

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2018278315 B2**

(54) Title
Therapeutics for glycogen storage disease type III

(51) International Patent Classification(s)
C12N 9/10 (2006.01) **C07K 14/81** (2006.01)
C07K 14/505 (2006.01)

(21) Application No: **2018278315** (22) Date of Filing: **2018.05.31**

(87) WIPO No: **WO18/222926**

(30) Priority Data

(31)	Number	(32)	Date	(33)	Country
	62/513,350		2017.05.31		US

(43) Publication Date: **2018.12.06**

(44) Accepted Journal Date: **2024.01.18**

(71) Applicant(s)
Ultragenyx Pharmaceutical Inc.

(72) Inventor(s)
Tachikawa, Kiyoshi;Perez-Garcia, Carlos Gustavo;Chivukula, Padmanabh;Bhaskaran, Hari;Cobaugh, Christian W.;Daugherty, Sean Christopher

(74) Agent / Attorney
Pizzeys Patent and Trade Mark Attorneys Pty Ltd, PO Box 291, WODEN, ACT, 2606, AU

(56) Related Art
WO 2014/130722 A1



(51) International Patent Classification:

C12N 9/10 (2006.01) C07K 14/81 (2006.01)
C07K 14/505 (2006.01)

(21) International Application Number:

PCT/US2018/035477

(22) International Filing Date:

31 May 2018 (31.05.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/513,350 31 May 2017 (31.05.2017) US

(71) Applicant: **ULTRAGENYX PHARMACEUTICAL INC.** [US/US]; 60 Leveroni Court, Novato, California 94949 (US).

(72) Inventors: **TACHIKAWA, Kiyoshi**; 9439 Fairgrove Lane #206, San Diego, California 92129 (US). **PEREZ-GAR-**

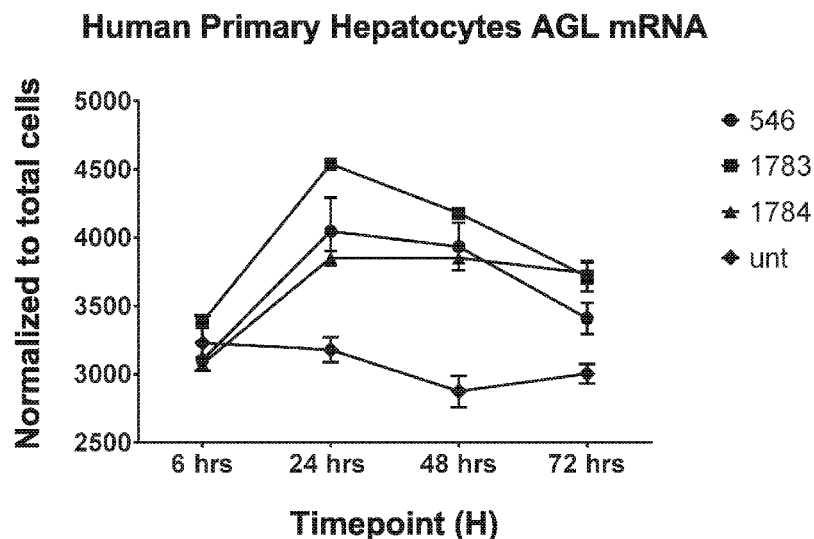
CIA, Carlos Gustavo; 4080 Florida Street, Apt. #2, San Diego, California 92104 (US). **CHIVUKULA, Padmanabh**; 1325 Pacific Highway Unit 2803, San Diego, California 92101 (US). **BHASKARAN, Hari Prakash**; 5927 Bark Street, San Diego, California 92105 (US). **COBAUGH, Christian W.**; 33 Beverly Road, Boston, Massachusetts 02461 (US). **DAUGHERTY, Sean Christopher**; 1400 Technology Lane, Apt. 305, Petaluma, California 94954 (US).

(74) Agent: **ROY, Madhuri** et al.; COOLEY LLP, 1299 Pennsylvania Avenue, NW Suite 700, Washington, District of Columbia 20004-2400 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,

(54) Title: THERAPEUTICS FOR GLYCOGEN STORAGE DISEASE TYPE III

FIG. 5



(57) Abstract: This invention provides a range of translatable polynucleotide and oligomer molecules for expressing a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity. The polynucleotide and oligomer molecules are expressible to provide the human AGL or a fragment thereof having AGL activity. The molecules can be used as active agents to express an active polypeptide or protein in cells or subjects. The agents can be used in methods for ameliorating, preventing, delaying onset, or treating a disease or condition associated with reduced activity of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) in a subject.



MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

THERAPEUTICS FOR GLYCOGEN STORAGE DISEASE TYPE III

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application Serial No. 62/513,350, filed May 31, 2017, which is herein incorporated by reference in its entirety for all purposes.

TECHNICAL FIELD OF THE INVENTION

[0002] This invention relates to the fields of molecular biology and genetics, as well as to biopharmaceuticals and therapeutics generated from translatable molecules. More particularly, this invention relates to methods, structures and compositions for molecules having the ability to be translated into active polypeptides or proteins, for use *in vivo* and as therapeutics.

DESCRIPTION OF TEXT FILE SUBMITTED ELECTRONICALLY

[0003] The contents of the text file submitted electronically herewith are incorporated herein by reference in their entirety: A computer readable format copy of the Sequence Listing (filename: ULPI_041_01WO_SeqList_ST25.txt, date recorded: May 30, 2018, file size: 226 kilobytes).

BACKGROUND OF THE INVENTION

[0004] Glycogen storage disease type III (also known as GSD III or Cori disease) is a rare (incidence 1:100,000) inborn error of glycogen metabolism caused by the deficiency of a glycogen debranching enzyme known as amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL). This autosomal recessive metabolic disorder is characterized by variable liver, cardiac muscle, and skeletal muscle involvement.

[0005] There are four subtypes GSD III, based on differences in tissue expression of the deficient enzyme, AGL. GSD IIIa accounts for approximately 85% of all GSD III and presents with liver and muscle involvement, resulting from enzyme deficiency in both liver and muscle. GSDIIIb accounts for approximately 15% and generally presents with only liver involvement, resulting from enzyme deficiency in the liver only. Meanwhile, GSDIIIc and GSDIIId are both extremely rare, with GSDIIIc believed to result from a deficiency of glucosidase debranching activity and GSDIIId believed to result from a deficiency of the transferase debranching activity.

[0006] In infancy and early childhood, liver involvement presents as ketotic hypoglycemia, hepatomegaly, hyperlipidemia, and elevated hepatic transaminases. In adolescence and adulthood, liver disease becomes less prominent. Hypertrophic cardiomyopathy develops in the majority of those with GSD IIIa, usually during childhood. Its clinical significance ranges from asymptomatic in the majority to severe cardiac dysfunction, congestive heart failure, and, rarely, sudden death. Skeletal myopathy manifesting as weakness is not usually evident in childhood, but slowly progresses, typically becoming prominent in adults.

[0007] In agreement with the organs affected, enhanced alanine transaminase (ALT), aspartate transaminase (AST), alkaline phosphatase (ALP), and/or creatine phosphokinase (CPK) activities are frequently present in the serum of GSD III patients.

[0008] As noted above, GSD III is caused by a deficiency in AGL. This deficiency is generally attributed to inherited mutation(s) of the AGL gene, which results in partial or total abolishment of AGL enzyme activity in a subject afflicted with GSD III. Molecular analyses of the AGL protein in GSD III patients have been performed in several ethnic populations, and over 100 different AGL mutations have been described. *See Goldstein et al.*, 2010, *Genet. Med.* 12: 424-430. *See also Sentner et al.*, 2013, *JIMD Rep.* 7: 19-26.

[0009] There is presently no effective treatment of GSD III. Attempts have been made to control hypoglycemia with frequent meals high in carbohydrates, often via the use of nocturnal gastric drip feedings or cornstarch supplements. Meanwhile, patients with myopathy have been treated with diets high in protein during the daytime plus overnight enteral infusions. Transient improvement in symptoms has been documented in a few patients but there is no long-term data demonstrating that the high protein diet prevents or treats the progressive myopathy. *See Chen YT, Burchell A, Glycogen storage disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D, The metabolic and molecular basis of inherited disease. New York: McGraw Hill, 1995: 935-65. The progressive myopathy and/or cardiomyopathy is a major cause of morbidity in adults, and patients presenting with progressive liver cirrhosis and hepatic carcinoma have been reported. Thus, there is an urgent need for therapy which can address the underlying cause of this disease, i.e., the deficiency of AGL enzyme activity.*

[0010] To date, enzyme replacement has not been explored in diseases in which the defective enzyme is present in the cytosol, such as AGL in GSD III, presumably due to the lack of an efficient and specific cellular uptake mechanism that delivers exogenous enzyme across the plasma membrane into the cytoplasm.

[0011] The present invention addresses the above-mentioned needs by providing molecules, structures, and compositions that have the ability to be translated in the cytoplasm to provide active AGL, which can ameliorate, prevent or treat a disease or condition associated with AGL deficiency, such as GSD III.

SUMMARY OF THE INVENTION

[0012] This invention provides compositions comprising novel molecules having the ability to be translated, which can be used to provide one or more active polypeptides and proteins, or fragments thereof. The invention further provides methods of using these compositions comprising novel molecules for the prevention or treatment of various disorders. More specifically, embodiments of this invention provide compositions comprising translatable molecules to provide active amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) and methods of using the compositions for the treatment of GSD III.

[0013] The translatable molecules of this invention can have functional cytoplasmic activity for producing AGL polypeptides or proteins. The peptides and proteins may be active for therapeutic modalities.

[0014] The translatable molecules of this invention can have long half-life, particularly in the cytoplasm of a cell. The translatable molecules can be expressible to provide a product that is active for ameliorating, preventing or treating a disease or condition associated with an AGL deficiency.

[0015] This disclosure provides a range of structures for translatable molecules for producing AGL polypeptides or proteins. In some embodiments, the translatable molecules can have an increased ability to be translated and/or an extended half-life over a native mRNA.

