGGATGGGGAAGCACCTGCGAG
GAGAGTTACTATTGATGCCTTTTACATGGATGCCTATGAAGTCAGTAATACTGAATTTGAG
AAGTTTGTGAACTCAACTGGCTATTTGACAGAGGCTGAGAAGTTTGGCGACTCCTTTGTCT
TTGAAGGCATGTTGAGTGAGCAAGTGAAGACCAATATTCAACAGGCAGTTGCAGCTGCTCC
CTGGTGGTTACCTGTGAAAGGCGCTAACTGGAGACACCCAGAAGGGCCTGACTCTACTATT
CTGCACAGGCCGGATCATCCAGTTCTCCATGTGTCCTGGAATGATGCGGTTGCCTACTGCA
CTTGGGCAGGGAAGCGGCTGCCCACGGAAGCTGAGTGGGAATACAGCTGTCGAGGAGGCCT
GCATAATAGACTTTTCCCCTGGGGCAACAAACTGCAGCCCAAAGGCCAGCATTATGCCAAC
ATTTGGCAGGGCGAGTTTCCGGTGACCAACACTGGTGAGGATGGCTTCCAAGGAACTGCGC
CTGTTGATGCCTTCCCTCCCAATGGTTATGGCTTATACAACATAGTGGGGAACGCATGGGA
ATGGACTTCAGACTGGTGGACTGTTCATCATTCTGTTGAAGAAACGCTTAACCCAAAAGGT
CCCCCTTCTGGGAAAGACCGAGTGAAGAAAGGTGGATCCTACATGTGCCATAGGTCTTATT
GTTACAGGTATCGCTGTGCTGCTCGGAGCCAGAACACACCTGATAGCTCTGCTTCGAATCT
GGGATTCCGCTGTGCAGCCGACCGCCTGCCCACTATGGACTGAGTCGACCCTAGAGCTCGC
TGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC
CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC
ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG
GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAATCTAGA Xn(60-100)
GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGC
CTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCTGGGG
GGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTCTTTG
CAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTCGGAC
CTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGCCATC
TGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGAGATC
CCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCACGTG
TCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTGCGCA
GCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAACCGG
GAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCCTGCC
TCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCCGATT
TCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTTTATG
GAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAGCCGG
GAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTCTGGT
GGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGCCTAG
CGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTGGGCG
CCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCGCACC
CTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGGGCTC
TCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTCTGGA
AAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTAAAAT
ACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTAGAGA
AGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTTTGAT
GATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGCAAGC
ATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGACTGTG
GGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTGTTGG
GCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGAAGCT
CAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATTCGCC
CTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCCTCAT
GCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAAATTT

ATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATC
GTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTAT
TTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGA
TGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCG
TTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGC
CCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAG
CGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCC
GAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCC
ATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACT
CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA
GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACTA
GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGC
CGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCAAAGT
GCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTA
ATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTG
ATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAGATAT
AGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAAC
CGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGG
TAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCGAAAA
GATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGAT
GCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTG
CTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAG
GCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACC
ACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTCCTCG
GCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCC
GACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTG
TCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGG
CCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACATATT
CTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAA
TATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACAAGGG
TGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGATGAG
CATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTT
TACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCA
ACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTAT
ATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCATCAAA
TGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTC
TGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGG

TTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTAT
GCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGC
ATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATCGCTG
TTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGCGCAT
CAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTCCGGG
GATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGA
AGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAA
CGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATA
GATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCA
TCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTTTTTT
CCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAGCAAT
AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC
CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA
ACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATT
CAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC
CACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG
CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATA
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG
AAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGA
GCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGG
TGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGT
GGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTC
GAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGGGGCT
CGAG
(SEQ ID NO: 19)

pXL028

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA
GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC
AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG
TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA
GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA
CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA
TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA
AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA
AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGCCACCACCTAGGACAGGCAGGGGCCTGCTTT
GGCTTGGACTGGTGCTGAGCTCTGTCTGTGTTGCCCTGGGCTCCGAGACCCAAGCCAACTC
TACAACCGATGCTCTCAATGTTCTGCTCATCATAGTGGATGACCTGCGGCCCTCTCTAGGC
TGCTATGGAGACAAGTTGGTGCGGAGCCCCAACATAGACCAGCTAGCCTCTCACTCCCTGC

TGTTCCAGAATGCCTTCGCCCAGCAAGCTGTGTGCGCCCCCTCTAGAGTGTCTTTCCTGAC
CGGGAGAAGGCCTGATACAACAAGGCTGTATGACTTTAACAGCTACTGGAGGGTGCACGCA
GGCAACTTCTCCACTATCCCCCAATACTTCAAGGAGAATGGCTATGTGACCATGAGCGTGG
GCAAGGTCTTCCACCCTGGAATCTCCTCCAACCACACTGATGATAGTCCCTACTCTTGGTC
TTTTCCTCCCTATCACCCTAGCAGTGAGAAGTATGAGAACACCAAAACCTGCAGAGGCCCT
GATGGGGAGCTGCATGCTAACCTCCTGTGTCCTGTAGATGTGCTGGACGTCCCAGAGGGCA
CCTTGCCAGATAAGCAGTCTACTGAGCAGGCTATCCAGCTGCTTGAGAAAATGAAGACTTC
TGCATCTCCCTTCTTTCTGGCTGTTGGCTACCACAAGCCTCACATCCCCTTCAGGTACCCT
AAGGAGTTCCAAAAGCTCTATCCTCTGGAAAACATCACACTTGCCCCCGATCCTGAGGTCC
CTGACGGCCTCCCACCAGTAGCCTACAATCCTTGGATGGACATTAGGCAGAGAGAGGATGT
CCAGGCTCTGAATATTTCTGTGCCCTATGGGCCCATCCCGGTGGACTTCCAGCGCAAAATC
AGACAGTCCTACTTTGCCTCTGTGAGCTATCTGGACACCCAGGTTGGGAGGCTCCTCTCCG
CCCTTGACGACCTCCAGTTGGCCAACAGCACCATTATAGCCTTCACCTCTGACCACGGCTG
GGCACTGGGGGAACACGGGGAGTGGGCTAAGTACTCTAACTTTGATGTGGCCACCCACGTG
CCCCTCATCTTTTATGTGCCTGGCAGGACTGCCAGCCTGCCCGAAGCTGGGGAAAAACTGT
TTCCATACCTGGACCCTTTTGACAGTGCTTCTCAGCTCATGGAACCTGGCCGTCAGAGCAT
GGATCTGGTGGAGCTAGTGTCCCTCTTCCCAACCTTGGCTGGCCTTGCTGGTCTCCAGGTG
CCTCCTAGATGCCCAGTCCCCTCCTTCCATGTTGAACTCTGCCGTGAGGGGAAGAATCTGC
TGAAGCACTTCAGATTCAGAGACTTGGAGGAGGACCCCTACCTTCCTGGGAACCCCAGGGA
GTTGATTGCATACTCCCAGTATCCCAGGCCAAGTGACATTCCCCAGTGGAACTCCGACAAA
CCAAGTCTGAAGGACATCAAGATCATGGGGTACAGCATCAGGACCATTGACTACAGATACA
CAGTGTGGGTTGGATTTAACCCAGATGAGTTCTTGGCAAACTTTTCTGACATCCATGCAAG
TCAGTTGTATTTTGTGGACAGCGACCCTCTGCAGGATCACAACATGTACAATGACAGCCAG
GGTGGGGACCTCTTTCAACTCCTCATGCCATAGCAATTGGCCCCTCTCCCTCCCCCCCCCC
TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTT
TCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGA
CGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGT
GAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGC
AGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAG
ATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAG
AGTCAAATGGCTCACCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCC
CATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGT
TAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGA
TAATCATATGGCCACCATGGCTGCTCCTGCCCTGGGGCTGGTGTGTGGAAGATGTCCTGAA
CTGGGCCTGGTTCTGTTACTGCTTCTGCTCAGCCTGCTCTGTGGTGCTGCCGGCAGCCAAG
AGGCAGGCACTGGCGCTGGAGCTGGAAGCCTGGCTGGGTCTTGTGGATGTGGCACACCACA
GAGGCCAGGGGCTCATGGCTCCTCTGCTGCAGCTCATAGGTACAGCAGAGAAGCCAATGCT
CCAGGCCCAGTGCCTGGAGAGAGACAGCTGGCTCACAGCAAGATGGTGCCCATCCCTGCTG
GGGTGTTCACAATGGGAACAGATGATCCCCAGATCAAGCAGGATGGGGAGGCGCCTGCCAG
GAGGGTGACCATTGATGCATTCTATATGGATGCCTATGAGGTGAGCAATACAGAATTTGAG
AAGTTTGTGAACTCTACTGGCTACCTGACTGAGGCTGAAAAATTTGGAGACTCTTTTGTGT
TTGAAGGAATGCTTAGTGAACAGGTTAAGACCAACATCCAGCAGGCTGTTGCAGCAGCCCC
CTGGTGGTTGCCTGTCAAGGGAGCTAACTGGAGGCACCCTGAGGGACCAGATTCTACAATC
CTGCATAGACCTGATCATCCTGTTCTGCATGTGTCTTGGAATGATGCTGTGGCTTACTGTA
CCTGGGCAGGAAAAAGGCTGCCAACAGAAGCTGAGTGGGAATACTCTTGCAGAGGAGGCCT
GCACAATAGACTGTTCCCATGGGGCAACAAGCTGCAACCCAAGGGCCAGCACTATGCTAAC
ATCTGGCAGGGAGAATTCCCTGTGACAAACACAGGAGAGGACGGCTTCCAGGGAACTGCCC
CTGTAGATGCTTTCCCTCCTAATGGCTATGGCCTGTATAACATTGTTGGCAACGCCTGGGA
GTGGACTTCTGATTGGTGGACAGTGCACCACTCTGTTGAGGAGACACTGAATCCTAAGGGG

CCACCTTCTGGAAAGGATAGAGTGAAGAAGGGGGGAAGCTACATGTGCCACAGGTCTTATT
GTTACAGATACAGGTGCGCTGCTAGGTCTCAGAACACCCCTGATAGCAGTGCTAGCAATCT
GGGCTTCAGGTGTGCCGCTGACAGACTGCCTACCATGGATTAAGTCGACCCTAGAGCTCGC
TGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC
CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC
ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGTGGGGTGGGGCAGGACAGCAAG
GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAATCTAGA Xn(60-
100)GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGC
GCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGG
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCT
GGGGGGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTC
TTTGCAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTC
GGACCTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGC
CATCTGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGA
GATCCCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCA
CGTGTCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTG
CGCAGCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAA
CCGGGAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCC
TGCCTCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCC
GATTTCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTT
TATGGAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAG
CCGGGAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTC
TGGTGGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGC
CTAGCGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTG
GGCGCCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCG
CACCCTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGG
GCTCTCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTC
TGGAAAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTA
AAATACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTA
GAGAAGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTT
TGATGATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGC
AAGCATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGAC
TGTGGGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTG
TTGGGCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGA
AGCTCAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATT
CGCCCTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCC
TCATGCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAA
ATTTATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTAT
TATCGTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGAT
TTATTTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAA
TAGATGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAG
TGCGTTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAG
GCGCCCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCC
GAAGCGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGG
GCCCGAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTC
CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA

TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG
CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTAC
ACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAA
GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
TCTGACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTT
GCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCA
AAGTGCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCAC
CGTAATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAG
ATTGATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAG
ATATAGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAA
CAACCGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCC
GAGGTAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCG
AAAAGATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAA
TGATGCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATC
GTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCC
CGAGGCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGT
TACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTC
CTCGGCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTC
TGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCAC
CTTGTCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATA
TTGGCCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACA
TATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTG
CGAATATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACA
AGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGA
TGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTT
TCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTG
AGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTG
GTATATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCAT
CAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGT
TTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATC
GGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAAT
AAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCAC
TCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATC
GCTGTTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGT
CGGAAGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTG
GCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGC
GATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATC
AGCATCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTT
TTTTCCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG
TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA

TCTGGCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC
CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG
CGCAACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTT
CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT
GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTC
AATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTAT
TTAGAAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATC
GCGAGCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGC
GCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTA
GTGTGGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGG
GGCTCGAG
(SEQ ID NO: 20)

pXL029

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA
GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC
AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG
TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA
GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA
CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA
TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA
AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA
AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGCCACCCCCCCGGACCGGGAGAGGCCTCTTGT
GGTTGGGCCTGGTGCTGAGCAGCGTGTGCGTGGCCCTGGGCAGTGAGACCCAGGCTAACTC
TACAACAGATGCCTTGAATGTGCTGCTGATCATTGTGGATGACCTGAGGCCAAGTCTGGGC
TGCTATGGGGACAAATTGGTGAGGTCCCCCAACATCGACCAGTTGGCCTCCCACTCTCTCC
TATTCCAAAATGCTTTCGCCCAGCAGGCAGTTTGTGCCCCCTCTAGGGTGAGCTTCCTCAC
TGGCAGGCGCCCTGACACCACTAGACTGTATGACTTTAACAGCTATTGGAGGGTGCACGCA
GGAAACTTCTCCACAATCCCTCAATACTTCAAGGAGAATGGTTATGTGACAATGTCTGTGG
GCAAGGTGTTCCACCCTGGCATCAGCAGCAACCACACCGATGACTCACCCTATAGTTGGTC
TTTTCCCCCCTACCATCCTTCATCTGAGAAATATGAAAACACAAAAACCTGCCGAGGCCCA
GACGGGGAACTGCATGCCAACCTACTCTGTCCTGTTGATGTACTGGACGTGCCCGAGGGCA
CCCTCCCTGATAAGCAGTCCACAGAACAGGCCATTCAGCTGCTTGAAAAGATGAAGACCTC
CGCATCCCCCTTCTTCTTGGCTGTCGGCTACCACAAGCCCCATATCCCCTTTAGATACCCC
AAGGAATTCCAGAAACTGTACCCACTGGAGAACATCACACTTGCTCCTGACCCTGAAGTGC
CTGACGGACTGCCTCCAGTGGCCTATAACCCTTGGATGGACATCCGGCAGCGCGAGGATGT
GCAGGCTCTGAACATTAGTGTGCCTTATGGGCCCATCCCTGTGGACTTTCAGAGGAAGATT
CGCCAGTCCTACTTTGCCTCTGTATCCTACCTGGACACACAGGTGGGACGCCTGCTGTCTG
CCCTTGATGATCTGCAACTGGCCAACAGCACCATTATAGCTTTCACATCAGACCATGGGTG

GGCTCTTGGGGAGCATGGTGAATGGGCTAAGTACTCCAACTTCGATGTGGCAACCCATGTC
CCTCTGATCTTCTATGTGCCAGGAAGGACCGCCTCTCTGCCAGAGGCAGGTGAGAAGCTGT
TCCCCTATCTGGACCCTTTTGACTCCGCCAGCCAGCTGATGGAGCCTGGCCGACAGTCTAT
GGACCTGGTTGAGCTGGTCAGCCTGTTTCCCACACTCGCTGGACTGGCTGGCCTGCAAGTA
CCCCCACGCTGCCCAGTGCCCTCCTTCCATGTGGAGCTTTGCAGGGAGGGGAAGAACCTCC
TCAAGCACTTCAGGTTCAGGGACCTAGAGGAGGATCCTTATCTGCCTGGAAACCCCAGAGA
GCTTATTGCTTACTCCCAGTATCCAAGGCCTAGTGACATTCCCCAATGGAACTCAGACAAA
CCAAGCCTGAAAGACATCAAGATCATGGGATACTCTATCAGGACCATTGACTACAGGTACA
CTGTGTGGGTTGGCTTCAACCCGGATGAGTTCCTGGCTAATTTCTCTGACATACATGCTGG
CGAGCTGTACTTCGTGGACAGTGACCCCCTGCAGGATCACAACATGTACAATGATTCCCAG
GGGGGTGACCTCTTCCAGCTTCTGATGCCCTAACAATTGGCCCCTCTCCCTCCCCCCCCCC
TAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTT
TCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGA
CGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGT
GAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGC
AGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAG
ATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAG
AGTCAAATGGCTCACCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCC
CATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGT
TAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGA
TAATACCGGTGCCACCATGGCTGCCCCTGCTCTGGGATTGGTTTGTGGCAGATGTCCTGAG
CTTGGTCTGGTGCTGTTGCTCCTTCTGTTGTCTCTGCTGTGTGGAGCAGCTGGGTCTCAGG
AAGCTGGCACAGGCGCTGGGGCTGGCTCTCTGGCCGGGTCATGTGGCTGTGGAACTCCCCA
GCGGCCTGGAGCCCATGGCAGCTCTGCCGCAGCACACAGGTATTCTAGGGAAGCCAATGCC
CCAGGCCCTGTGCCTGGGGAGAGACAGCTAGCTCATTCTAAGATGGTGCCTATCCCAGCCG
GGGTTTTTACAATGGGCACTGATGATCCTCAGATTAAGCAGGATGGAGAGGCCCCCGCCAG
AAGAGTGACCATTGATGCTTTCTACATGGATGCATATGAAGTGTCCAACACAGAGTTTGAG
AAATTTGTGAACTCTACTGGATACTTGACCGAGGCTGAGAAGTTTGGAGATTCCTTTGTCT
TTGAAGGCATGCTGTCTGAGCAGGTCAAGACCAACATTCAGCAAGCAGTGGCCGCTGCACC
TTGGTGGCTTCCTGTGAAGGGCGCCAACTGGAGACATCCAGAGGGGCCAGATAGTACCATC
CTCCACAGACCTGATCACCCAGTCCTTCATGTTTCCTGGAATGATGCAGTTGCTTACTGCA
CTTGGGCCGGCAAGAGGCTCCCTACTGAGGCAGAGTGGGAATACTCCTGCAGAGGAGGCCT
GCACAACAGACTGTTCCCTTGGGGGAACAAGCTTCAGCCCAAAGGCCAGCACTATGCTAAC
ATCTGGCAGGGTGAGTTTCCAGTCACCAATACAGGGGAGGACGGATTCCAGGGAACCGCAC
CAGTAGATGCCTTCCCTCCTAATGGCTATGGCCTGTATAATATTGTGGGCAATGCATGGGA
GTGGACCTCTGACTGGTGGACTGTGCACCACTCAGTGGAGGAAACCCTGAACCCTAAGGGA
CCCCCTTCAGGCAAAGATAGAGTCAAAAAGGGAGGGAGCTATATGTGTCACAGATCCTATT
GCTACAGATATAGATGTGCAGCCAGGTCCCAGAACACCCCTGACTCTTCTGCTAGCAACCT
GGGCTTTCGGTGTGCTGCTGATAGACTGCCCACCATGGACTAAGTCGACCCTAGAGCTCGC
TGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC
CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC
ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG
GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAATCTAGA Xn(60-
100)GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGC
GCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGG
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCT
GGGGGGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTC
TTTGCAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTC
GGACCTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGC

CATCTGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGA
GATCCCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCA
CGTGTCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTG
CGCAGCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAA
CCGGGAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCC
TGCCTCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCC
GATTTCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTT
TATGGAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAG
CCGGGAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTC
TGGTGGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGC
CTAGCGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTG
GGCGCCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCG
CACCCTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGG
GCTCTCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTC
TGGAAAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTA
AAATACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTA
GAGAAGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTT
TGATGATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGC
AAGCATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGAC
TGTGGGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTG
TTGGGCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGA
AGCTCAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATT
CGCCCTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCC
TCATGCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAA
ATTTATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTAT
TATCGTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGAT
TTATTTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAA
TAGATGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAG
TGCGTTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAG
GCGCCCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCC
GAAGCGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGG
GCCCGAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTC
CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG
CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTAC
ACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAA
GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
TCTGACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTT
GCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCA
AAGTGCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCAC

CGTAATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAG
ATTGATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAG
ATATAGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAA
CAACCGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCC
GAGGTAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCG
AAAAGATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAA
TGATGCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATC
GTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCC
CGAGGCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGT
TACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTC
CTCGGCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTC
TGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCAC
CTTGTCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATA
TTGGCCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACA
TATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTG
CGAATATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACA
AGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGA
TGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTT
TCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTG
AGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTG
GTATATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCAT
CAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGT
TTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATC
GGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAAT
AAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCAC
TCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATC
GCTGTTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGT
CGGAAGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTG
GCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGC
GATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATC
AGCATCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTT
TTTTCCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG
TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA
TCTGGCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC
CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG
CGCAACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTT
CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT
GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTC
AATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTAT

TTAGAAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATC
GCGAGCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGC
GCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTA
GTGTGGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGG
GGCTCGAG
(SEQ ID NO: 21)

pXL030

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAACGCGTGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCCA
GTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGGTGAATATTAACC
AAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACCGGGGGAGGCTGCTGG
TGAATATTAACCAAGGTCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACTA
GTAAATGACCTATTAAGAATATTTCATAGAACGAATGTTCCGATGCTCTAATCTCTCTAGA
CAAGGTTCATATTTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAA
TCAGCAGGTTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAAGTGAGTATA
AAAGCCCCAGGCTGGGAGCAGCCATCACAGAAGTCCACTCATTCTTGGCAGGCCGCGGCTA
AGGTAAGTTGGCGCCGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTT
TTTTACAGGCCTGGGCGCGCCGCCACCATGGCTGCGCCCGCACTAGGGCTGGTGTGTGGAC
GTTGCCCTGAGCTGGGTCTCGTCCTCTTGCTGCTGCTGCTCTCGCTGCTGTGTGGAGCGGC
AGGGAGCCAGGAGGCCGGGACCGGTGCGGGCGCGGGGTCCCTTGCGGGTTCTTGCGGCTGC
GGCACGCCCCAGCGGCCTGGCGCCCATGGCAGTTCGGCAGCCGCTCACCGATACTCGCGGG
AGGCTAACGCTCCGGGCCCCGTACCCGGAGAGCGGCAACTCGCGCACTCAAAGATGGTCCC
CATCCCTGCTGGAGTATTTACAATGGGCACAGATGATCCTCAGATAAAGCAGGATGGGGAA
GCACCTGCGAGGAGAGTTACTATTGATGCCTTTTACATGGATGCCTATGAAGTCAGTAATA
CTGAATTTGAGAAGTTTGTGAACTCAACTGGCTATTTGACAGAGGCTGAGAAGTTTGGCGA
CTCCTTTGTCTTTGAAGGCATGTTGAGTGAGCAAGTGAAGACCAATATTCAACAGGCAGTT
GCAGCTGCTCCCTGGTGGTTACCTGTGAAAGGCGCTAACTGGAGACACCCAGAAGGGCCTG
ACTCTACTATTCTGCACAGGCCGGATCATCCAGTTCTCCATGTGTCCTGGAATGATGCGGT
TGCCTACTGCACTTGGGCAGGGAAGCGGCTGCCCACGGAAGCTGAGTGGGAATACAGCTGT
CGAGGAGGCCTGCATAATAGACTTTTCCCCTGGGGCAACAAACTGCAGCCCAAAGGCCAGC
ATTATGCCAACATTTGGCAGGGCGAGTTTCCGGTGACCAACACTGGTGAGGATGGCTTCCA
AGGAACTGCGCCTGTTGATGCCTTCCCTCCCAATGGTTATGGCTTATACAACATAGTGGGG
AACGCATGGGAATGGACTTCAGACTGGTGGACTGTTCATCATTCTGTTGAAGAAACGCTTA
ACCCAAAAGGTCCCCCTTCTGGGAAAGACCGAGTGAAGAAAGGTGGATCCTACATGTGCCA
TAGGTCTTATTGTTACAGGTATCGCTGTGCTGCTCGGAGCCAGAACACACCTGATAGCTCT
GCTTCGAATCTGGGATTCCGCTGTGCAGCCGACCGCCTGCCCACTATGGACTGAGTCGACC
CTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC
CTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAAT
GAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAATCTAGA
Xn(60-
100)GTTTAAACATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGC
GCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGG
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCT
GGGGGGCCCTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTC
TTTGCAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTC
GGACCTGGCGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGC

CATCTGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGA
GATCCCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCA
CGTGTCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTG
CGCAGCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAA
CCGGGAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCC
TGCCTCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCC
GATTTCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTT
TATGGAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAG
CCGGGAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTC
TGGTGGTGGCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGC
CTAGCGTTGACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTG
GGCGCCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCG
CACCCTGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGG
GCTCTCTCCCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTC
TGGAAAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTA
AAATACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTA
GAGAAGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTT
TGATGATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGC
AAGCATTCCTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGAC
TGTGGGTGGCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTG
TTGGGCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGA
AGCTCAAGGGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATT
CGCCCTCTTAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCC
TCATGCCTTCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAA
ATTTATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTAT
TATCGTTGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGAT
TTATTTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAA
TAGATGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAG
TGCGTTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAG
GCGCCCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCC
GAAGCGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGG
GCCCGAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTC
CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG
CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTAC
ACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAA
GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
TCTGACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTT
GCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCA
AAGTGCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCAC

CGTAATCTGCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAG
ATTGATGAGCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAG
ATATAGATCTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAA
CAACCGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCC
GAGGTAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCG
AAAAGATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAA
TGATGCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATC
GTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCC
CGAGGCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGT
TACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTC
CTCGGCGTTTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTC
TGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCAC
CTTGTCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATA
TTGGCCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACA
TATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTG
CGAATATATGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACA
AGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGA
TGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTT
TCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTG
AGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTG
GTATATCCAGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCAT
CAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGT
TTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATC
GGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAAT
AAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGT
TTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCAC
TCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATC
GCTGTTAAAAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGT
CGGAAGTGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTG
GCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGC
GATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATC
AGCATCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTT
TTTTCCTCCTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG
TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA
TCTGGCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC
CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG
CGCAACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTT
CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT
GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTC
AATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTAT

TTAGAAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATC
GCGAGCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGC
GCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTA
GTGTGGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGG
GGCTCGAG
(SEQ ID NO: 22)

pXL032

CGCGTTGTAGTTAATGATTAACCCGCCATGCTACTTATCTACGTAGCCATGCTCTAGGAAG
ATCGGAATTCGCCCTTAAGCTAGCAGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGG
CCCTTGGCAGCATTTACTCTCTCTGTTTGCTCTGGTTAATAATCTCAGGAGCACAAACATT
CCAGATCCAGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGGCCCTTGGCAGCATTTA
CTCTCTCTGTTTGCTCTGGTTAATAATCTCAGGAGCACAAACATTCCAGATCCGGCGCGCC
AGGGCTGGAAGCTACCTTTGACATCATTTCCTCTGCGAATGCATGTATAATTTCTACAGAA
CCTATTAGAAAGGATCACCCAGCCTCTGCTTTTGTACAACTTTCCCTTAAAAAACTGCCAA
TTCCACTGCTGTTTGGCCCAATAGTGAGAACTTTTTCCTGCTGCCTCTTGGTGCTTTTGCC
TATGGCCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGGACTTAAACCCCTCCAGC
TCTGACAATCCTCTTTCTCTTTTGTTTTACATGAAGGGTCTGGCAGCCAAAGCAATCACTC
AAAGTTCAAACCTTATCATTTTTTGCTTTGTTCCTCTTGGCCTTGGTTTTGTACATCAGCT
TTGAAAATACCATCCCAGGGTTAATGCTGGGGTTAATTTATAACTAAGAGTGCTCTAGTTT
TGCAATACAGGACATGCTATAAAAATGGAAAGATGTTGCTTTCTGAGAGACTGCAGAAGTT
GGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATA
GAAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTT
ACTGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCAGGCGGCCGCATGCCCCCGCCCCG
CACCGGCCGCGGCCTGCTGTGGCTGGGCCTGGTGCTGAGCAGCGTGTGCGTGGCCCTGGGC
AGCGAGACCCAGGCCAACAGCACCACCGACGCCCTGAACGTGCTGCTGATCATCGTGGACG
ACCTGCGCCCCAGCCTGGGCTGCTACGGCGACAAGCTGGTGCGCAGCCCCAACATCGACCA
GCTGGCCAGCCACAGCCTGCTGTTCCAGAACGCCTTCGCCCAGCAGGCCGTGTGCGCCCCC
AGCCGCGTGAGCTTCCTGACCGGCCGCCGCCCCGACACCACCCGCCTGTACGACTTCAACA
GCTACTGGCGCGTGCACGCCGGCAACTTCAGCACCATCCCCCAGTACTTCAAGGAGAACGG
CTACGTGACCATGAGCGTGGGCAAGGTGTTCCACCCCGGCATCAGCAGCAACCACACCGAC
GACAGCCCCTACAGCTGGAGCTTCCCCCCCTACCACCCCAGCAGCGAGAAGTACGAGAACA
CCAAGACCTGCCGCGGCCCCGACGGCGAGCTGCACGCCAACCTGCTGTGCCCCGTGGACGT
GCTGGACGTGCCCGAGGGCACCCTGCCCGACAAGCAGAGCACCGAGCAGGCCATCCAGCTG
CTGGAGAAGATGAAGACCAGCGCCAGCCCCTTCTTCCTGGCCGTGGGCTACCACAAGCCCC
ACATCCCCTTCCGCTACCCCAAGGAGTTCCAGAAGCTGTACCCCCTGGAGAACATCACCCT
GGCCCCCGACCCCGAGGTGCCCGACGGCCTGCCCCCCGTGGCCTACAACCCCTGGATGGAC
ATCCGCCAGCGCGAGGACGTGCAGGCCCTGAACATCAGCGTGCCCTACGGCCCCATCCCCG
TGGACTTCCAGCGCAAGATCCGCCAGAGCTACTTCGCCAGCGTGAGCTACCTGGACACCCA
GGTGGGCCGCCTGCTGAGCGCCCTGGACGACCTGCAGCTGGCCAACAGCACCATCATCGCC
TTCACCAGCGACCACGGCTGGGCCCTGGGCGAGCACGGCGAGTGGGCCAAGTACAGCAACT
TCGACGTGGCCACCCACGTGCCCCTGATCTTCTACGTGCCCGGCCGCACCGCCAGCCTGCC
CGAGGCCGGCGAGAAGCTGTTCCCCTACCTGGACCCCTTCGACAGCGCCAGCCAGCTGATG
GAGCCCGGCCGCCAGAGCATGGACCTGGTGGAGCTGGTGAGCCTGTTCCCCACCCTGGCCG
GCCTGGCCGGCCTGCAGGTGCCCCCCCGCTGCCCCGTGCCCAGCTTCCACGTGGAGCTGTG
CCGCGAGGGCAAGAACCTGCTGAAGCACTTCCGCTTCCGCGACCTGGAGGAGGACCCCTAC
CTGCCCGGCAACCCCCGCGAGCTGATCGCCTACAGCCAGTACCCCCGCCCCAGCGACATCC
CCCAGTGGAACAGCGACAAGCCCAGCCTGAAGGACATCAAGATCATGGGCTACAGCATCCG
CACCATCGACTACCGCTACACCGTGTGGGTGGGCTTCAACCCCGACGAGTTCCTGGCCAAC

TTCAGCGACATCCACGCCGGCGAGCTGTACTTCGTGGACAGCGACCCCCTGCAGGACCACA
ACATGTACAACGACAGCCAGGGCGGCGACCTGTTCCAGCTGCTGATGCCCTAGAAGCCTGG
ATCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTT
GCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCC
GTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTT
GTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACT
GGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTA
TTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTT
GGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCC
TGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATC
CAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGAGA
TCTGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTT
CCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGG
GAGGATTGGGAAGACAATAGCAGGCATGCTGGGGACTCGAGTTAAGGGCGAATTCCCGATT
AGGATCTTCCTAGAGCATGGCTACGTAGATAAGTAGCATGGCGGGTTAATCATTAACTACA
GTTTAAAC Xn(60-
100)ATTTAAATAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTC
GCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCA
GTGAGCGAGCGAGCGCGCAGAGAGTATACATCGATGTGAGTTCGCGGGTGGCTGGGGGGCC
CTGGGCTGCGACCGCCCCCGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTCTTTGCAGT
CGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTCGGACCTGG
CGGAGCCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCGCCATCTGCC
CGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTCCCTACTCCCGCCCGAGATCCCGA
CCCGGACCCCTAGGTGGGGGACGCTTTCTTTCCTTTCGCGCTCTGCGGGGTCACGTGTCGC
AGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCTAGTGCGCAGCGG
GTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGTTGCAACCGGGAAG
GAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTGCGCTCCTGCCTCGA
TGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGCGAGGCTAGCTGGCCCGATTTCTC
CTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGAAGTGGGTTTATGGAGG
TCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCATGGTGCCAAGCCGGGAGA
AGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAAATGGCCCCTCTGGTGGTG
GCCCCTTCCTGCAGCGCCGGCTCACCTCACGGCCCCGCCCTTCCCCTGCCAGCCTAGCGTT
GACCCGACCCCAAAGGCCAGGCTGTAAATGTCACCGGGAGGATTGGGTGTCTGGGCGCCTC
GGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGATCTCCCACCGCACCCTGG
TCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGCTGGGGGCCTGGGCTGGGGCTCTCTC
CCATCCCTTCTCCCCACACACATGCACTTACCTGTGCTCCCACTCCTGATTTCTGGAAAAG
AGCTAGGAAGGACAGGCAACTTGGCAAATCAAAGCCCTGGGACTAGGGGGTTAAAATACAG
CTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTCCCTTTTGTAGGAGGGACTTAGAGAAGGG
GTGGGCTTGCCCTGTCCAGTTAATTTCTGACCTTTACTCCTGCCCTTTGAGTTTGATGATG
CTGAGTGTACAAGCGTTTTCTCCCTAAAGGGTGCAGCTGAGCTAGGCAGCAGCAAGCATTC
CTGGGGTGGCATAGTGGGGTGGTGAATACCATGTACAAAGCTTGTGCCCAGACTGTGGGTG
GCAGTGCCCCACATGGCCGCTTCTCCTGGAAGGGCTTCGTATGACTGGGGGTGTTGGGCAG
CCCTGGAGCCTTCAGTTGCAGCCATGCCTTAAGCCAGGCCAGCCTGGCAGGGAAGCTCAAG
GGAGATAAAATTCAACCTCTTGGGCCCTCCTGGGGGTAAGGAGATGCTGCATTCGCCCTCT
TAATGGGGAGGTGGCCTAGGGCTGCTCACATATTCTGGAGGAGCCTCCCCTCCTCATGCCT
TCTTGCCTCTTGTCTCTTAGGCATGCAAAAGAGTCGAATAAGGGCGACACAAAATTTATTC
TAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATCGTTG
ACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTATTTTC

TTAATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGATGAA
TAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCGTTGC
TTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGCCCTT
AAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAGCGGG
TTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCCGAGC
TTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCCATCC
GTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCTTCCT
CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAA
GGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA
GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC
GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACC
CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACTAGAAG
AACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGC
TCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA
TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGC
TCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGCCGTC
CCGTCAAGTCAGCGTAATGCTCTGCTTAGGTGGCGGTACTTGGGTCGATATCAAAGTGCAT
CACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTAATCT
GCTTGCACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTGATGA
GCGCGGTGGCAATGCCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAGATATAGAT
CTCACTACGCGGCTGCTCAAACTTGGGCAGAACGTAAGCCGCGAGAGCGCCAACAACCGCT
TCTTGGTCGAAGGCAGCAAGCGCGATGAATGTCTTACTACGGAGCAAGTTCCCGAGGTAAT
CGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCACCGAACTCACGACCGAAAAGATC
AAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACATGTGCGAATGATGCCC
ATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGC
TCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAGGCAT
AGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCACCG
CTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATTTTTTTTTCCTCCTCGGCGT
TTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCCGACA
TGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGC
CTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGGCCAC
GTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCGCTGACGAAAAACATATTCTCA
ATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAATATA
TGTGTAGAAACTGCCGGAAATCGTCGTGTGCACTCATGGAAAACGGTGTAACAAGGGTGAA
CACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGATGAGCATT
CATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTTTACG
GTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTG
ACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCC
AGTGATTTTTTTCTCCATTTTTTTTTCCTCCTTTAGAAAAACTCATCGAGCATCAAATGAA
ACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAA
TGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCG
ATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTAT
CAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCAT
TTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCA

ACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGGCGAAATACGCGATCGCTGTTAA
AAGGACAATTACAAACAGGAATCGAGTGCAACCGGCGCAGGAACACTGCCAGCGCATCAAC
AATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAACGCTGTTTTTCCGGGGATC
GCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGTG
GCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCT
ACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATAGATT
GTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCA
TGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATTTTTTTTCCTC
CTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCA
TAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCC
CAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGT
CTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGT
TGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGC
TCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCACGTTGTCAGAAG
TAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT
AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACA
TAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGG
ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAG
CATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA
AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATATTAT
TGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAA
ATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGAGCCC
ATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGGTGGT
CCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGG
ACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAA
GACTGGGCCTTTCGCCCGGGCTAATTGAGGGGTGTCGCCCTTATTCGACTCGGGGCTCGAG
CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTTTAATTAAA
(SEQ ID NO: 23)

Use of rAAV Vectors that Encode I2S for Treatment of Disease

[0123] Described herein are methods of treating a disease associated with I2S enzyme deficiency. Accordingly, in some embodiments, the rAAV vectors described herein are suitable for treating a subject that has an I2S deficiency, such as Hunter syndrome (MPSII). The method of treating includes administering to the subject in need thereof a recombinant adeno-associated virus (rAAV) vector as described herein.