[0016] The translatable molecules of this invention can be used in medicines, and for methods and compositions for producing and delivering active polypeptides and proteins. The translatable molecules of this invention can be used to provide polypeptides or proteins *in vitro*, *ex vivo*, or *in vivo*.

[0017] Embodiments of this disclosure provide a range of novel polynucleotides for expressing a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity. The polynucleotides can include natural nucleotides and chemically modified nucleotides. The polynucleotides can be expressible to provide a human AGL or a fragment thereof having AGL activity.

[0018] In further aspects, this invention provides a range of novel translatable oligomers comprising one or more unlocked nucleic acid (UNA) monomers for expressing a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity. A translatable oligomer can contain one or more UNA monomers, along with natural nucleotides and chemically modified nucleotides. A translatable oligomer comprising one or more UNA monomers can be expressible to provide the human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity.

[0019] In certain aspects, the translatable molecules of this invention can provide high-efficiency expression of a polypeptide or protein, or a fragment thereof. The expression can be *in vitro*, *ex vivo*, or *in vivo*.

[0020] In some embodiments, a molecule of this invention can have increased cytoplasmic half-life over a native, mature mRNA that encodes the same polypeptide or protein. The inventive molecules and compositions can provide increased functional cellular activity with respect to a native, mature mRNA.

[0021] In further aspects, a translatable molecule of this invention can provide increased activity as a drug agent providing a peptide or protein product, as compared to a native, mature mRNA. A translatable molecule of this invention may reduce the dose level required for efficacious therapy.

[0022] Embodiments of this invention include the following.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] FIG. 1 shows the results of expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL, NM_000028) *in vitro* using translatable molecules of this invention. FIG. 1 shows the relative expression of AGL in AML12 and C2C12 cells normalized to reference molecule 534. The vertical axis reflects the fold-increase relative to the reference, e.g., 10 being a 10-fold increase over the reference. The molecules, including the reference molecule, comprise a tobacco etch virus (TEV) 5' UTR and XenopusXenopus beta-globin (XBG) 3' UTR. The molecules were capped during transcription and synthesized with N¹-methylpseudouridine, so that 100% of uridines were replaced with N¹-methylpseudouridine. The translatable molecules encoding AGL were transfected in two cell lines (AML12, C2C12). Cells were lysed and harvested at 6 h post-transfection. Quantitative Western Blot was performed to detect AGL by using an antibody specific for AGL (ab133720, rabbit).

[0024] FIG. 2 shows the results of expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL, NM_000028) *in vitro* using translatable molecules of this invention. FIG. 2 shows the relative expression of AGL in AML12 and C2C12 cells normalized to reference molecule 534. The vertical axis reflects the fold-increase relative to the reference, e.g., 10 being a 10-fold increase over the reference. The molecules, including the reference molecule, comprise a tobacco etch virus (TEV) 5' UTR and Xenopus beta-globin (XBG) 3' UTR. The molecules were capped during transcription and synthesized with N¹-methylpseudouridine, so that 100% of uridines were replaced with N¹-methylpseudouridine. The translatable molecules encoding AGL were transfected in two cell lines (AML12, C2C12). Cells were lysed and harvested at 24 h post-transfection. Quantitative Western Blot was performed to detect AGL by using an antibody specific for AGL (ab133720, rabbit).

[0025] FIG. 3 shows the results of expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL, NM_000028) *in vivo* using translatable molecules of this invention. FIG. 3 shows the relative expression of AGL in WT mice for translatable molecules 528 and 534 (reference) at timepoints 24 h and 48 h. The molecules comprise a tobacco etch virus (TEV) 5' UTR and Xenopus beta-globin (XBG) 3' UTR. The molecules were capped during transcription and synthesized with N¹-methylpseudouridine, so that 100% of uridines were replaced with N¹-methylpseudouridine. The translatable molecules encoding AGL were each prepared in a lipid nanoparticle formulation and intravenously injected into WT mice at 10 mg/kg. Mice livers were harvested at 24 h and 48 h, and Quantitative Western Blot was performed to detect AGL by using an antibody specific for AGL (ab133720, rabbit).

[0026] FIG. 4 shows the results of expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL, NM_000028) *in vivo* using translatable molecules of this invention. FIG. 4 shows the relative liver expression of AGL in WT mice post-dose of translatable molecules 525, 527, 528, 529, and 546, as compared to baseline PBS (100). The synthesized translatable molecules 525, 527, 528, 529, and 546 encoding AGL were prepared in a lipid nanoparticle formulation and injected via IP in WT mice. Dose injected was 10 mpk, and livers were collected at 6 h for further analysis. Quantitative Western Blot was performed to detect AGL by using an antibody specific for AGL (ab133720, rabbit).

[0027] FIG. 5 shows the result of expressing human AGL from three translatable molecules: 546, 1783, and 1784. Human primary hepatocytes were transfected with codon-optimized mRNA and AGL protein expression was measured by In-Cell Western™ at 6, 24, 48, and 72

hours post-transfection. The expression of the mRNA sequences was compared with an untreated control (“unt”).

[0028] FIG. 6 shows expression of human AGL from various mRNA molecules formulated with lipid nanoparticles in wild-type C57BL/6 mice. The protein concentration (ng/mg) of exogenous human AGL expressed from the mRNA molecules in homogenates from liver biopsy samples was determined by a multiple reaction monitoring assay. The translatable molecules shown in the graph as 546.1, 736.1, 738.1, 737.1, 731.1, and 1783.1 are the same as 546, 736, 738, 737, 731, and 1783, respectively, as described in Example 2. Meanwhile, the translatable molecule 546.7 has the same nucleobase sequence as 546, which is described in Example 2, but was synthesized with 5-methoxyuridine in place of uridine instead of N¹-methylpseudouridine, which was used to synthesize translatable molecule 546.

[0029] FIG. 7 shows expression of endogenous mouse AGL in wild-type C57BL/6 mice treated with various mRNA molecules formulated with lipid nanoparticles. The protein concentration (ng/mg) of endogenous mouse AGL expressed in homogenates from liver biopsy samples was determined by a multiple reaction monitoring assay. The translatable molecules shown in the graph as 546.1, 736.1, 738.1, 737.1, 731.1, and 1783.1 are the same as 546, 736, 738, 737, 731, and 1783, respectively, as described in Example 2. Meanwhile, the translatable molecule 546.7 has the same nucleobase sequence as 546, which is described in Example 2, but was synthesized with 5-methoxyuridine in place of uridine instead of N¹-methylpseudouridine, which was used to synthesize translatable molecule 546.

[0030] FIG. 8 shows the histopathology of a liver from an AGL knockout mice treated with vehicle (“VEH”). Marked to severe vacuolation of hepatocytes and moderate to marked increases in glycogen accumulation was observed within hepatocytes treated with vehicle.

[0031] FIG. 9 shows the histopathology of a liver from an AGL knockout mice treated with translatable molecule 546 formulated with ATX2 lipid nanoparticles. Only mild to moderate vacuolation of hepatocytes and only mild to moderate increases in glycogen accumulation within hepatocytes was observed within hepatocytes treated with translatable molecule 546 formulated with ATX2 lipid nanoparticles.

DETAILED DESCRIPTION OF THE INVENTION

[0032] This invention provides a range of novel agents and compositions to be used for therapeutic applications. The molecules and compositions of this invention can be used for

ameliorating, preventing or treating GSD III and/or a disease associated reduced presence or function of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) in a subject.

[0033] In some embodiments, this invention encompasses synthetic, purified, translatable polynucleotide molecules for expressing a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase. The molecules may contain natural and chemically modified nucleotides, and encode the human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity.

[0034] In certain embodiments, this disclosure includes synthetic, purified, translatable oligomer molecules comprising one or more UNA monomers for expressing a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity. A translatable oligomer may contain one or more UNA monomers, as well as natural and chemically-modified nucleotides. A translatable oligomer comprising one or more UNA monomers can be expressible to provide the human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL), or a fragment thereof having AGL activity.

[0035] As used herein, the term “translatable” may be used interchangeably with the term “expressible” and refers to the ability of polynucleotide, or a portion thereof, to be converted to a polypeptide by a host cell. As is understood in the art, translation is the process in which ribosomes in a cell’s cytoplasm create polypeptides. In translation, messenger RNA (mRNA) is decoded by tRNAs in a ribosome complex to produce a specific amino acid chain, or polypeptide. Furthermore, the term “translatable” when used in this specification in reference to an oligomer, means that at least a portion of the oligomer, *e.g.*, the coding region of an oligomer sequence (also known as the coding sequence or CDS), is capable of being converted to a protein or a fragment thereof.

[0036] As used herein, the term “monomer” refers to a single unit, *e.g.*, a single nucleic acid, which may be joined with another molecule of the same or different type to form an oligomer. In some embodiments, a monomer may be an unlocked nucleic acid, *i.e.*, a UNA monomer.

[0037] Meanwhile, the term “oligomer” may be used interchangeably with “polynucleotide” and refers to a molecule comprising at least two monomers and includes oligonucleotides such as DNAs and RNAs. In the case of oligomers containing RNA monomers and/or unlocked nucleic acid (UNA) monomers, the oligomers of the present invention may contain sequences in addition to the coding sequence (CDS). These additional sequences may be untranslated sequences, *i.e.*, sequences which are not converted to protein by a host cell. These untranslated

sequences can include a 5' cap, a 5' untranslated region (5' UTR), a 3' untranslated region (3' UTR), and a tail region, *e.g.*, a polyA tail region. As described in further detail herein, any of these untranslated sequences may contain one or more UNA monomers – these UNA monomers are not capable of being translated by a host cell's machinery. In the context of the present invention, a “translatable oligomer”, a “translatable molecule”, “translatable polynucleotide”, or “translatable compound” refers to a sequence that comprises a region, *e.g.*, the coding region of an RNA (*e.g.*, the coding sequence of human AGL or a codon-optimized version thereof), that is capable of being converted to a protein or a fragment thereof, *e.g.*, the human AGL protein or a fragment thereof.

[0038] As used herein, the term “codon-optimized” means a natural (or purposefully designed variant of a natural) coding sequence which has been redesigned by choosing different codons without altering the encoded protein amino acid sequence increasing the protein expression levels (Gustafsson et al., *Codon bias and heterologous protein expression*. 2004, Trends Biotechnol 22: 346-53). Variables such as high codon adaptation index (CAI), LowU method, mRNA secondary structures, cis-regulatory sequences, GC content and many other similar variables have been shown to somewhat correlate with protein expression levels (Villalobos et al., *Gene Designer: a synthetic biology tool for constructing artificial DNA segments*. 2006, BMC Bioinformatics 7:285). High CAI (codon adaptation index) method picks a most frequently used synonymous codon for an entire protein coding sequence. The most frequently used codon for each amino acid is deduced from 74218 protein-coding genes from a human genome. The LowU method targets only U-containing codons that can be replaced with a synonymous codon with fewer U moieties. If there are a few choices for the replacement, the more frequently used codon will be selected. The remaining codons in the sequence are not changed by the LowU method. This method may be used in conjunction with the disclosed mRNAs to design coding sequences that are to be synthesized with 5-methoxyuridine.

[0039] As will be appreciated by the skilled artisan equipped with the present disclosure, the translatable molecules of the present invention may be used to ameliorate, prevent, or treat any disease or disorder associated with reduced activity (*e.g.*, resulting from reduced concentration, presence, and/or function) of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) in a subject. In some embodiments, the translatable molecules of this invention can be used in methods for ameliorating, preventing or treating one or more of GSD IIIa, GSD IIIb, GSDIIIc, and GSDIIId (collectively or individually referred to herein as “GSD III” or “glycogen storage

disease type III"). The disease or disorder to be treated herein (*e.g.*, GSD IIIa, GSD IIIb, GSDIIIc, or GSDIIId) may be associated with low blood sugar (hypoglycemia), enlargement of the liver (hepatomegaly), excessive amounts of fat in the blood (hyperlipidemia), elevated blood levels of liver enzymes, chronic liver disease (cirrhosis), liver failure, slow growth, short stature, benign tumors (adenomas), hypertrophic cardiomyopathy, cardiac dysfunction, congestive heart failure, skeletal myopathy, and/or poor muscle tone (hypotonia). In some embodiments, the translatable molecules of the present invention may be used to ameliorate, prevent, or treat any or all of these aforementioned symptoms.

[0040] As is understood by the skilled artisan, GSD III may be referred to by any number of alternative names in the art, including, but not limited to, AGL deficiency, Cori disease, Cori's disease, debrancher deficiency, Forbes disease, glycogen debrancher deficiency, GSD3, or limit dextrinosis (due to the limit dextrin-like structures in the cytosol). Accordingly, GSD III may be used interchangeably with any of these alternative names in the specification, the examples, the drawings, and the claims.

[0041] A translatable molecule of this invention encoding a functional AGL moiety can be delivered to the liver, in particular to hepatocytes, of a patient in need (*e.g.*, a GSD III patient), and can elevate active AGL levels of the patient. The translatable molecule can be used for preventing, treating, ameliorating or reversing any symptoms of GSD III in the patient.

[0042] In further aspects, a translatable molecule of this invention can also be used for reducing the dependence of a GSD III patient on a particular diet to control the disease. For instance, a translatable molecule of this invention can be used to reduce a GSD III patient's dependence on frequent high carbohydrate meals and/or diets abnormally high in protein.

[0043] Embodiments of this invention further encompass processes for making a translatable molecule for expressing a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL). The processes include transcribing *in vitro* an AGL DNA template in the presence of natural and chemically-modified nucleoside triphosphates to form a product mixture, and purifying the product mixture to isolate the translatable molecule. A translatable molecule may also be made by methods as are known in the art.

[0044] The molecules of this invention can be translatable molecules containing RNA and/or UNA monomers. These translatable molecules can have long half-life, particularly in the cytoplasm. The long duration translatable molecules can be used for ameliorating, preventing, or treating disease or disorder associated with reduced activity (*e.g.*, resulting from reduced

concentration, presence, and/or function) of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) in a subject.

[0045] The properties of the translatable molecules of this invention arise according to their molecular structure, and the structure of the molecule in its entirety, as a whole, can provide significant benefits based on those properties. Embodiments of this invention can provide translatable molecules having one or more properties that advantageously provide enhanced protein concentration or increased protein activity. The molecules and compositions of this invention can provide formulations comprising therapeutic agents for ameliorating, preventing, or treating any disease or disorder associated with reduced activity (*e.g.*, resulting from reduced concentration, presence, and/or function) of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) in a subject.

[0046] This invention provides a range of translatable molecules that are surprisingly translatable to provide active polypeptide or protein, *in vitro*, *ex vivo*, and *in vivo*.

[0047] A translatable molecule of this invention is expressible to provide one or more active polypeptides or proteins, or fragments thereof.

[0048] The translatable structures and compositions can have increased translational activity or cytoplasmic half-life. In these embodiments, the translatable structures and compositions can provide increased functional half-life in the cytoplasm of mammalian cells, as compared to a native mRNA.

[0049] As used herein, the term “half-life” is the time required for a quantity such as nucleic acid or protein concentration or activity to fall to half of its value as measured at the beginning of a time period.

[0050] A range of structures for translatable molecules of this invention are provided herein, including oligomers containing one or more UNA monomers. An oligomer containing one or more UNA monomers can incorporate specialized linker groups. The linker groups can be attached in a chain in the translatable molecule. Each linker group can also be attached to a nucleobase.

[0051] In some aspects, a linker group can be a monomer. Monomers can be attached to form a chain molecule. In a chain molecule of this invention, a linker group monomer can be attached at any point in the chain.

[0052] In certain aspects, linker group monomers can be attached in a chain molecule of this invention so that the linker group monomers reside near the ends of the chain, or at any position in the chain.

[0053] In further aspects, the linker groups of a chain molecule can each be attached to a nucleobase. The presence of nucleobases in the chain molecule can provide a sequence of nucleobases in the chain molecule.

[0054] In certain embodiments, this invention provides translatable oligomer molecules having chain structures that incorporate novel combinations of the linker group monomers, along with certain natural nucleotides, or non-natural nucleotides, or modified nucleotides, or chemically modified nucleotides.

[0055] The oligomer molecules of this invention can display a sequence of nucleobases, and can be designed to express a polypeptide or protein, *in vitro*, *ex vivo*, or *in vivo*. The expressed polypeptide or protein can have activity in various forms, including activity corresponding to a protein expressed from a natural, native or wild type mRNA, or activity corresponding to a negative or dominant negative protein.

[0056] In some aspects, this invention can provide active, translatable oligomer molecules having a base sequence that is identical to at least a fragment of a native nucleic acid molecule of a cell.

[0057] In some embodiments, the cell can be a eukaryotic cell, a mammalian cell, or a human cell.

[0058] This invention provides structures, methods and compositions for translatable oligomeric agents that incorporate the linker group monomers. The oligomeric molecules of this invention can be used as active agents in formulations for therapeutics.

[0059] This invention provides a range of translatable molecules that are useful for providing therapeutic effects because of their ability to be expressed as polypeptide or protein in a cell in a subject.

[0060] In certain embodiments, a translatable molecule can be structured as an oligomer composed of monomers. The oligomeric structures of this invention may contain one or more linker group monomers, along with certain nucleotides.

[0061] In certain embodiments, a translatable molecule may contain a sequence of nucleobases, and can be designed to express a peptide or protein of any isoform, in part by having sufficient homology with a native polynucleotide sequence.

[0062] In some embodiments, a translatable molecule can be from about 200 to about 12,000 monomers in length, or more. In certain embodiments, a translatable molecule can be from 1,000 to 9,000 monomers in length, from 3,000 to 7,000 monomers in length, or from 4,000 to 6,000 monomers in length. In an exemplary embodiment, the translatable molecule is from 4,500 to 5,500 monomers in length. In a further exemplary embodiment, the translatable molecule is about 5,000 monomers in length.

[0063] In some embodiments, a translatable molecule can contain from 1 to about 800 UNA monomers. In certain embodiments, a translatable molecule can contain from 1 to 600 UNA monomers, or 1 to 100 UNA monomers, or 1 to 12 UNA monomers.

[0064] In some embodiments, a translatable molecule can contain from 1 to about 800 locked nucleic acid (LNA) monomers. In certain embodiments, a translatable molecule can contain from 1 to 600 LNA monomers, or 1 to 100 LNA monomers, or 1 to 12 LNA monomers.

[0065] A translatable molecule of this invention may comprise a 5' cap, a 5' untranslated region of monomers, a coding region of monomers, a 3' untranslated region of monomers, and a tail region of monomers.

[0066] A translatable molecule of this invention may comprise a 3' untranslated region of monomers containing one or more UNA monomers.

[0067] A translatable molecule of this invention may comprise a tail region of monomers containing one or more UNA monomers.

[0068] A translatable molecule of this invention may comprise regions of sequences or structures that are operable for translation in a cell, or which have the functionality of regions of an mRNA including, for example, a 5' cap, a 5' untranslated region, a coding region, a 3' untranslated region, and a polyA tail.

[0069] This invention further contemplates methods for delivering one or more vectors comprising one or more translatable molecules to a cell. In further embodiments, the invention also contemplates delivering one or more translatable molecules to a cell.

[0070] In some embodiments, one or more translatable molecules can be delivered to a cell, *in vitro*, *ex vivo*, or *in vivo*. Viral and non-viral transfer methods as are known in the art can be

used to introduce translatable molecules in mammalian cells. Translatable molecules can be delivered with a pharmaceutically acceptable vehicle, or for example, with nanoparticles or liposomes.