[0124] The rAAV vector described herein can be used to treat any disease associated with I2S deficiency or disorder.

[0125] In some embodiments, the rAAV vector remains episomal following administration to a subject in need thereof. In some embodiments, the rAAV vector does not remain episomal following administration to a subject in need thereof. For example, in some embodiments, the rAAV vector integrates into the genome of the subject. Such integration can be achieved, for example, by using various gene-editing technologies, such as, zinc finger nucleases (ZFNs), Transcription activator-like effector nucleases (TALENS), ARCUS genome editing, and/or CRISPR-Cas systems.

[0126] In some embodiments, a pharmaceutical composition comprising an rAAV vector described herein is used to treat subjects in need thereof. The pharmaceutical composition containing an rAAV vector or particle of the invention contains a pharmaceutically acceptable excipient, diluent or carrier. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions and the like. Such carriers can be formulated by conventional methods and are administered to the subject at a therapeutically effective amount.

[0127] The rAAV vector is administered to a subject in need thereof via a suitable route. In embodiments, the rAAV vector is administered by intravenous, intraperitoneal, subcutaneous, or intradermal administration. In embodiments, the rAAV vector is administered intravenously. In embodiments, the intradermal administration comprises administration by use of a "gene gun" or biolistic particle delivery system. In some embodiments, the rAAV vector is administered via a non-viral lipid nanoparticle. For example, a composition comprising the rAAV vector may comprise one or more diluents, buffers, liposomes, a lipid, a lipid complex. In some embodiments, the rAAV vector is comprised within a microsphere or a nanoparticle, such as a lipid nanoparticle. In some embodiments, the rAAV vectors and/or the transgene expression cassette and/or the optimized *IDS* transgene sequences and/or any compositions of the gene expression cassette are administered via non-viral chemical particles such as lipid nanoparticles, non-viral biological molecules such as exosomes and/or extracellular vesicle.

[0128] In some embodiments, functional I2S is detectable in plasma or serum of the subject at about 2 to 6 weeks post administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at about 2 weeks. In some embodiments, functional I2S is detectable in plasma or serum of the subject at about 3 weeks.

In some embodiments, functional I2S is detectable in plasma or serum of the subject at about 4 weeks. In some embodiments, functional I2S is detectable in plasma or serum of the subject at about 5 weeks. In some embodiments, functional I2S is detectable in plasma or serum of the subject at about 6 weeks. In some embodiments, functional I2S is detectable in hepatocytes of the subject at about 2 to 6 weeks post administration of the rAAV vector.

[0129] In some embodiments, functional I2S is detectable in plasma of the subject at least 3 months, 6 months, 12 months, 2 years, 3 years, 4 years, 5 years, 6 years, 7 years, 8 years, 9 years, or 10 years after administration of the rAAV vector. Accordingly, in some embodiments, functional I2S is detectable in plasma or serum of the subject at least 3 months after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 6 months after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 12 months after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 2 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 3 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 4 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 5 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 6 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 7 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 8 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 9 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject at least 10 years after administration of the rAAV vector. In some embodiments, functional I2S is detectable in plasma or serum of the subject for the remainder of the subject's life following administration of the rAAV vector.

[0130] In some embodiments, the administered rAAV comprising I2S results in the production of active I2S to the same extent as found following administration of purified I2S protein delivered intravenously. In some embodiments, the administered rAAV comprising

I2S results in production of a greater amount of active I2S as compared to administration of purified I2S protein delivered intravenously.

[0131] In some embodiments, the administered rAAV comprising I2S results in the reduction of glycosaminoglycan (GAG) in the subject. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, or about 10% in comparison to the subject's baseline GAG level prior to administering the rAAV comprising I2S. Accordingly, in some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 95%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 90%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 85%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 80%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 75%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 70%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 65%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 60%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 55%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 50%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 45%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 40%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 35%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 30%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 25%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 20%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 15%. In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject by about 10%.

[0132] In some embodiments, the administered rAAV comprising I2S reduces GAG in the subject for at least about 2 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 12 months, 1 year, 2 years, 3 years, 4 years, 5 years, or more than 5 years.

[0133] In some embodiments, following administration of the AAV vector to the subject the levels of functional I2S detectable in the circulation are between about 2 and 10 times greater than the amount of functional I2S detectable in the subject before administration of the rAAV comprising I2S.

[0134] In some embodiments, following administration of the AAV vector to the subject the levels of detectable active I2S meets or exceeds human therapeutic level. In some embodiments, the levels of active I2S post administration of the rAAV vector is about between 2 and 35 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 2 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 3 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 4 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 5 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 6 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 6 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 7 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 8 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 9 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 10 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 15 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 20 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 25 times the human therapeutic level. In some embodiments, the levels of active I2S post administration is about 30 times the human therapeutic level. In some embodiments, the levels of ac active I2S post administration is about 35 times the human therapeutic level.

[0135] Thus, administration of rAAV vector comprising the I2S results in sustained robust expression in comparison to a single administration of purified I2S to a subject in need.

[0136] In some embodiments, the rAAV I2S vector is delivered as a single dose per subject. In some embodiments, the subject is delivered the minimal effective dose (MED). As

used herein, MED refers to the rAAV I2S vector dose required to achieve I2S activity resulting in reduced GAG levels in a subject.

[0137] The vector titer is determined on the basis of the DNA content of the vector preparation. In some embodiments, quantitative PCR or optimized quantitative PCR is used to determine the DNA content of the rAAV I2S vector preparations. In one embodiment, the dosage is about 1 x 10^{11} vector genome (vg)/kg body weight to about 1 x 10^{13} vg/kg, inclusive of endpoints.

[0138] In one embodiment, the dosage is selected in the range of 1 x 10^{9} vg/kg to 3 x 10^{15} vg/kg (for example, 1 x 10^{9} vg/kg, 3 x 10^{9} vg/kg, 1 x 10^{10} vg/kg, 3 x 10^{10} vg/kg, 1 x 10^{11} vg/kg, 3 x 10^{11} vg/kg, 1 x 10^{12} vg/kg, 3 x 10^{12} vg/kg, 1 x 10^{13} vg/kg, 3 x 10^{13} vg/kg, 1 x 10^{14} vg/kg, 3 x 10^{14} vg/kg, 1 x 10^{15} vg/kg, 3 x 10^{15} vg/kg). In some embodiments, the dosage is 5 x 10^{13} vg/kg . In another embodiment, the dosage is 5 x $l0^{12}$ vg/kg In specific embodiments, the dose of rAAV administered to a subject is at least 5 x 10^{11} vg/kg , 1 x 10^{12} vg/kg , 1.5 x 10^{12} vg/kg , 2.0 x 10^{12} vg/kg , 2.5 x 10^{12} vg/kg , 3.0 x 10^{12} vg/kg , 3.5 x 10^{12} vg/kg , 4.0 x 10^{12} vg/kg , 4.5 x 10^{12} vg/kg , 5.0 x 10^{12} vg/kg , 5.5 x 10^{12} vg/kg , 6.0 x 10^{12} vg/kg , 6.5 x 10^{12} vg/kg , 7.0 x 10^{12} vg/kg , or 7.5 x 10^{12} vg/kg.

[0139] In some embodiments, the rAAV I2S vector compositions can be formulated in dosage units to contain an amount of replication-defective virus that is in the range of about 1.0 x 10^{9} vg to about 1.0 x 10^{15} vg. As used herein, the term "dosage" can refer to the total dosage delivered to the subject in the course of treatment, or the amount delivered in a single (of multiple) administration.

[0140] In some embodiments, the dosage is sufficient to decrease plasma GAG levels in the patient by 25% or more. In some embodiments, rAAV I2S is administered in combination with one or more therapies for the treatment of Hunter syndrome.

Production of rAAV Viral Vectors

[0141] Methods for generating and isolating AAV viral vectors suitable for delivery to a subject are known in the art. *See*, e.g., US Patent 7790449; US Patent 7282199; WO 2003/042397; WO 2005/033321, WO 2006/1 10689; and US 7588772 B2. In a one system, a producer cell line is transiently transfected with a construct that encodes the transgene flanked by ITRs and a construct(s) that encodes rep and cap. In a second system, a packaging cell line that stably supplies rep and cap is transiently transfected with a construct encoding

the transgene flanked by ITRs. In each of these systems, AAV virions are produced in response to infection with helper adenovirus or herpesvirus, requiring the separation of the rAAVs from contaminating virus. More recently, systems have been developed that do not require infection with helper virus to recover the AAV (i.e., adenovirus El, E2a, VA, and E4 or herpesvirus UL5, UL8, UL52, and UL29, and herpesvirus polymerase) are also supplied, in trans, by the system. In these newer systems, the helper functions can be supplied by transient transfection of the cells with constructs that encode the required helper functions, or the cells can be engineered to stably contain genes encoding the helper functions, the expression of which can be controlled at the transcriptional or posttranscriptional level.

[0142] In some embodiments, the expression cassette flanked by ITRs and rep/cap genes are introduced into a desired cell or cell line by infection with baculovirus-based vectors.

[0143] In some embodiments, the expression cassette flanked by ITRs and rep/cap genes are introduced into insect cells by infection with baculovirus-based vectors. For reviews on these production systems, *see* generally, e.g., Zhang et al, 2009, "Adenovirus-adeno-associated virus hybrid for large-scale recombinant adeno-associated virus production," *Human Gene Therapy* 20:922-929, the contents of which is incorporated herein by reference in its entirety. Methods of making and using these and other AAV production systems are also described in the following U.S. patents, the contents of each of which is incorporated herein by reference in its entirety: 5, 139,941 ; 5,741,683; 6,057, 152; 6,204,059; 6,268,213; 6,491,907; 6,660,514; 6,951,753; 7,094,604; 7, 172,893; 7,201,898; 7,229,823; and 7,439,065. *See* generally, e.g., Grieger & Samulski, 2005, "Adeno-associated virus as a gene therapy vector: Vector development, production and clinical applications," *Adv. Biochem. Engin/Biotechnol.* 99: 119-145; Buning et al, 2008, "Recent developments in adeno-associated virus vector technology," *J. Gene Med* 10:717-733; and the references cited below, each of which is incorporated herein by reference in its entirety.

[0144] The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. *See*, e.g., Green and Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY (2012). Similarly, methods of generating rAAV virions are well known and the selection of a suitable

method is not a limitation on the present invention. See, e.g., K. Fisher et al, (1993) J. Virol, 70:520-532 and US Patent No. 5,478,745.

[0145] Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties, construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

[0146] In one embodiment, the production plasmid is that described herein, or as described in WO2012/158757, which is incorporated herein by reference. Various plasmids are known in the art for use in producing rAAV vectors, and are useful herein. The production plasmids are cultured in the host cells which express the AAV cap and/or rep proteins. In the host cells, each rAAV genome is rescued and packaged into the capsid protein or envelope protein to form an infectious viral particle.