[0071] In some embodiments, translatable structures and compositions of this invention can reduce the number and frequency of transfections required for cell-fate manipulation in culture as compared to utilizing native compositions.

[0072] In further aspects, this invention provides increased activity for translatable molecules as active agent, as compared to utilizing a native mRNA.

[0073] In some aspects, this invention can provide translatable molecules that may reduce the cellular innate immune response, as compared to that induced by a native nucleic acid, polypeptide or protein.

[0074] This invention can provide synthetic translatable molecules that are refractory to deadenylation as compared to native molecules.

[0075] In certain embodiments, this invention can provide synthetic translatable molecules with increased specific activity and longer functional half-life as compared to native molecules. The synthetic translatable molecules of this invention can provide increased levels of ectopic protein expression. When expressing a translatable molecule using a vector, cellular-delivery can be at increased levels, and cytotoxic innate immune responses can be restrained so that higher levels of ectopic protein expression can be achieved. The translatable molecules of this invention can have increased specific activity and longer functional half-life than native mRNAs.

[0076] In certain aspects, a translatable molecule may have a number of mutations relative to a native mRNA.

[0077] In further embodiments, this invention can provide translatable molecules having cleavable delivery and targeting moieties attached at a 3' end and/or a 5' end.

[0078] In general, the specific activity for a synthetic translatable molecule delivered by transfection can be viewed as the number of molecules of protein expressed per delivered transcript per unit time.

[0079] As used herein, translation efficiency refers to a measure of the production of a protein or polypeptide by translation of a translatable molecule *in vitro* or *in vivo*.

[0080] This invention provides a range of translatable oligomer molecules, which can contain one or more UNA monomers, and a number of nucleic acid monomers, wherein the translatable molecule can be expressible to provide a polypeptide or protein.

[0081] In some embodiments, this invention includes a range of translatable oligomer molecules, which can contain one or more UNA monomers in one or more untranslated regions, and a number of nucleic acid monomers, wherein the translatable molecule can be expressible to provide a polypeptide or protein.

[0082] In some embodiments, this invention includes a range of translatable molecules, which contain one or more UNA monomers in a tail region, and a number of nucleic acid monomers, wherein the translatable molecule can be expressible to provide a polypeptide or protein.

[0083] In some embodiments, a translatable molecule can contain a modified 5' cap.

[0084] In further embodiments, a translatable molecule can contain a translation enhancing 5' untranslated region of monomers.

[0085] In additional embodiments, a translatable molecule can contain a translation enhancing 3' untranslated region of monomers.

[0086] In additional embodiments, a translatable molecule can contain one or more UNA monomers in a 3' untranslated region of monomers.

[0087] In further embodiments, a translatable molecule can contain one or more UNA monomers in a tail region of monomers.

[0088] In further embodiments, a translatable molecule can contain one or more UNA monomers in a polyA tail.

[0089] In some embodiments, a translatable molecule can contain one or more LNA monomers in a 3' untranslated region of monomers or in a tail region of monomers, *e.g.*, in a polyA tail.

[0090] In another aspect, a translatable molecule of this invention can exhibit at least 2-fold, 3-fold, 5-fold, or 10-fold increased translation efficiency *in vivo* as compared to a native mRNA that encodes the same translation product.

[0091] In a further aspect, a translatable molecule can produce at least a 2-fold, 3-fold, 5-fold, or 10-fold increased polypeptide or protein level *in vivo* as compared to a native mRNA that encodes the same polypeptide or protein.

[0092] In certain embodiments, a translatable molecule can provide increased levels of a polypeptide or protein *in vivo* as compared to a native mRNA that encodes the same polypeptide or protein. For example, the level of a polypeptide or protein can be increased by 10%, or 20%, or 30%, or 40%, or 50%, or more.

[0093] In additional embodiments, this invention provides methods for treating a disease or condition in a subject by administering to the subject a composition containing a translatable molecule of the invention.

[0094] A translatable molecule of this invention may be used for ameliorating, preventing or treating a disease or disorder, *e.g.*, a disease or disorder associated with reduced activity (*e.g.*, resulting from reduced concentration, presence, and/or function) of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) in a subject. In these embodiments, a composition comprising a translatable molecule of this invention can be administered to regulate, modulate, or increase the concentration or effectiveness of the AGL enzyme in a subject. In some aspects, the enzyme can be an unmodified, natural enzyme for which the patient has an abnormal quantity (*e.g.*, a patient with a mutated version of AGL which partially or totally abolishes AGL activity). In some aspects, the enzyme can be an unmodified, natural AGL enzyme which can be used to treat a patient harboring a mutated version of AGL. In exemplary embodiments, a translatable molecule of this invention may be used for ameliorating, preventing or treating GSD III.

[0095] In some embodiments, a translatable molecule may be delivered to cells or subjects, and translated to increase AGL levels in the cell or subject.

[0096] 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.

[0097] In an exemplary embodiment, a subject of the present invention is a subject with reduced activity (*e.g.*, resulting from reduced concentration, presence, and/or function) of amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL). In a further exemplary embodiment, the subject is a human.

[0098] In some embodiments, administering a composition comprising a translatable molecule of the invention can result in increased liver AGL protein levels in a treated subject. In some embodiments, administering a composition comprising a translatable molecule of the invention results in about a 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95% increase in liver AGL protein levels relative to a baseline AGL protein level in the subject prior to treatment. In an exemplary embodiment, administering a composition comprising a translatable molecule of the invention results in an increase in liver AGL levels relative to baseline liver AGL levels in the subject prior to treatment. In some embodiments, the increase in liver AGL levels can be at least about 5%, 10%, 20%, 30%, 40%, 50%, 100%, 200%, or more.

[0099] In some embodiments, the AGL protein which is expressed from a translatable molecule of the invention is detectable in the liver, serum, plasma, kidney, heart, muscle, brain, cerebrospinal fluid, or lymph nodes. In exemplary embodiments, the AGL protein is expressed in the liver cells, *e.g.*, hepatocytes of a treated subject.

[00100] In some embodiments, administering a composition comprising a translatable molecule of the invention results in the expression of a natural, non-mutated human AGL (*i.e.*, normal or wild-type AGL as opposed to abnormal or mutated AGL) protein level at or above about 10 ng/mg, about 20 ng/mg, about 50 ng/mg, about 100 ng/mg, about 150 ng/mg, about 200 ng/mg, about 250 ng/mg, about 300 ng/mg, about 350 ng/mg, about 400 ng/mg, about 450 ng/mg, about 500 ng/mg, about 600 ng/mg, about 700 ng/mg, about 800 ng/mg, about 900 ng/mg, about 1000 ng/mg, about 1200 ng/mg or about 1500 ng/mg of the total protein in the liver of a treated subject.

[00101] As used herein, the term “about” or “approximately” 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 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).

[00102] In some embodiments, the expression of the natural, non-mutated human AGL protein is detectable 6, 12, 18, 24, 30, 36, 48, 60, and/or 72 hours after administration of a composition comprising a translatable molecule of the invention. In some embodiments, the expression of the natural, non-mutated human AGL protein is detectable 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, and/or 7 days after administration of a composition comprising a translatable molecule of the invention. In some embodiments, the expression of the natural, non-mutated human AGL protein is detectable 1 week, 2 weeks, 3 weeks, and/or 4 weeks after the administration. In some embodiments, the expression of the natural, non-mutated human AGL protein is detectable after administration of a composition comprising a translatable molecule of the invention. In some embodiments, expression of natural, non-mutated human AGL protein is detectable in the liver, *e.g.*, hepatocytes, after administration of a composition comprising a translatable molecule of the invention.

Variant Templates for Making Translatable Molecules

[00103] In various embodiments described herein, the translatable oligomer may comprise a mRNA encoding AGL, wherein the mRNA encoding AGL is codon-optimized. In some embodiments, the AGL is human AGL (*i.e.*, hAGL). In some embodiments, the human AGL comprises an amino acid sequence of SEQ ID NO: 2. In some embodiments, the human AGL consists of an amino acid sequence of SEQ ID NO: 2.

[00104] In some embodiments, a variant DNA template may be utilized to make a translatable molecule capable of encoding AGL. A variant DNA template of this disclosure may exhibit advantages in processes for making a translatable molecule, and the efficiency of the translatable molecule. Variation of the template can be utilized to enhance incorporation of modified nucleotides or monomers in a translatable molecule of this invention. In certain aspects, variation of the template can be utilized to enhance the structural features of the translatable molecule. The enhanced structural features of the translatable molecule can provide unexpectedly advantageous properties, including translation efficiency to provide a polypeptide or protein product.

[00105] In some aspects of this invention, variation of the template may include reducing the occurrence or frequency of appearance of certain nucleotides in the template strand. Reducing the occurrence of a certain nucleotide can alter the structures and processes of this

disclosure to provide non-native forms, which achieve surprisingly improved properties of a translatable RNA product encoding AGL.

[00106] Aspects of this invention may require a variant DNA template in processes for making a translatable molecule. A DNA molecule can have a non-coding template strand of nucleotides that can be transcribed to provide a target translatable molecule encoding AGL.

[00107] A target translatable molecule can be any RNA, whether native or modified, synthetic or derived from a natural source.

[00108] In some embodiments, a variant DNA template can be used for which an open reading frame of the template strand is transformed to an alternative form, while preserving codon assignment.

[00109] In certain embodiments, a DNA template can be used for which alternative nucleotides are used based on alternative codon optimization and/or sequence degeneracy.

[00110] In additional embodiments, a DNA template may have certain nucleotides replaced with alternative nucleotides, while preserving codon assignment.

[00111] Embodiments of this invention advantageously utilize alternative codons in a DNA template of this invention to be used in processes for making a translatable molecule encoding AGL. The variations that can be achieved in a DNA template of this invention can be far greater in scope than for cells and organisms, which may require preferred codons in many processes. In this invention, a wide range of alternative codons and positions can be used in a DNA template for transcribing a translatable molecule.

[00112] In further aspects of this invention, variation of the template may include reducing the occurrence or frequency of appearance of certain nucleotides in the template strand. For example, the occurrence of a nucleotide in a template may be reduced to a level below 25% of nucleotides in the template. In further examples, the occurrence of a nucleotide in a template may be reduced to a level below 20% of nucleotides in the template. In some examples, the occurrence of a nucleotide in a template may be reduced to a level below 16% of nucleotides in the template. In certain examples, the occurrence of a nucleotide in a template may be reduced to a level below 12% of nucleotides in the template.

Human AGL

[00113] The human AGL gene encodes a 1532 amino acid protein with a molecular mass of approximately 174.8 kDa. AGL is a multifunctional enzyme acting as a 1,4-alpha-D-glucan:1,4-alpha-D-glucan-4-alpha-D-glycosyltransferase and an amylo-1,6-glucosidase in glycogen degradation. As noted above, genetic deficiency of normal AGL activity causes glycogen storage disease III.

[00114] The consensus human AGL coding sequence has an RNA sequence of 4,599 nucleobases, shown in SEQ ID NO: 1.

[00115] The consensus human AGL coding sequence – found at NCBI Accession No. NP_000019.2 – translates into SEQ ID NO: 2.

[00116] In some embodiments, a translatable molecule can be made and used for expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (hAGL) with advantageously increased efficiency of translation, as compared to a native mRNA of hAGL. The translatable molecule expressing hAGL may exhibit activity suitable for use in methods for ameliorating, preventing or treating disease. In some embodiments, the translatable molecule may comprise one or more UNA monomers.

[00117] In some embodiments, a translatable molecule may include a 5' cap, a 5' UTR, a translation initiation sequence, *e.g.*, a Kozak sequence, a human AGL CDS, a 3' UTR, and/or a tail region. In an exemplary embodiment, a translatable molecule may include a 5' cap (m7GpppGm), a 5' UTR of tobacco etch virus (TEV), a Kozak sequence, a human AGL CDS, a 3' UTR of *Xenopus* beta-globin, and a tail region. In further exemplary embodiments, the human AGL CDS may comprise a codon-optimized sequence of SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45, described in further detail below. In any of these and other embodiments described herein, the translatable molecule may comprise one or more UNA monomers. In any of these and other embodiments described herein, the translatable molecule may comprise one or more LNA monomers.

[00118] The translation efficiency of the molecule can be increased as compared to a native mRNA of AGL. In particular, after 48 hours, the translation efficiency of the molecule may be more than doubled as compared to the native mRNA of AGL.

[00119] In some embodiments, a suitable mRNA sequence for the present invention comprises an mRNA sequence encoding the human AGL protein. The sequence of the naturally occurring human AGL protein is shown in SEQ ID NO: 2.

[00120] In some embodiments, a suitable mRNA sequence may be an mRNA sequence that encodes a homolog or variant of human AGL. As used herein, a homolog or a variant of human AGL protein may be a modified human AGL protein containing one or more amino acid substitutions, deletions, and/or insertions as compared to a wild-type or naturally-occurring human AGL protein while retaining substantial AGL protein activity. In some embodiments, an mRNA suitable for the present invention encodes a protein substantially identical to human AGL protein. In some embodiments, an mRNA suitable for the present invention encodes an amino acid sequence at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO: 2. In some embodiments, an mRNA suitable for the present invention encodes a fragment or a portion of human AGL protein.

[00121] In some embodiments, an mRNA suitable for the present invention encodes a fragment or a portion of human AGL protein, wherein the fragment or portion of the protein still maintains AGL activity similar to that of the wild-type protein.

[00122] In some embodiments, an mRNA suitable for the present invention comprises a sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45.

[00123] In some embodiments, a translatable oligomeric molecule of the present invention comprises a coding sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45. In some embodiments, a translatable oligomeric molecule comprising a coding sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45 further comprises one or more sequences selected from a 5' cap, a 5' UTR, a translation initiation sequence, a 3' UTR, and a tail region.

[00124] In some embodiments, a translatable oligomeric molecule of the present invention comprises a coding sequence that is less than 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and expresses a functional human AGL protein. In an exemplary embodiment, a translatable oligomeric molecule of the present invention comprises a coding sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and expresses a functional human AGL protein. In another exemplary embodiment, a

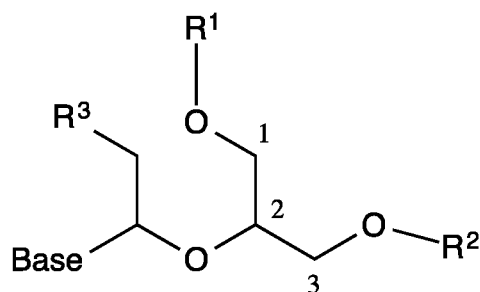
translatable oligomeric molecule of the present invention comprises a coding sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and expresses a functional human AGL protein, wherein the coding sequence is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to a sequence selected from SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45. In yet another exemplary embodiment, a translatable oligomeric molecule of the present invention comprises a coding sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and expresses a functional human AGL protein, wherein the coding sequence is at least 95% identical to a sequence selected from SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45. In yet another exemplary embodiment, a translatable oligomeric molecule of the present invention comprises a coding sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and expresses a functional human AGL protein, wherein the coding sequence is at least 95% identical to a sequence selected from SEQ ID NO: 19, SEQ ID NO: 31, or SEQ ID NO: 45. Accordingly, in some embodiments, the present application provides a polynucleotide comprising of or consisting of a nucleobase sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and wherein the human AGL coding sequence is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to a sequence selected from SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45. In an exemplary embodiment, the present application provides a polynucleotide comprising of or consisting of a nucleobase sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and wherein the human AGL coding sequence is at least 95% identical to a sequence selected from SEQ ID NO: 19, SEQ ID NO: 31, or SEQ ID NO: 45. In a specific embodiment, the present application provides a polynucleotide comprising of a nucleobase sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and wherein the human AGL coding sequence is at least 95% identical to SEQ ID NO: 19. In another specific embodiment, the present application provides a polynucleotide comprising of a nucleobase sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and wherein the human AGL coding sequence is at least 95% identical to SEQ ID NO: 31. In another specific embodiment, the present application provides a polynucleotide comprising of a nucleobase

sequence that is less than 80% identical to the wild-type human AGL coding sequence over the full length human AGL coding sequence of SEQ ID NO: 1, and wherein the human AGL coding sequence is at least 95% identical to SEQ ID NO: 45.

[00125] In some embodiments, a translatable oligomeric molecule of the invention encodes a fusion protein comprising a full length, fragment or portion of a AGL protein fused to another sequence (*e.g.*, an N or C terminal fusion). In some embodiments, the N or C terminal sequence is a signal sequence or a cellular targeting sequence.

UNA Monomers and Oligomers

[00126] In some embodiments, linker group monomers can be unlocked nucleomonomers (UNA monomers), which are small organic molecules based on a propane-1,2,3-tri-yl-trisoxy structure as shown below:

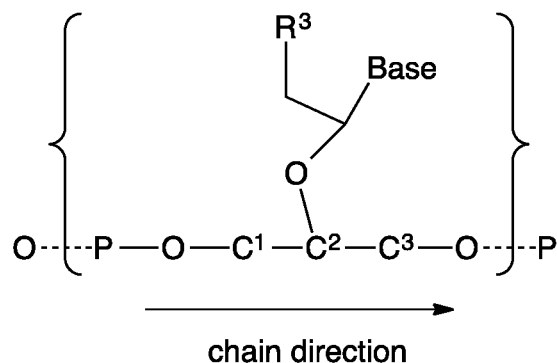


UNA MONOMER

where R^1 and R^2 are H, and R^1 and R^2 can be phosphodiester linkages, Base can be a nucleobase, and R^3 is a functional group described below.

[00127] In another view, the UNA monomer main atoms can be drawn in IUPAC notation as follows:

UNA monomer unit



where the direction of progress of the oligomer chain is from the 1-end to the 3-end of the propane residue.

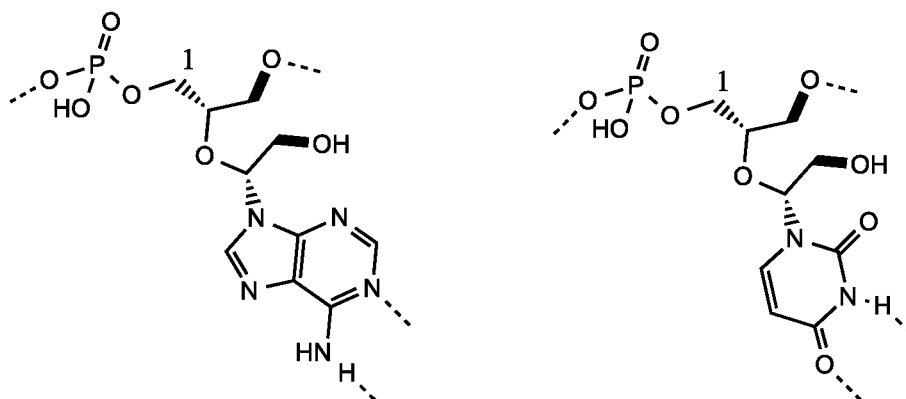
[00128] Examples of a nucleobase include uracil, thymine, cytosine, 5-methylcytosine, adenine, guanine, inosine, and natural and non-natural nucleobase analogues.

[00129] Examples of a nucleobase include pseudouracil, 1-methylpseudouracil (m1Ψ), *i.e.*, N¹-methylpseudouracil, and 5-methoxyuracil.

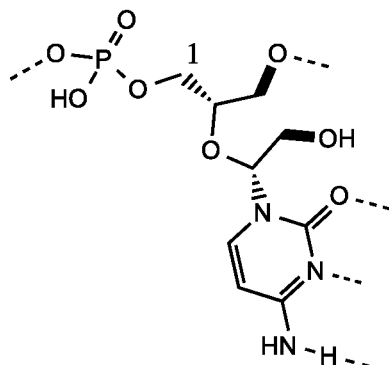
[00130] In general, a UNA monomer, which is not a nucleotide, can be an internal linker monomer in an oligomer. An internal UNA monomer in an oligomer is flanked by other monomers on both sides.