[0147] In certain embodiments, the rAAV expression cassette, the vector (such as rAAV vector), the virus (such as rAAV), the production plasmid comprises AAV inverted terminal repeat sequences, a codon optimized nucleic acid sequence that encodes *IDS* and/or *SUMF-1*, and expression control sequences that direct expression of the encoded proteins in a host cell. In other embodiments, the rAAV expression cassette, the virus, the vector (such as rAAV vector), the production plasmid further comprise one or more of an intron, a Kozak sequence, a polyA, posttranscriptional regulatory elements and others. In one embodiment, the post-transcriptional regulatory element is Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE).

[0148] Various methods are known in the art relating to the production and purification of AAV vectors. *See*, e.g., Mizukami, Hiroaki, et al. A Protocol for AAV vector production and purification; U.S. Patent Publication Numbers US20070015238 and US20120322861. For example, a plasmid comprising a gene of interest may be combined with one or more helper plasmids, e.g., that contain a rep gene (e.g., encoding Rep78, Rep68, Rep52 and Rep40) and a cap gene (encoding VP1, VP2, and VP3, including a modified VP2 region as described herein), and transfected into a recombinant cells such that the rAAV can be packaged and subsequently purified.

[0149] In some embodiments, the packaging is performed in a helper cell or producer cell, such as a mammalian cell or an insect cell. Exemplary mammalian cells include, but are not limited to, HEK293 cells, COS cells, HeLa cells, BHK cells, or CHO cells (see, e.g., ATCC® CRL-1573™, ATCC® CRL-1651™, ATCC® CRL-1650™, ATCC® CCL-2, ATCC® CCL-10™, or ATCC® CCL-61™). Exemplary insect cells include, but are not limited to Sf9 cells (see, e.g., ATCC® CRL-1711™). The helper cell may comprises rep and/or cap genes that encode the Rep protein and/or Cap proteins for use in a method described herein. In some embodiments, the packaging is performed in vitro.

[0150] In some embodiments, a plasmid containing comprising the gene of interest is combined with one or more helper plasmids, e.g., that contain a rep gene of a first serotype and a cap gene of the same serotype or a different serotype, and transfected into helper cells such that the rAAV is packaged.

[0151] In some embodiments, the one or more helper plasmids include a first helper plasmid comprising a rep gene and a cap gene, and a second helper plasmid comprising one or more of the following helper genes: Ela gene, Elb gene, E4 gene, E2a gene, and VA gene. For clarity, helper genes are genes that encode helper proteins Ela, Elb, E4, E2a, and VA. In some embodiments, the cap gene is modified such that one or more of the proteins VP1, VP2 and VP3 do not get expressed. In some embodiments, the cap gene is modified such that VP2 does not get expressed. Methods for making such modifications are known in the art (Lux et al. (2005), *J Virology*, 79: 11776-87).

[0152] Helper plasmids, and methods of making such plasmids, are generally known in the art and generally commercially available (*see*, e.g., pDF6, pRep, pDM, pDG, pDPlrs, pDP2rs, pDP3rs, pDP4rs, pDP5rs, pDP6rs, pDG(R484E/R585E), and pDP8.ape plasmids from PlasmidFactory, Bielefeld, Germany; other products and services available from Vector Biolabs, Philadelphia, PA; Cellbiolabs, San Diego, CA; Agilent Technologies, Santa Clara, Ca; and Addgene, Cambridge, MA; pxx6; Grimm et al. (1998), Novel Tools for Production and Purification of Recombinant Adeno associated Virus Vectors, *Human Gene Therapy*, Vol. 9, 2745-2760; Kem, A. et al. (2003), Identification of a Heparin-Binding Motif on Adeno- Associated Virus Type 2 Capsids, *Journal of Virology*, Vol. 77, 11072-11081.; Grimm et al. (2003), Helper Virus-Free, Optically Controllable, and Two-Plasmid-Based Production of Adeno-associated Virus Vectors of Serotypes 1 to 6, *Molecular Therapy*, Vol. 7, 839-850; Kronenberg et al. (2005).

EXAMPLES

[0153] Other features, objects, and advantages of the present invention are apparent in the examples that follow. It should be understood, however, that the examples, while indicating embodiments of the present invention, are given by way of illustration only, not limitation. Various changes and modifications within the scope of the invention will become apparent to those skilled in the art from the examples.

Example 1. Vector Design

[0154] Exemplary methods and designs of generating rAAV expression constructs (rAAV vectors) comprising coding sequences of human iduronate 2-sulfatase (IDS or I2S) and variations of the same are provided in this Example. In this study, recombinant AAV vector (rAAV8) was used. The basic design of an rAAV vector comprises an expression cassette flanked by inverted terminal repeats (ITRs): a 5'-ITR and a 3'-ITR. These ITRs mediate the replication and packaging of the vector genome by the AAV replication protein Rep and associated factors in vector producer cells. Typically, an expression cassette contains a promoter, a coding sequence, a polyA tail and/or a tag, as shown in Figure 1A. An expression construct encoding human IDS (hIDS) was designed and prepared using standard molecular biology techniques. The coding sequence for the hIDS was inserted downstream of a promoter, hTTR (human transthyrethin promoter). Additionally, liver-specific cis-acting regulatory module (CRM) was inserted upstream of the promoter, and a minute virus of mice (MVM) intron sequence was inserted downstream of the promoter. This regulatory and promoter combination was tested for high transduction level, as shown in the examples that follow. Furthermore, the WPRE sequence was inserted downstream of the coding region. Without wishing to be bound by theory, this element creates a tertiary structure that increases the mRNA stability. Other mechanisms of function have been described for WPRE including, for example, improving transcript termination and facilitating mRNA nuclear export. Figure 1B shows schematic representations of the expression constructs described above. The expression construct was then cloned into an AAV plasmid backbone and confirmed by sequencing. Vectors were packaged in viral particles and stored.

[0155] Any number of variations of the above scheme can be performed. Alternative constructs can be obtained by replacing the WPRE sequence with a SUMF1 (sulfatase-modifying factor 1) coding sequence preceded by an internal ribosomal entry site (IRES), as

shown in Figure 2. Based on published data (Fraldi A et al. 2007. Biochem J. 403:305-312), the presence of SUMF1 increases I2S activities when tested in *in vitro* cell-based assays. Additional *in vivo* data published in the same study demonstrated more significant improvement of sulfamidase (SGSH) activity in the presence of SUMF1. Provided that SUMF1 is required to activate the Formyl glycine catalytic residue of its substrate, including IDS, in the endoplasmic reticulum, co-expression of IDS and SUMF1 from the same vector can ensure that there is increased amount of SUMF1 to process the increased amount of IDS at a cellular level, therefore increasing the probability of activated IDS to be processed prior to trafficking to the lysosome, or exiting the cells to travel to other cells for update.

Codon Optimization

[0156] Additionally, the coding sequences for the IDS or SUMF1 were codon-optimized based on multiple parameters, such as codon adaptation index (CAI), CpG site count, GC content, and repetitious base sequences. High CAI was preferred to utilize more frequently used codons and to potentially increase transgene product expression level from the vector. CpG sites, which can elicit immune response, were reduced. Repetitious bases were also removed. A web-based multi-objective optimization platform for synthetic gene design called COOL (Codon Optimization Online) and internal codon usage frequency table were used for this purpose. Additionally, potential splicing sites were manually removed. The characteristics of the optimized hIDS and SUMF1 coding sequences are summarized in Table 4, and the schematics for the representative constructs of hIDS-WPRE and hIDS-IRES-SUMF1 are shown in Figure 1B and Figure 2, respectively. Any number of variations of the above scheme can be performed. For example, more than one promoter may be used, and/or an IRES sequence may be introduced upstream of the coding region. Additionally, different combinations of regulatory region, promotor, and intron can be contemplated.

Table 4: Exemplary characteristics of optimized hIDS and SUMF1 coding sequences

Coding sequence	CpG	CAI	CG%	Repeats
hIDS WT	56	0.7645	51.97	896
hIDS COOLopt	24	0.8403	54.14	0
hIDS AUSopt	30	0.8519	54.51	0

SUMF1 WT	49	0.75	55.91	455
SUMF1 COOLopt	7	0.9038	54.04	0
SUMF1 AUSopt	13	0.8572	54.58	0

Example 2. Expression of rAAV-driven hIDS-WPRE expression in vivo

[0157] This example illustrates the potency of the optimized constructs for rAAV-driven IDS expression *in vivo*. Mice were injected with control vector (rAAV-XL032) (Group A); or test samples rAAV-XL024 (hIDS wt-WPRE) construct (Group B) or rAAV-XL026 (hIDS-AUSopt-WPRE) construct (Group C), as depicted in Figure 1A and Figure 1B. Mice of six weeks of age received 5x 10^9 vg of vectors in a volume of 200 µl via the tail vain, and serum samples were collected at 2 days, 7 days, 21 days, 8 weeks, and 12 weeks post injection. Mice were sacrificed at 12 weeks, and the tissue samples were harvested. A group of age-matched wild-type mice and a group of the age-matched, untreated *IDS*-KO mice were used as positive and negative controls, respectively. The experimental design is summarized in Table 5, below.

Table 5. Exemplary *in vivo* study using rAAV vectors that encode hIDS

Group	Condition	Treatment	Volume	Dose	N/group
A	Control vector	rAAV-XL032 (hIDS wt-WPRE)	200µl	5x 10^9 vg	8
B	Test	rAAV-XL024 (hIDS wt-WPRE)	200µl	5x 10^9 vg	8
C	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	5x 10^9 vg	8
D	Positive Control	WT; Uninjected	-	-	6
E	Negative Control	IDS-KO; Uninjected	-	-	6
F	Positive Control	IDS Enzyme (Intrathecal to KO mice)	10µl	250 µg/dose	5

[0158] Vector-mediated expression quantity was determined by ELISA using B85 antibody. Results are depicted in Figure 3. Mice injected with rAAV vectors of the optimized constructs, rAAV-XL024 (Group B) and rAAV-XL026 (Group C) showed higher hIDS concentration in serum compared to mice injected with the control vector, rAAV-XL032 (Group A) which is a vector that encodes a codon-optimized hIDS by a different algorithm that does not reduce CpG sites and a shortened version of wild type WPRE sequence. For all groups injected with vectors (Groups A-C), the hIDS concentration increased until day 21, and the hIDS level was maintained up to 12 weeks (84 days) after a single injection. hIDS was not detected for all WT and IDS-KO groups as they do not express human IDS.

[0159] Similar results were obtained when the level of hIDS activity was tested as shown in Figure 4A. For all groups injected with vectors (Groups A-C), the activity levels increase until day 21. Similar to the expression results, the hIDS activity in mouse groups injected with rAAV vectors of optimized constructs (Group B and Group C) was higher than in mouse group with the control vector rAAV-XL032 (Group A). The hIDS activity level of WT mouse group was relatively stable throughout the study period. The rAAV vector used in this example was rAAV8.

[0160] In tissues, hI2S was measured using ELISA as shown in Figure 4B. For rAAV-XL024 (Group B), hI2S was detectable in liver and kidney but not brain. Heart and spleen were not analyzed for this group. For rAAV-XL026 (Group C), hI2S was detectable in liver, kidney, heart, spleen, and brain. For rAAV-XL032 (Group A), hI2S was detectable in liver, kidney, and brain. Heart and spleen were not analyzed for this group.

[0161] In tissues, hI2S was measured using activity as shown in Figure 4C. For rAAV-XL024 (Group B), hI2S activity was detectable in liver, kidney, and brain. Heart and spleen were not analyzed for this group. For rAAV-XL026 (Group C), hI2S was detectable in liver, kidney, heart, spleen, and brain. For rAAV-XL032 (Group A), hI2S was detectable in liver, kidney, and brain. Heart and spleen were not analyzed for this group. Enzyme activity levels were compared to the WT group as shown in Figure 4D. In serum and somatic tissues, all vector dosed groups contained levels of activity greater than WT. In brain, levels of rAAV-XL026 and rAAV-XL032 were less than WT.

Example 3. GAG clearance by gene therapy with hIDS-WPRE constructs

[0162] The enzyme iduronate-2-sulfatase (IDS) removes the sulfate group from the glycosaminoglycans (GAGs), dermatan and heparan sulfates, and its absence or inactivity results in mucopolysaccharidosis type II (MPSII), or Hunter syndrome, a lysosomal storage disorder. Therefore, GAG clearance was measured to evaluate the potency of hIDS expressed by the optimized rAAV constructs. Brain, liver, and kidney tissues were extracted from mouse groups shown in Table 5 at week 12, and the GAG level in each tissue was measured.

[0163] As shown in Figure 5A, in the brain the GAG level was slightly reduced for rAAV-XL026 (Group C) and rAAV-XL032 (Group A) compared to untreated IDS-KO mice. As shown in Figure 5B and Figure 5C, the GAG level was significantly reduced in liver and kidney when the mice were injected with rAAV vectors encoding hIDS. The GAG level was similar to the GAG level found in untreated WT mice.

Example 4. Reduction in heparin sulfate (HS) and dermatan sulfate (DS) by gene therapy with hIDS-WPRE constructs

[0164] This example illustrates the reduction in heparan sulfate and dermatan sulfate levels by expression of hIDS in mice. The enzyme iduronate-2-sulfatase (IDS) removes the sulfate group from the glycosaminoglycans (GAGs), dermatan and heparan sulfate, and its absence or inactivity results in accumulation of GAGs resulting in mucopolysaccharidosis type II (MPSII), or Hunter syndrome, a lysosomal storage disorder.

[0165] In order to evaluate the potency of hI2S expressed by optimized rAAV constructs, GAG clearance was measured.

[0166] Glycosaminoglycans (GAGs) were measured using a liquid chromatography - mass spectrometry (LC/MS) assay that can detect heparan sulfate (HS) and dermatan sulfate (DS) in liver, kidney, heart, spleen, and brain tissues extracted at week 12 from the mouse groups shown in Table 5. The results are shown in Figure 6A-B.

[0167] For rAAV-XL024 (Group B), HS and DS GAGs were reduced compared to untreated KO (Group E) in liver and kidney, and to a lesser extent in brain. Heart and spleen were not analyzed for this group. For rAAV-XL026 (Group C), HS and DS GAGs were reduced compared to untreated KO (Group E) liver, kidney, heart, spleen, and to a lesser

extent in brain. For rAAV-XL032 (Group A), HS and DS GAGs were reduced compared to untreated KO (Group E) in liver and kidney, and to a lesser extent in brain.

[0168] The HS and DS GAG levels were normalized to untreated IKO = 0% reduction and WT = 100% reduction as shown in Figure 6B.

[0169] Overall, all vector dosed groups showed % reduction in HS and DS GAGs in somatic tissues and % GAG reduction in brain HS and DS GAGs. The results showed that expression of I2S resulted in reduced GAG levels in mouse tissue.

Example 5. Reduction in lysosomal storage compartment by gene therapy with hIDS-WPRE constructs in mice

[0170] This example illustrates a reduction in lysosomal storage compartment as detected by LAMP1 staining upon enzyme replacement therapy with h*IDS*-WPRE constructs in mice, for example, in treatment of Mucopolysaccharidosis (MPSII).

[0171] Briefly, LAMP1 staining was used to measure the lysosomal storage compartment in mice. Various tissues such as liver and brain hippocampus, thalamus, corpus callosum, cortex, cerebellum and stratum were stained. In this example, a control mouse group was administered with I2S enzyme intrathecally (IT) five times during the experimental period, at day 0, day 7, day 14, day 21 and day 28 (Group F in Table 5). Reduced LAMP1 staining indicates substrate reduction, an improvement in the pathology of the KO mice.

[0172] The results are shown in Figure 7A and Figure 7B. Significant reduction in LAMP1 was observed in liver tissues of mice injected with vectors as seen in Figure 7A. The effect in liver was comparable to WT mice, and mouse group treated with ERT.

[0173] Brain is shown in Figure 7B. Compared to the untreated KO (Group E), reductions in LAMP1 staining positivity reached statistical significance in hippocampus with rAAV-XL026, corpus callosum with rAAV-XL024 and rAAV-XL026, and thalamus with rAAV-XL024 and rAAV-XL026.

[0174] Overall, the results showed that in mouse liver and brain tissues, administration of IDS resulted in reduced lysosomal storage compartment as measured by LAMP1 staining.

Example 6. Expression of rAAV-driven hIDS-IRES-SUMF1 expression in vivo

[0175] This example compares the *in vivo* expression and activity of hIDS from administration of hIDS-IRES-SUMF1 vectors relative to hIDS-WPRE vectors in mice.

[0176] SUMF1 is required to activate the FGly catalytic residue of IDS, so a comparison was performed using vectors that expressed hIDS and SUMF1 relative to vectors that expressed hIDS-WPRE.

[0177] Mice were injected with rAAV vectors expressing hIDS AUSopt-WPRE construct (rAAV-XL026; Group C), hIDS-IRES-SUMF1 constructs (rAAV-XL027; Group G, and XL029; Group H), or SUMF1 construct (rAAV-XL030; Group I) as a negative control. The schematics for these constructs are depicted in Figure 1B and Figure 2.

[0178] Mice of six weeks of age received 5x 10^9 vg of vectors in a volume of 200 µl via the tail vain, and serum samples were collected at 2 days, 7 days and 21 days post injection. A group of age-matched wild-type mice were used as positive and negative controls, respectively. The exemplary *in vivo* study is summarized in Table 6.

Table 6. Exemplary *in vivo* study using rAAV vectors that encode hIDS and SUMF1 relative to hIDS and WPRE

Group	Condition	Treatment	Volume	Dose
C	Control	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	5x 10^9 vg
D	Positive Control	WT; Uninjected	-	-
G	Test	rAAV-XL027 (hIDS wt-IRES-SUMF1 wt)	200µl	200µl
H	Test	rAAV-XL029 (hIDS AUSopt-IRES-SUMF1AUSopt)	200µl	200µl
I	Negative Control	rAAV-XL030 (SUMF1 wt)	200µl	200µl

[0179] Expression levels of IDS in serum was quantified by ELISA. Results are depicted in Figure 8. Unexpectedly, mice injected with rAAV-XL027 and rAAV-XL029

expressing both hIDS and SUMF1 showed lower hI2S concentration in serum at day 21 compared to mice injected with the rAAV-XL026 vector (Group C), which expresses only the hIDS.

[0180] Similar results were obtained when the level of hI2S activity was tested as shown in Figure 9. For all groups injected with vectors expressing both hIDS and SUMF1 (Groups G-H), the activity level was lower than the group injected with vector expressing hIDS-WPRE (Group C). Without wishing to be bound by a particular theory, it is believed that the WPRE element in rAAV-XL026 vector creates a tertiary structure that increase the mRNA stability and subsequently yields higher protein expression and therefore increased activity. Other mechanisms of function have been described for WPRE including, for example, improving transcript termination and facilitating mRNA nuclear export. SUMF1, on the other hand, creates more active forms of hIDS, but the quantity of hIDS expressed with WPRE element is significantly higher. The rAAV vector used in this example is rAAV8.

[0181] The results showed that IDS enzyme levels and activity were higher upon expression of h*IDS*-WPRE vectors relative to h*IDS*-IRES-SUMF1 vectors.

Example 7. Long-term expression of rAAV-driven hIDS-WPRE in serum and tissue in vivo in mice

[0182] This example compares the *in vivo* expression and activity of hIDS in mice administered one of three doses of h*IDS*-WPRE. Expression levels and activity were evaluated over about 12-13 months in serum and tissues.

[0183] Mice were injected with rAAV vectors expressing h*IDS* AUSopt-WPRE construct (rAAV-XL026) at doses of 5x 10^9 vg (Group D), 2.5x 10^{10} vg (Group E), 1.25x 10^{11} vg (Group F), null vector construct rAAV-MY011 at a dose of 1.25x 10^{11} vg (Group G) as a negative control. The schematics for these constructs are depicted in Figure 1B and Figure 2.

[0184] Five- to seven-week-old mice received vectors in a volume of 200 µl via the tail vein, and serum samples were collected at days 7, 14, 28, 56, 84, 112, 140, 168, 196, 224, 252, 280, 308, 336, and 364 post injection. A group of age-matched wild-type mice and a group of the age-matched, untreated *IDS*-KO mice were used as positive and negative controls, respectively. An exemplary *in vivo* study is summarized in Table 7.

Table 7. Exemplary *in vivo* study using rAAV vectors and monitored for 12 months

Group	Condition	Treatment	Volume	Dose	N/group
D	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	5x 10^9 vg	12
E	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	2.5x 10^{10} vg	12
F	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	1.25x 10^{11} vg	12
G	Negative Control	rAAV-MY011 (null vector)	200µl	1.25x 10^{11} vg	11
H	Negative Control	IDS-KO; Uninjected	-	-	9
I	Positive Control	WT; Uninjected	-	-	6

[0185] ***Serum hI2S levels and activity:*** Expression levels of serum hI2S was determined by ELISA. Results are depicted in Figure 10. Concentrations of hI2S increased with increasing doses of the rAAV-XL026 vector (Groups D, E, F), and hI2S remained detectable until the final timepoint of 364 days post-injection. The rAAV vector used in this example is rAAV8.

[0186] Similar results were obtained when the level of hI2S activity was tested as shown in Figure 11. Concentrations of hI2S activity increased with increasing doses of the rAAV-XL026 vector (Groups D, E, F), and hI2S activity remained detectable until the final timepoint of 364 days post-injection.

[0187] ***Tissue hI2S levels and activity:*** In tissues, hI2S was measured using ELISA as shown in Figure 12A. For rAAV-XL026 (Group D, E, F), hI2S was detectable in liver, spleen, kidney, heart, lung, and brain and increased with increasing doses.

[0188] In tissues, hI2S activity was measured as shown in Figure 12B. rAAV-XL026 (Group D, E, F), hI2S activity was detectable in liver, spleen, kidney, heart, lung, and brain and increased with increasing doses.

[0189] Glycosaminoglycans (GAGs) were measured using a liquid chromatography - mass spectrometry (LC/MS) assay that can detect heparan sulfate (HS) and dermatan sulfate (DS) in liver, kidney, heart, spleen, lung, and brain tissues extracted at day 364 from the mouse groups shown in Table 7. The results are shown in Figure 12C.

[0190] In the somatic tissues, rAAV-XL026 at all doses (Group D, E, F) greatly reduced HS and DS GAG levels. The HS and DS GAG levels were normalized to untreated IKO = 0% reduction and WT = 100% reduction as shown in Figure 12D. All three doses of rAAV-XL026 (Group D, E, F) showed 100% reductions in HS and DS GAGs in somatic tissues. In brain HS, GAGs were decreased by 17% for rAAV-XL026 Group D, 38% for rAAV-XL026 Group E, and 50% for rAAV-XL026 Group F.

[0191] Various tissues such as liver and brain hippocampus, thalamus, white matter, cortex, cerebellum and striatum were stained for LAMP1. Brain is shown in Figure 12E. Compared to the untreated KO (Group H), reductions in LAMP1 staining positivity reached statistical significance in hippocampus with rAAV-XL026 Group D, Group E, and Group F, in striatum with rAAV-XL026 Group E and Group F, in white matter with rAAV-XL026 Group E and Group F, in thalamus with rAAV-XL026 Group E and Group F, in cortex with rAAV-XL026 Group E and Group F, and in cerebellum with rAAV-XL026 Group E and Group F.

[0192] ***Bone volume:*** Body structure was investigated using micro-computed tomography (micro-CT). Bone volume was measured over time at mouse ages 7, 9, 11, and 13 months using micro-CT as shown in Figure 13A and Figure 13B.

[0193] In Figure 13A and Figure 13B, line A corresponds to rAAV-XL026 Group D, line B to rAAV-XL026 Group E, line C to rAAV-XL026 Group F, line D to rAAV-MY011 Group G negative control, line E to untreated *IDS*-KO Group H negative control, and line F to untreated wild-type Group I positive control.

[0194] In the humerus (Figure 13A), the measurements of bone volumes for rAAV-XL026 Group D, Group E, and Group F overlay with the positive control wild-type Group I and volumes for these groups are lower than the negative controls rAAV-MY011 Group G and untreated *IDS*-KO Group H.