[00131] A UNA monomer can participate in base pairing when the oligomer forms a complex or duplex, for example, and there are other monomers with nucleobases in the complex or duplex.

[00132] Examples of UNA monomer as internal monomers flanked at both the propane-1-yl position and the propane-3-yl position, where R³ is -OH, are shown below.

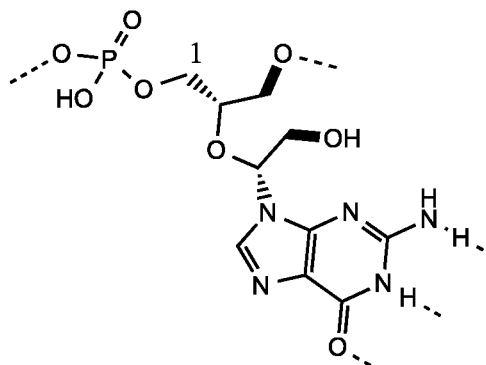


UNA-A



UNA-C

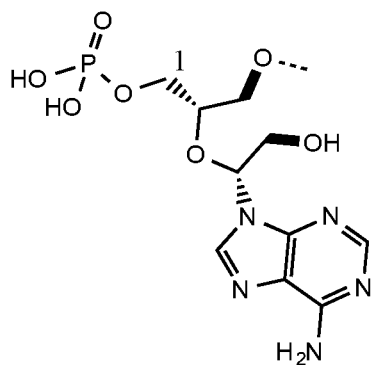
UNA-U



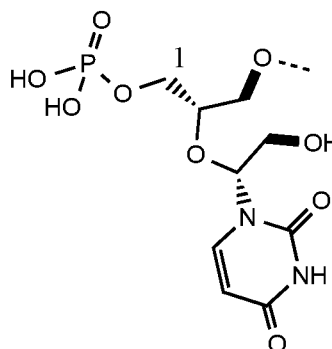
UNA-G

[00133] A UNA monomer can be a terminal monomer of an oligomer, where the UNA monomer is attached to only one monomer at either the propane-1-yl position or the propane-3-yl position. Because the UNA monomers are flexible organic structures, unlike nucleotides, the terminal UNA monomer can be a flexible terminator for the oligomer.

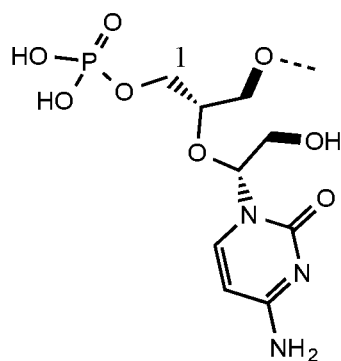
[00134] Examples of a UNA monomer as a terminal monomer attached at the propane-3-yl position are shown below.



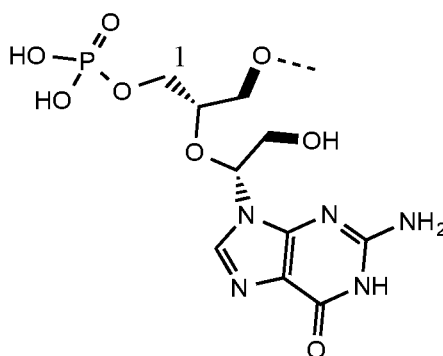
terminal UNA-A



terminal UNA-U

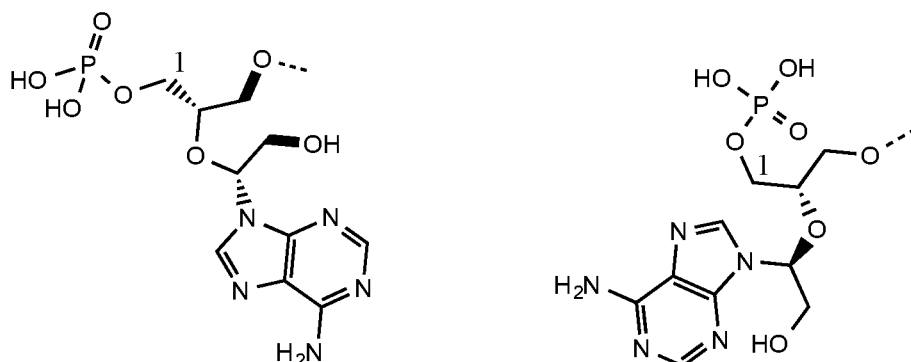


terminal UNA-C



terminal UNA-G

[00135] Because a UNA monomer can be a flexible molecule, a UNA monomer as a terminal monomer can assume widely differing conformations. An example of an energy minimized UNA monomer conformation as a terminal monomer attached at the propane-3-yl position is shown below.



UNA-A terminal forms: the dashed bond shows the propane-3-yl attachment

[00136] Among other things, the structure of the UNA monomer allows it to be attached to naturally-occurring nucleotides.

[00137] A UNA oligomer can be a chain composed of UNA monomers, as well as various nucleotides that may be based on naturally-occurring nucleosides.

[00138] In some embodiments, the functional group R^3 of a UNA monomer can be OR^4 , SR^4 , NR^4_2 , $NH(C=O)R^4$, morpholino, morpholin-1-yl, piperazin-1-yl, or 4-alkanoyl-piperazin-1-yl, where R^4 is the same or different for each occurrence, and can be H, alkyl, a cholesterol, a lipid molecule, a polyamine, an amino acid, or a polypeptide.

[00139] The UNA monomers are organic molecules. UNA monomers are not nucleic acid monomers or nucleotides, nor are they naturally-occurring nucleosides or modified naturally-occurring nucleosides.

[00140] A UNA oligomer of this invention is a synthetic chain molecule.

[00141] In some embodiments, as shown above, a UNA monomer can be UNA-A (designated $\tilde{A}$), UNA-U (designated $\tilde{U}$), UNA-C (designated $\tilde{C}$), and UNA-G (designated $\tilde{G}$).

[00142] Designations that may be used herein include mA, mG, mC, and mU, which refer to the 2'-O-Methyl modified ribonucleotides.

[00143] Designations that may be used herein include dT, which refers to a 2'-deoxy T nucleotide.

[00144] As used herein, in the context of oligomer sequences, the symbol N can represent any natural nucleotide monomer, or any modified nucleotide monomer.

[00145] As used herein, in the context of oligomer sequences, the symbol Q represents a non-natural, modified, or chemically-modified nucleotide monomer.

[00146] As used herein, in the context of oligomer sequences, the symbol X may be used to represent a UNA monomer.

Modified and Chemically-Modified Nucleotides

[00147] In the examples of modified or chemically-modified nucleotides herein, an alkyl, cycloalkyl, or phenyl substituent may be unsubstituted, or further substituted with one or more alkyl, halo, haloalkyl, amino, or nitro substituents.

[00148] Examples of nucleic acid monomers include non-natural, modified, and chemically-modified nucleotides, including any such nucleotides known in the art.

[00149] Examples of modified or chemically-modified nucleotides include 5-hydroxycytidines, 5-alkylcytidines, 5-hydroxyalkylcytidines, 5-carboxycytidines, 5-formylcytidines, 5-alkoxycytidines, 5-alkynylcytidines, 5-halocytidines, 2-thiocytidines, N⁴-alkylcytidines, N⁴-aminocytidines, N⁴-acetylcytidines, and N⁴,N⁴-dialkylcytidines.

[00150] Examples of modified or chemically-modified nucleotides include 5-hydroxycytidine, 5-methylcytidine, 5-hydroxymethylcytidine, 5-carboxycytidine, 5-formylcytidine, 5-methoxycytidine, 5-propynylcytidine, 5-bromocytidine, 5-iodocytidine, 2-

thiocyridine; N⁴-methylcyridine, N⁴-aminocyridine, N⁴-acetylcyridine, and N⁴,N⁴-dimethylcyridine.

[00151] Examples of modified or chemically-modified nucleotides include 5-hydroxyuridines, 5-alkyluridines, 5-hydroxyalkyluridines, 5-carboxyuridines, 5-carboxyalkylesteruridines, 5-formyluridines, 5-alkoxyuridines, 5-alkynyluridines, 5-halouridines, 2-thiouridines, and 6-alkyluridines.

[00152] Examples of modified or chemically-modified nucleotides include 5-hydroxyuridine, 5-methyluridine, 5-hydroxymethyluridine, 5-carboxyuridine, 5-carboxymethylesteruridine, 5-formyluridine, 5-methoxyuridine, 5-propynyluridine, 5-bromouridine, 5-fluorouridine, 5-iodouridine, 2-thiouridine, and 6-methyluridine.

[00153] Examples of modified or chemically-modified nucleotides include 5-methoxycarbonylmethyl-2-thiouridine, 5-methylaminomethyl-2-thiouridine, 5-carbamoylmethyluridine, 5-carbamoylmethyl-2'-O-methyluridine, 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine, 5-methylaminomethyl-2-selenouridine, 5-carboxymethyluridine, 5-methyldihydrouridine, 5-taurinomethyluridine, 5-taurinomethyl-2-thiouridine, 5-(isopentenylaminomethyl)uridine, 2'-O-methylpseudouridine, 2-thio-2'-O-methyluridine, and 3,2'-O-dimethyluridine.