[0195] In the zygomatic arch of the cranium (Figure 13B) at 13 months, measurements of bone volumes for rAAV-XL026 Group D, Group E, and Group F overlay

with the positive control wild-type Group I and volumes for these groups are lower than the negative controls rAAV-MY011 Group G and untreated *IDS*-KO Group H.

[0196] These results showed lower bone volume in the humerus and zygomatic arch bone upon expression of I2S, which was maintained over 13 months.

[0197] Overall, the results showed that serum and tissue I2S expression and activity levels and resultant effects on bone volume were maintained over about 12-13 months.

Example 8. Monitoring of rAAV-driven hIDS-WPRE expression in vivo for 3 months after low-dose administration in mice

[0198] This example illustrates the *in vivo* expression and activity of hI2S in mice administered hI2S over 3 months at a dose of between 5 x 10^6- 5 x 10^9 vg.

[0199] Mice were injected with rAAV vectors expressing h*IDS* AUSopt-WPRE construct (rAAV-XL026) at doses of 5x 10^6 vg (Group C), 5x 10^7 vg (Group D), 5x 10^8 vg (Group E), 5x 10^9 vg (Group F), null vector construct rAAV-MY011 at a dose of 5x 10^9 vg (Group B) as a negative control.

[0200] Five- to seven-week-old mice were administered 200 µl of vector via the tail vain, and serum samples were collected at days 14, 28, 56, and 84 post injection. A group of age-matched wild-type mice and a group of the age-matched, untreated *IDS*-KO mice were used as positive and negative controls, respectively. The exemplary in vivo study is summarized in Table 8.

[0201] Table 8. Exemplary *in vivo* study using rAAV vectors and monitored for 3 months

Group	Condition	Treatment	Volume	Dose	N/group
A	Negative Control	IDS-KO; Uninjected	-	-	8
B	Negative Control	rAAV-MY011 (null vector)	200µl	2.5x 10^9 vg	8
C	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	5x 10^6 vg	8
D	Test	rAAV-XL026	200µl	5x 10^7 vg	8

		(hIDS AUSopt-WPRE)			
E	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	5x 10^8 vg	8
F	Test	rAAV-XL026 (hIDS AUSopt-WPRE)	200µl	5x 10^9 vg	8
G	Positive Control	WT; Uninjected	-	-	5

[0202] ***Serum IDS levels and activity:*** Expression levels of hIDS in serum of mice administered low doses of the h*IDS*-WPRE vector was determined by ELISA, and shown in Figure 14A. Concentrations of hI2S decreased with decreasing doses of the rAAV-XL026 vector for Groups D, E, F, and hI2S was undetectable for group C. The rAAV vector used in this example is rAAV8.

[0203] Serum I2S activity levels were measured as in Figure 14B. Concentrations of hI2S decreased with decreasing doses of the rAAV-XL026 vector for Groups D, E, F, and hI2S was undetectable for group C.

[0204] Overall, Groups D and E administered 5 x 10^8 vg and 5 x 10^9 vg doses respectively, showed an increase in I2S expression and activity over untreated mice, and levels and activity of I2S were maintained over 84 days or 3 months.

[0205] ***Tissue I2S levels and activity:*** In liver, spleen, kidney, heart, lung, bone marrow, quadriceps, and brain, as measured by ELISA, the untreated KO (Group A) and untreated WT (Group J) and negative control rAAV-MY011 (Group B) groups did not contain detectable human I2S (hI2S) protein (Figure 14C). From 5x 10^7 vg to 5x 10^9vg, the rAAV-XL026 Groups D, E, and F contained increasing concentrations of hI2S in tissues with increasing dose.

[0206] I2S activity was measured in liver, spleen, kidney, heart, lung, bone marrow, quadriceps, and brain as shown in Figure 14D. The rAAV-MY011 (Group B) contained I2S activity in all tissues that was similar to untreated KO (Group A). From 5x 10^7 vg to 5x 10^9 vg, the rAAV-XL026 Groups D, E, and F contained increasing concentrations of I2S activity in tissues with increasing dose. I2S activity levels in each tissue of the dosed KO mice can be compared to the untreated WT levels as shown in Figure 14E.

[0207] Tissue levels and activity of I2S were increased in a dose-dependent manner in mice administered low doses of rAAV vector comprising hI2S-WPRE.

[0208] ***Tissue GAG levels:*** Glycosaminoglycans (GAGs) were measured using a liquid chromatography - mass spectrometry (LC/MS) assay that can detect heparan sulfate (HS) and dermatan sulfate (DS) in liver, spleen, kidney, heart, lung, bone marrow, quadriceps, skin, and brain tissues extracted at day 84 from the mouse groups shown in Table 8. The HS and DS GAG levels were normalized to untreated IKO = 0% reduction and WT = 100% reduction as shown in Table 9 and Table 10.

[0209] The 5x 10^8 vg rAAV-XL026 Group E reduced HS and GAG levels by >90% in all somatic tissues measured. In brain, the 5x 10^8 vg rAAV-XL026 Group E reduced HS and GAG levels in brain by 8.8% and 41%, respectively. The 5x 10^9 vg rAAV-XL026 Group F reduced HS and GAG levels by >95% in all somatic tissues measured. In brain, the 5x 10^9 vg rAAV-XL026 Group F reduced HS and GAG levels in brain by 39% and 78%, respectively.

[0210] HS and GAG levels were reduced in somatic tissues and brain upon administration of 5 x 10^7, 5 x 10^8 or 5 x 10^9 vg rAAV comprising *IDS*-WPRE.

Table 9. Normalized HS GAG reduction

Tissue	**rAAV-XL026 5 x 10^7 vg**	**rAAV-XL026 5 x 10^8 vg**	**rAAV-XL026 5 x 10^9 vg**
Liver	99%	100%	100%
Spleen	79%	98%	100%
Kidney	37%	91%	101%
Heart	20%	97%	99%
Lung	50%	97%	100%
Bone marrow	92%	99%	100%
Quadriceps	59%	99%	102%

Skin	62%	94%	100%
Brain	11%	8.8%	39%

Table 10. Normalized DS GAG reduction

Tissue	**rAAV-XL026 5 x 10^7 vg**	**rAAV-XL026 5 x 10^8 vg**	**rAAV-XL026 5 x 10^9 vg**
Liver	92%	99%	99%
Spleen	66%	95%	100%
Kidney	22%	97%	100%
Heart	43%	99%	99%
Lung	32%	100%	100%
Bone marrow	88%	101%	101%
Quadriceps	86%	100%	100%
Skin	92%	94%	106%
Brain	17%	41%	78%

[0211] ***Lysosomal storage:*** Various tissues such as liver, spleen, kidney, heart, and brain hippocampus, thalamus, white matter, cortex, cerebellum and striatum were stained for LAMP1. Somatic tissues are shown in Figure 14F. In the somatic tissues including liver, spleen, and heart, LAMP1 IHC staining positivity was significantly reduced compared to the naïve KO control for rAAV-XL026 5x 10^7 vg dose Group D, 5x 10^8 vg dose Group E, and 5x 10^9 vg dose Group F. In the kidney, LAMP1 staining positivity was significantly reduced compared to the naïve KO control for rAAV-XL026 5x 10^8 vg dose Group E, and 5x 10^9 vg dose Group F.

[0212] LAMP1 IHC staining in brain regions is shown in Figure 14G. In the brain, LAMP1 positivity staining was measured in hippocampus, striatum, thalamus, white matter, cortex, and cerebellum. In all of these regions the LAMP1 staining positivity for the rAAV-XL026 dose Groups C, D, E, and F was not significantly different from the untreated KO control Group A.

[0213] The results showed that lysosomal storage was decreased in somatic tissues upon administration of low doses of *IDS*-WPRE, but not significantly in the brain.

[0214] Overall, the results showed that upon administration of low doses of rAAV comprising h*IDS*-WPRE in mice, serum and tissue I2S expression and activity was maintained over about three months, and lysosomal storage was decreased in somatic tissues as determined by LAMP.

Example 9. Serum expression levels of hI2S transgene product in non-human primates (NHPs)

[0215] This example illustrates an exemplary PK/distribution study in non-human primates administered a hI2S transgene product.

[0216] Non-human male primates between about 1.8 to 2 years of age were administered the hI2S product via intravenous infusion. Low dose cohort animals received a low dose of hI2S of 1.25 x 10^{12} vg/kg (n=3) and high dose cohort animals received a higher dose of hI2S of 6.25 x 10^{12} vg/kg (n=6). Control animals received only formulation buffer.

[0217] Serum samples were collected at various time-points, starting from pre-dose, prior to administration of hI2S transgene, followed by sampling every 20 days up to 240 days. Three animals each from the low dose cohort and the high dose cohort, respectively, were sacrificed at 3 months.

[0218] Figure 15A demonstrates the hI2S transgene product serum concentration from the low dose cohort in NHP serum post-dose day 1-to 3-month necropsy. Sustained hI2S transgene product concentration was observed at 1.25 x 12 vg/kg (low dose).

[0219] Correspondingly, in Figure 15B, the hI2S transgene product enzyme activity in serum showed sustained levels up to 3 months. At 3 months, about 5-fold higher levels of I2S enzyme activity was observed relative to the NHP endogenous I2S activity levels measured in control animals that received only formulation buffer.

[0220] Figure 15C demonstrates the hI2S transgene product serum concentrations from high dose cohorts in three NHP post-dose day 1 to about 90 days or 3-months necropsy, and remaining three high-dose animals from post-dose day 1 to about 240 days or 8 months.

[0221] The individual animals showed a variable profile of hI2S transgene product and variability was observed in the initial maximum serum concentration level achieved. A rapid decline in hI2S transgene product serum concentration was observed in the animals that were followed until 3-month necropsy. The corresponding hI2S enzyme activity in serum showed an increase then a decline after day 21 to a plateau level of about 5200-7500 nmol/hr/mL) (Figure 15D), demonstrating sustained hI2S enzyme activity in serum even though the I2S transgene product concentration was reduced in the three animals that were terminated at 3 months.

[0222] The animals where serum hI2S transgene product concentration and serum I2S activity was followed to about 240 days (8 months) showed variability in their hI2S transgene product concentration profile. Two animals showed a decline in I2S transgene product concentration to about 100ng/mL or below by about 8 months. This was also associated with an observed decline in hI2S enzyme activity to about endogenous levels (formulation buffer) by about 8 months. One of the animals in this study showed sustained hI2S transgene product concentration and enzyme activity out to about 8 months.

[0223] Overall, the data demonstrated that even in the absence of immunosuppressants, sustained expression of the hI2S transgene product was achieved in non-human primates.

Example 10. Comparison of serum concentrations of hI2S transgene product and I2S enzyme activity in low and high dose non-human primates relative to anti-transgene product antibody (a.k.a. anti-hI2S ADA (Anti-Drug Antibody)) and anti-AAV8 ADA titration

[0224] This example illustrates a comparison between anti-hI2S ADA and anti-AAV8 ADA data from individual non-human primates plotted with hI2S transgene product concentration and enzyme activity.

[0225] Anti-hI2S ADA and anti-AAV8 ADA data from individual animals were plotted against hI2S transgene product concentration and enzyme activity (Figure 16). Only up to 3 months of ADA data were illustrated in these graphs. An initial decline in hI2S level

was observed within the first month in the absence of anti-hI2S ADA in some animals (2001, 3002, 4003, 4001 and 4002). The decline in hI2S levels was not dose dependent. The initial decline of hI2S transgene product concentration in NHP serum (within the first month) varied from animal to animal (10% - 80%).

[0226] The decline in hI2S transgene product in serum after the equilibrium state was correlated with the presence of anti-hI2S ADA in the cohort that received a high dose of hI2S. "Equilibrium state" or "re-established state" refers to "steady state" or plateau levels of I2S enzyme or I2S enzyme activity. The presence of anti-AAV8 ADA in some animals (3002 and 4001) may prevent re-dosing. None of the animals in the low dose cohort showed the presence of anti-hI2S ADA in serum.

[0227] The results showed that the decline in serum I2S levels was correlated with the presence of anti-hI2S ADA in non-human primates that received a high dose of rAAV-XL026 comprising h*IDS*-WPRE.

Example 11. Individual non-human primate liver hI2S transgene product concentration profile

[0228] This example illustrates the hI2S transgene product concentration profile in the liver of non-human primates.

[0229] Liver biopsies were carried out from both low dose (Figure 17A) and high dose (Figure 17B) animals at 1 month and 2 months and terminal liver samples were collected at 3 months. The three animals from high dose cohort that were intended for 12 months necropsy had a biopsy at 3 months and 6 months.

[0230] Each biopsy was taken from left and right lobes. The average hI2S transgene product concentration between the left and right lobes in the liver from low dose cohort ranged from an average of 38.2ng/mL to 42.5ng/mL between 1 month and 3 months.

[0231] The hI2S transgene product concentration in the liver from high dose cohort ranged from 175.1ng/mL to 367.6ng/mL at 1 month. One of the animals (animal 3001) showed a large difference between the left (101.8ng/mL) and right (528.3 ng/mL) lobes.

[0232] At 2 months, two animals (animal 3001 and animal 3002) showed a rapid decline of liver hI2S transgene product concentration from 1 month to 2 months. Other

animals also showed decline from 2 months to 6 months, including animal 3003, but to a lesser extent.

[0233] Overall, these results showed individual concentration profiles of I2S transgene product in non-human primate tissues from 1 to 3 months post-administration of rAAV-XL026.

Example 12. Comparative concentration profile of hI2S transgene product in non-human primate tissues

[0234] This example illustrates a comparative concentration profile of hI2S transgene product in non-human primate tissues.

[0235] Tissue necropsy was performed on animals from the low dose cohort and 3 animals from the high dose cohort at 3 months.

[0236] The hI2S transgene product concentration was measured by ELISA in tissue homogenate from various organs including kidney, spleen, lung, heart and bone marrow. Although the hI2S transgene product concentration in serum from high dose cohort at 3 months showed a significant lower level (<250ng/mL) in the presence of anti-hI2S ADA when compared to those of low dose cohort (>640ng/mL), the average hI2S concentration in kidney (Figure 18A) and spleen (Figure 18B) showed a higher level in the high dose cohort than that of lower dose.

[0237] In the presence of anti-hI2S ADA, the 3 animals from the high dose cohort showed lower hI2S transgene product concentration in the lung (Figure 18C) and heart (Figure 18D), whereas in the bone marrow (Figure 18E), hI2S tissue concentration was comparable between animals in the high and low dose cohort. Since the tissues were from healthy NHP, endogenous I2S is detected, for example, in the formulation buffer group in the lung.

[0238] The results showed the comparative concentration profile of I2S in different non-human primate tissues post-treatment of rAAV-XL026.

Example 13. Comparative hI2S tissue exposure in non-human primates relative to IDS knockout mice

[0239] This example illustrates the comparative I2S concentration and enzyme activity in various target tissues between non-human primates and *IDS* KO mice.

[0240] The hI2S tissue enzyme activity and corresponding percentage of HS GAG reduction from $2.5x10^{11}$vg/kg and $2.5x10^{10}$vg/kg in *IDS* KO mice are shown in Table 11.

[0241] In mice, greater than 95% HS GAG reduction was observed at a dose of about $2.5x10^{10}$ vg/kg for most tissues except for the kidney, where a reduction of about 91% was observed. In the heart and bone marrow, data from the low dose cohort of non-human primates showed higher hI2S tissue concentration and enzyme activity than those of *IDS* KO mice at $2.5x10^{10}$ vg/kg.

[0242] In lung tissue, in *IDS* KO mice, there was only 30% of WT I2S enzyme activity yet this level showed HS GAG reduction of 97%. In non-human primates, administered a low dose of hI2S, enzyme activity in the lung showed a greater % of WT hI2S enzyme activity at 40%. By administering rAAV-XL026 via IV infusion at 1.25×10^{12}vg/kg to NHP, sufficient hI2S transgene production exposure in the lung was achieved that can result in a reduction of HS GAG to ≥95% in a disease model.

Table 11. Comparison of hI2S concentration and enzyme activity and GAG reduction in knock-out mice and non-human primates.

	hI2S concentration in NHP tissues at 3 months necropsy above vehicle control	hI2S concentration (at 2.5×10^{10} vg/kg) in KO mouse tissues at 3 months necropsy.	NHP tissue hI2S enzyme activity expressed as a % of WT	KO mouse tissue hI2S enzyme activity (at 2.5×10^{10} vg/kg) expressed as % of WT	KO mouse tissue HS GAG reduction (at 2.5×10^{10} vg/kg) expressed as % of the KO vehicle	KO mouse hI2S enzyme activity (at 2.5×10^{11} vg/kg) expressed as % of WT	KO mouse tissue HS GAG reduction (at 2.5×10^{11} vg/kg) expressed as % of the KO vehicle.
Lung (low)	7.1 ng/mg	2 ng/mg	40%	30%	~97%	270%	100%
Lung (high)	0.9 ng/mg		20%				
Heart (low)	12.5 ng/mg	1.4 ng/mg	232%	210%	~97%	2480%	100%
Heart (high)	4.8 ng/mg		174%				

Bone marrow (low)	29.6 ng/mg	15.1 ng/mg	210%	310% (10% at 2.5 x 10^9 vg/kg)	100% (92% at 2.5 x 10^9 vg/kg)	46,000%	100%
Bone marrow (high)	30.4 ng/mg		170%				
Kidney (low)	4.5 ng/mg	3 ng/mg	10%	90%	~91%	1380%	100%
Kidney (high)	31.6 ng/mg		180%				
Spleen (low)	36.2 ng/mg	17 ng/mg	100%	140%	~95%	1170%	100%
Spleen (high)	58.4 ng/mg		190%				

EQUIVALENTS AND SCOPE

[0243] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the following claims:

We claim:

1. A recombinant adeno-associated virus (rAAV) vector comprising an AAV8 capsid and a codon-optimized sequence encoding a human iduronate-2-sulfatase (I2S) enzyme.

2. A recombinant adeno-associated virus (rAAV) vector comprising an AAV9 capsid and a codon-optimized sequence encoding human iduronate-2 sulfatase (I2S) enzyme.

3. The rAAV vector of claim 1 or 2, wherein the codon-optimized sequence encoding a human I2S enzyme comprises a sequence having at least about 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to SEQ ID NO: 6.

4. The rAAV vector of claim 3, wherein the codon-optimized sequence encoding a human I2S enzyme comprises a sequence identical to SEQ ID NO: 11 or 12.

5. The rAAV vector of claim 1, wherein the vector further comprises a liver-specific promoter.

6. The rAAV vector of claim 5, wherein the liver-specific promoter is transthyretin promoter (TTR).

7. The rAAV vector of claim 5 or 6, wherein the vector further comprises a 5' and a 3' inverted terminal repeat (ITR), an intron upstream of the I2S sequence, and a cis-acting regulatory module (CRM).

8. The rAAV vector of claim 2, wherein the vector further comprises a ubiquitous promoter.

9. The rAAV vector of claim 8, wherein the vector further comprises a 5' and a 3' inverted terminal repeat, an intron upstream of the I2S sequence, and a cis-acting regulatory module (CRM).

10. The rAAV vector of any one of the preceding claims, wherein the rAAV vector comprises a sulfatase modifying factor 1 (SUMF1).

11. The rAAV vector of claim 10, wherein the SUMF1 is preceded by an internal ribosome entry site (IRES).

12. The rAAV vector of any one of the preceding claims, wherein the vector further comprises a WPRE sequence.

13. The rAAV vector of claim 7 or 9, wherein the intron is a minute virus of mice (MVM) or SV40 intron.

14. The rAAV vector of claim 7 or 9, wherein the CRM is liver-specific CRM.

15. The rAAV vector of claim 7 or 9, wherein the CRM is a neuronal-specific CRM.

16. The rAAV vector of claim 7 or 9, wherein the CRM is CRM8.

17. The rAAV vector of claim 7 or 9, wherein the vector comprises at least three CRMs.

18. A recombinant adeno-associated virus (rAAV) comprising an AAV8 capsid and an rAAV vector, said vector comprising:

a. a 5' inverted terminal repeat (ITR);

b. a cis-acting regulatory module (CRM);

c. a liver specific promoter;

d. a minute virus of mice (MVM);

e. a sequence encoding a human iduronate-2-sulfatase (I2S) enzyme;

f. a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE); and

g. a 3' ITR.

19. The rAAV of claim 18, wherein the sequence encoding a human I2S enzyme is a wild type sequence or a codon-optimized sequence.

20. The rAAV of claim 19, wherein the sequence encoding a human I2S enzyme is a codon-optimized sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity with SEQ ID NO: 6.

21. The rAAV of claim 18, further comprising a sequence encoding a sulfatase modifying factor 1 (SUMF1) and an internal ribosome entry site (IRES).

22. A recombinant adeno-associated virus (rAAV) comprising an AAV9 capsid and an rAAV vector, said vector comprising:

a. a 5' inverted terminal repeat (ITR);

b. a cis-acting regulatory module (CRM);

c. a ubiquitous promoter;

d. a minute virus of mice (MVM);

e. a sequence encoding a human iduronate-2-sulfatase (I2S) enzyme;

f. a woodchuck hepatitis virus posttranscriptional regulatory element(WPRE); and

g. a 3' ITR.

23. The rAAV of claim 22, wherein the sequence encoding a human I2S enzyme is a wild type sequence or a codon-optimized sequence.

24. The rAAV of claim 23, wherein the sequence encoding a human I2S enzyme is a codon-optimized sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity with SEQ ID NO: 6.

25. The rAAV of claim 22, further comprising a sequence encoding a sulfatase modifying factor 1 (SUMF1) and an internal ribosome entry site (IRES).

26. The rAAV of any one of the preceding claims, wherein the rAAV vector does not comprise ApoB.

27. A method of treating a subject having Hunter syndrome (MPS II), comprising administering to the subject in need thereof an rAAV of any one of the preceding claims.

28. A method of treating a subject having Hunter syndrome (MPS II), comprising administering to the subject in need thereof a recombinant adeno-associated virus (rAAV) vector comprising an AAV8 or AAV9 capsid, and a promoter operably linked to a nucleic acid sequence that encodes iduronate-2-sulfatase (I2S), and wherein administering results in an increase in I2S enzymatic activity in the subject.

29. The method of claim 28, wherein the increase in I2S activity is detected in the serum of the subject.

30. The method of any one of claims 28-29, wherein the increase in I2S activity is detected in the liver of the subject.

31. The method of any one of claims 28-30, wherein the I2S activity is detected in the central nervous system (CNS).

32. The method of any one of claims 28-31, wherein the increase in I2S activity is detected in the brain of the subject.

33. The method of any one of claims 28-32, wherein the increase in I2S activity is detected in the hippocampus, thalamus, corpus callosum, cortex, cerebellum, or stratum of the brain.

34. The method of any one of claims 28-33, wherein the increase of I2S activity is detected in the kidney of the subject.

35. The method of any one of the claims 28-34, wherein the increase of I2S activity is maintained for at least 30, 60, 90, 120, 150, 180 days or more after a single administration.

36. The method of any one of the claims 28-35, wherein the level of I2S activity is measured by heparin sulfate assay.

37. The method of any one of the claims 28-35, wherein the level of I2S activity is measured by dermatan sulfate assay.

38. The method of any one of the claims 28-37, wherein the administering the AAV reduces the level of glycosaminoglycan (GAG) in the subject.

39. The method of claim 38, wherein the administering the AAV reduces the level of GAG in the serum of the subject.

40. The method of any one of claims 38-39, wherein the administering the AAV reduces the level of GAG in the liver of the subject.

41. The method of any one of claims 38-40, wherein the administering the AAV reduces the level of GAG in the kidney of the subject.

42. The method of any one of claims 38-41, wherein the administering the AAV reduces the level of GAG in the CNS of the subject.

43. The method of any one of claims 38-41, wherein the administering the AAV reduces the level of GAG in the brain of the subject.

44. The method of claim 43, wherein the administering the AAV reduces the level of GAG in in the hippocampus, thalamus, corpus callosum, cortex, cerebellum, or stratum of the brain.

45. The method of any one of claims 28-44, wherein the AAV is administered intravenously.

46. The method of any one of claims 30-44, wherein the AAV is administered intrathecally.

47. The method of claim 45 or 46, wherein the AAV is administered at dose of about 5 x 10^9 vg.

48. The method of any one of claims 30-47, wherein the administering of the rAAV does not elicit immune response.