[00154] Examples of modified or chemically-modified nucleotides include N⁶-methyladenosine, 2-aminoadenosine, 3-methyladenosine, 8-azaadenosine, 7-deazaadenosine, 8-oxoadenosine, 8-bromoadenosine, 2-methylthio-N⁶-methyladenosine, N⁶-isopentenyladenosine, 2-methylthio-N⁶-isopentenyladenosine, N⁶-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N⁶-(cis-hydroxyisopentenyl)adenosine, N⁶-glycylcarbamoyladenosine, N⁶-threonylcarbamoyl-adenosine, N⁶-methyl-N⁶-threonylcarbamoyl-adenosine, 2-methylthio-N⁶-threonylcarbamoyl-adenosine, N⁶,N⁶-dimethyladenosine, N⁶-hydroxynorvalylcarbamoyladenosine, 2-methylthio-N⁶-hydroxynorvalylcarbamoyl-adenosine, N⁶-acetyl-adenosine, 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, alpha-thio-adenosine, 2'-O-methyl-adenosine, N⁶,2'-O-dimethyl-adenosine, N⁶,N⁶,2'-O-trimethyl-adenosine, 1,2'-O-dimethyl-adenosine, 2'-O-ribosyladenosine, 2-amino-N⁶-methyl-purine, 1-thio-adenosine, 2'-F-ara-adenosine, 2'-F-adenosine, 2'-OH-ara-adenosine, and N⁶-(19-amino-pentaoxonadecyl)-adenosine.

[00155] Examples of modified or chemically-modified nucleotides include N¹-alkylguanosines, N²-alkylguanosines, thienoguanosines, 7-deazaguanosines, 8-oxoguanosines, 8-bromoguanosines, O⁶-alkylguanosines, xanthosines, inosines, and N¹-alkylinosines.

[00156] Examples of modified or chemically-modified nucleotides include N¹-methylguanosine, N²-methylguanosine, thienoguanosine, 7-deazaguanosine, 8-oxoguanosine, 8-bromoguanosine, O⁶-methylguanosine, xanthosine, inosine, and N¹-methylinosine.

[00157] Examples of modified or chemically-modified nucleotides include pseudouridines. Examples of pseudouridines include N¹-alkylpseudouridines, N¹-cycloalkylpseudouridines, N¹-hydroxypseudouridines, N¹-hydroxyalkylpseudouridines, N¹-phenylpseudouridines, N¹-phenylalkylpseudouridines, N¹-aminoalkylpseudouridines, N³-alkylpseudouridines, N⁶-alkylpseudouridines, N⁶-alkoxypseudouridines, N⁶-hydroxypseudouridines, N⁶-hydroxyalkylpseudouridines, N⁶-morpholinopseudouridines, N⁶-phenylpseudouridines, and N⁶-halopseudouridines. Examples of pseudouridines include N¹-alkyl-N⁶-alkylpseudouridines, N¹-alkyl-N⁶-alkoxypseudouridines, N¹-alkyl-N⁶-hydroxypseudouridines, N¹-alkyl-N⁶-hydroxyalkylpseudouridines, N¹-alkyl-N⁶-morpholinopseudouridines, N¹-alkyl-N⁶-phenylpseudouridines, and N¹-alkyl-N⁶-halopseudouridines. In these examples, the alkyl, cycloalkyl, and phenyl substituents may be unsubstituted, or further substituted with alkyl, halo, haloalkyl, amino, or nitro substituents.

[00158] Examples of pseudouridines include N¹-methylpseudouridine, N¹-ethylpseudouridine, N¹-propylpseudouridine, N¹-cyclopropylpseudouridine, N¹-phenylpseudouridine, N¹-aminomethylpseudouridine, N³-methylpseudouridine, N¹-hydroxypseudouridine, and N¹-hydroxymethylpseudouridine.

[00159] Examples of nucleic acid monomers include modified and chemically-modified nucleotides, including any such nucleotides known in the art.

[00160] Examples of modified and chemically-modified nucleotide monomers include any such nucleotides known in the art, for example, 2'-O-methyl ribonucleotides, 2'-O-methyl purine nucleotides, 2'-deoxy-2'-fluoro ribonucleotides, 2'-deoxy-2'-fluoro pyrimidine nucleotides, 2'-deoxy ribonucleotides, 2'-deoxy purine nucleotides, universal base nucleotides, 5-C-methyl-nucleotides, and inverted deoxyabasic monomer residues.

[00161] Examples of modified and chemically-modified nucleotide monomers include 3'-end stabilized nucleotides, 3'-glyceryl nucleotides, 3'-inverted abasic nucleotides, and 3'-inverted thymidine.

[00162] Examples of modified and chemically-modified nucleotide monomers include locked nucleic acid nucleotides (LNA), 2'-O,4'-C-methylene-(D-ribofuranosyl) nucleotides, 2'-methoxyethoxy (MOE) nucleotides, 2'-methyl-thio-ethyl, 2'-deoxy-2'-fluoro nucleotides, and 2'-O-methyl nucleotides. In an exemplary embodiment, the modified monomer is a locked nucleic acid nucleotide (LNA).

[00163] Examples of modified and chemically-modified nucleotide monomers include 2',4'-constrained 2'-O-methoxyethyl (cMOE) and 2'-O-Ethyl (cEt) modified DNAs.

[00164] Examples of modified and chemically-modified nucleotide monomers include 2'-amino nucleotides, 2'-O-amino nucleotides, 2'-C-allyl nucleotides, and 2'-O-allyl nucleotides.

[00165] Examples of modified and chemically-modified nucleotide monomers include N⁶-methyladenosine nucleotides.

[00166] Examples of modified and chemically-modified nucleotide monomers include nucleotide monomers with modified bases 5-(3-amino)propyluridine, 5-(2-mercapto)ethyluridine, 5-bromouridine; 8-bromoguanosine, or 7-deazaadenosine.

[00167] Examples of modified and chemically-modified nucleotide monomers include 2'-O-aminopropyl substituted nucleotides.

[00168] Examples of modified and chemically-modified nucleotide monomers include replacing the 2'-OH group of a nucleotide with a 2'-R, a 2'-OR, a 2'-halogen, a 2'-SR, or a 2'-amino, where R can be H, alkyl, alkenyl, or alkynyl.

[00169] Some examples of modified nucleotides are given in Saenger, Principles of Nucleic Acid Structure, Springer-Verlag, 1984.

[00170] Example of base modifications described above can be combined with additional modifications of nucleoside or nucleotide structure, including sugar modifications and linkage modifications.

[00171] Certain modified or chemically-modified nucleotide monomers may be found in nature.

Translatable Molecules Containing One or More UNA Monomers

[00172] Aspects of this invention provide structures and compositions for translatable molecules that are oligomeric compounds containing one or more UNA monomers. The translatable oligomers can be active agents for pharmaceutical compositions. In some embodiments, the translatable oligomers encode human AGL or a variant thereof.

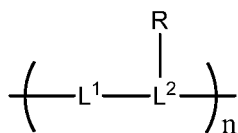
[00173] An oligomeric, translatable molecule of this invention may contain one or more UNA monomers. Oligomeric molecules of this invention can be used as active agents in formulations for supplying peptide and protein therapeutics. In some embodiments, the translatable oligomers encode human AGL or a variant thereof.

[00174] In some embodiments, this invention provides oligomeric, translatable compounds having a structure that incorporates novel combinations of UNA monomers with certain natural nucleotides, non-natural nucleotides, modified nucleotides, or chemically-modified nucleotides.

[00175] Translatable oligomeric compounds of this invention can have a length of from about 200 to about 12,000 bases in length. Translatable oligomeric compounds of this invention can have a length of about 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, or about 9000 bases. In some embodiments, translatable oligomeric compounds of this invention can have a length of about 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000, 5100, 5200, 5300, 5400, or about 5500 bases. In an exemplary embodiment, the translatable oligomeric compound of the invention has a length of about 5000 bases.

[00176] In further aspects, the oligomeric, translatable compounds of this invention which comprise one or more UNA monomers can be pharmacologically active molecules. A translatable oligomeric molecule can be used as an active pharmaceutical ingredient for generating a peptide or protein active agent *in vitro*, *in vivo*, or *ex vivo*. In an exemplary embodiment, the translatable oligomeric compound of this invention encodes human AGL or a variant thereof.

[00177] A translatable oligomeric molecule of this invention can have the structure of Formula I:



Formula I

wherein L^1 is a linkage, n is from 200 to 12,000, and for each occurrence L^2 is a UNA linker group having the formula $\text{---C}^1\text{---C}^2\text{---C}^3\text{---}$, where R is attached to C^2 and has the formula $\text{---OCH}(\text{CH}_2\text{R}^3)\text{R}^5$, where R^3 is ---OR^4 , ---SR^4 , ---NR^4_2 , $\text{---NH}(\text{C=O})\text{R}^4$, morpholino, morpholin-1-yl, piperazin-1-yl, or 4-alkanoyl-piperazin-1-yl, where R^4 is the same or different for each occurrence and is H, alkyl, a cholesterol, a lipid molecule, a polyamine, an amino acid, or a polypeptide, and where R^5 is a nucleobase, or $\text{L}^2(\text{R})$ is a sugar such as a ribose and R is a nucleobase, or L^2 is a modified sugar such as a modified ribose and R is a nucleobase. In certain embodiments, a nucleobase can be a modified nucleobase. L^1 can be a phosphodiester linkage.

[00178] The base sequence of a translatable oligomeric molecule can be any sequence of nucleobases.

[00179] In some aspects, a translatable oligomeric molecule of this invention can have any number of phosphorothioate intermonomer linkages in any intermonomer location.

[00180] In some embodiments, any one or more of the intermonomer linkages of a translatable oligomeric molecule can be a phosphodiester, a phosphorothioate including dithioates, a chiral phosphorothioate, and other chemically modified forms.

[00181] When a translatable oligomeric molecule terminates in a UNA monomer, the terminal position has a 1-end, or the terminal position has a 3-end, according to the positional numbering shown above.

Enhanced Translation

[00182] A translatable molecule of this invention can incorporate a region that enhances the translational efficiency of the molecule.

[00183] In general, translational enhancer regions as known in the art can be incorporated into the structure of a translatable molecule to increase peptide or protein yields.

[00184] A translatable molecule containing a translation enhancer region can provide increased production of peptide or protein.

[00185] In some embodiments, a translation enhancer region can comprise, or be located in a 5' or 3' untranslated region of a translatable molecule.

[00186] Examples of translation enhancer regions include naturally-occurring enhancer regions from TEV 5'UTR and Xenopus beta-globin 3'UTR.

Molecular Structures and Sequences

[00187] A translatable molecule can be designed to express a target peptide or protein. In some embodiments, the target peptide or protein can be associated with a condition or disease in a subject.

[00188] In some aspects, the base sequence of a translatable molecule can include a portion that is identical to at least an effective portion or domain of a base sequence of an mRNA, where an effective portion is sufficient to impart a therapeutic activity to a translation product of the translatable molecule.

[00189] In some aspects, this invention provides active translatable molecules having a base sequence identical to at least a fragment of a native nucleic acid molecule of a cell.

[00190] In certain embodiments, the base sequence of a translatable molecule can include a portion that is identical to a base sequence of an mRNA, except for one or more base mutations. The number of mutations for the translatable molecule should not exceed an amount that would produce a translation product of the translatable molecule having substantially less activity than the mRNA.

[00191] The oligomeric, translatable UNA molecules of this invention can display a sequence of nucleobases, and can be designed to express a peptide or protein, in vitro, ex vivo, or in vivo. The expressed peptide or protein can have activity in various forms, including activity corresponding to protein expressed from a native or natural mRNA.

[00192] In some embodiments, a translatable molecule of this invention may have a chain length of about 400 to 15,000 monomers, where any monomer that is not a UNA monomer can be an N or Q monomer.

Molecular Cap Structure

[00193] A translatable molecule of this invention may have a 5'-end capped with various groups and their analogues as are known in the art. In an exemplary embodiment, the 5' cap may be a m7GpppGm cap. In further embodiments, the 5' cap may be selected from m7GpppA, m7GpppC; unmethylated cap analogs (*e.g.*, GpppG); dimethylated cap analog (*e.g.*, m2,7GpppG), a trimethylated cap analog (*e.g.*, m2,2,7GpppG), dimethylated symmetrical cap analogs (*e.g.*, m7Gpppm7G), or anti reverse cap analogs (*e.g.*, ARCA; m7, 2'OmeGpppG, m72'dGpppG, m7,3'OmeGpppG, m7,3'dGpppG and their tetraphosphate derivatives) (*see, e.g.*, Jemielity, J. et al., RNA 9: 1108-1122 (2003)). In other embodiments, the 5' cap may be an ARCA cap (3'-OMe-m7G(5')pppG). The 5' cap may be an mCAP (m7G(5')ppp(5')G, N⁷-Methyl-Guanosine-5'-Triphosphate-5'-Guanosine). The 5' cap may be resistant to hydrolysis.

[00194] Some examples of 5' cap structures are given in WO2015/051169A2, WO/2015/061491, and US Patent Nos. 8,093,367 and 8,304,529.

Tail Region

[00195] In some embodiments, the translatable oligomer encoding AGL comprises a tail region, which can serve to protect the mRNA from exonuclease degradation. In some embodiments, the tail region can be a polyA tail.

[00196] PolyA tails can be added using a variety of methods known in the art, *e.g.*, using poly A polymerase to add tails to synthetic or *in vitro* transcribed RNA. Other methods include the use of a transcription vector to encode poly A tails or the use of a ligase (*e.g.*, via splint ligation using a T4 RNA ligase and/or T4 DNA ligase), wherein polyA may be ligated to the 3' end of a sense RNA. In some embodiments, a combination of any of the above methods is utilized.

[00197] In some embodiments, a translatable oligomer comprises a 3' polyA tail structure. In some embodiments, the length of the polyA tail can be at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, or 300 nucleotides. In some embodiments, a 3' polyA tail contains about 5 to 300 adenosine nucleotides (*e.g.*, about 30 to 250 adenosine nucleotides, about 60 to 220 adenosine nucleotides, about 80 to 200 adenosine nucleotides, about 90 to about 150 adenosine nucleotides, or about 100 to about 120 adenosine nucleotides). In an

exemplary embodiment, the 3' polyA tail is about 100 nucleotides in length. In another exemplary embodiment, the 3' polyA tail is about 115 nucleotides in length. In another exemplary embodiment, the 3' polyA tail is about 250 nucleotides in length.

[00198] In some embodiments, the 3' polyA tail comprises one or more UNA monomers. In some embodiments, the 3' polyA tail contains 2, 3, 4, 5, 10, 15, 20, or more UNA monomers. In an exemplary embodiment, the 3' polyA tail contains 2 UNA monomers. In a further exemplary embodiment, the 3' polyA tail contains 2 UNA monomers which are found consecutively, *i.e.*, contiguous to each other in the 3' polyA tail.

[00199] In an exemplary embodiment, the 3' polyA tail comprises or consists of a sequence shown in SEQ ID NO: 6. In another exemplary embodiment, the 3' polyA tail comprises or consists of a sequence shown in SEQ ID NO: 38. In yet another exemplary embodiment, the 3' polyA tail comprises or consists of a sequence shown in SEQ ID NO: 39.

[00200] In some embodiments, the translatable oligomer comprises a 3' polyC tail structure. In some embodiments, the length of the polyC tail can be at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, or 300 nucleotides. In some embodiments, a 3' polyC tail contains about 5 to 300 cytosine nucleotides (*e.g.*, about 30 to 250 cytosine nucleotides, about 60 to 220 cytosine nucleotides, about 80 to about 200 cytosine nucleotides, about 90 to 150 cytosine nucleotides, or about 100 to about 120 cytosine nucleotides). In an exemplary embodiment, the 3' polyC tail is about 100 nucleotides in length. In another exemplary embodiment, the 3' polyC tail is about 115 nucleotides in length. The polyC tail may be added to the polyA tail or may substitute the polyA tail. The polyC tail may be added to the 5' end of the polyA tail or the 3' end of the polyA tail.

[00201] In some embodiments, the length of the poly A and/or poly C tail is adjusted to control the stability of a modified translatable oligomeric molecule of the invention and, thus, the transcription of protein. For example, since the length of the polyA tail can influence the half-life of a translatable molecule, the length of the polyA tail can be adjusted to modify the level of resistance of the mRNA to nucleases and thereby control the time course of polynucleotide expression and/or polypeptide production in a target cell.

5' and 3' Untranslated Regions (UTRs)

[00202] In some embodiments, the translatable oligomer encoding AGL may comprise a 5' untranslated region and/or a 3' untranslated region. As is understood in the art, the 5'

and/or 3' UTR may affect an mRNA's stability or efficiency of translation. In an exemplary embodiment, the translatable oligomer comprises a 5' UTR and a 3' UTR.

[00203] In some embodiments, the translatable oligomer may comprise a 5' UTR that is at least about 25, 50, 75, 100, 125, 150, 175, 200, 300, 400, or 500 nucleotides. In some embodiments, a 5' UTR contains about 50 to 300 nucleotides (*e.g.*, about 75 to 250 nucleotides, about 100 to 200 nucleotides, about 120 to 150 nucleotides, or about 135 nucleotides). In an exemplary embodiment, the 5' UTR is about 135 nucleotides in length.

[00204] In some embodiments, the 5' UTR is derived from an mRNA molecule known in the art to be relatively stable (*e.g.*, histone, tubulin, globin, GAPDH, actin, or citric acid cycle enzymes) to increase the stability of the translatable oligomer. In other embodiments, a 5' UTR sequence may include a partial sequence of a CMV immediate-early 1 (IE1) gene. Examples of 5' UTR sequences may be found in US Patent No. 9,149,506. In some embodiments, the 5' UTR comprises a sequence selected from the 5' UTRs of human IL-6, alanine aminotransferase 1, human apolipoprotein E, human fibrinogen alpha chain, human transthyretin, human haptoglobin, human alpha-1-antichymotrypsin, human antithrombin, human alpha-1-antitrypsin, human albumin, human beta globin, human complement C3, human complement C5, SynK, AT1G58420, mouse beta globin, mouse albumin, and a tobacco etch virus, or fragments of any of the foregoing. In an exemplary embodiment, the 5' UTR is derived from a tobacco etch virus (TEV). In a further exemplary embodiment, the 5' UTR comprises or consists of a sequence set forth in SEQ ID NO: 3. In yet another exemplary embodiment, the 5' UTR is a fragment of a sequence set forth in SEQ ID NO: 3, such as a fragment of at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, or 125 contiguous nucleotides of SEQ ID NO: 3.

[00205] In some embodiments, the translatable oligomeric molecule comprises an internal ribosome entry site (IRES). As is understood in the art, an IRES is an RNA element that allows for translation initiation in an end-independent manner. In exemplary embodiments, the IRES is in the 5' UTR. In other embodiments, the IRES may be outside the 5' UTR.

[00206] In some embodiments, the translatable oligomer may comprise a 3' UTR that is at least about 25, 50, 75, 100, 125, 150, 175, 200, 300, 400, or 500 nucleotides. In some embodiments, a 3' UTR contains about 50 to 300 nucleotides (*e.g.*, about 75 to 250 nucleotides,

about 100 to 200 nucleotides, about 140 to 175 nucleotides, or about 160 nucleotides). In an exemplary embodiment, the 3' UTR is about 160 nucleotides in length.

[00207] In some embodiments, the 3' UTR comprises one or more UNA monomers. In some embodiments, the 3' UTR contains 2, 3, 4, 5, 10, 15, 20, or more UNA monomers.

[00208] Examples of 3' UTR sequences may be found in US Patent No. 9,149,506. In some embodiments, the 3' UTR comprises a sequence selected from the 3' UTRs of alanine aminotransferase 1, human apolipoprotein E, human fibrinogen alpha chain, human haptoglobin, human antithrombin, human alpha globin, human beta globin, human complement C3, human growth factor, human hepcidin, MALAT-1, mouse beta globin, mouse albumin, and *Xenopus* beta globin, or fragments of any of the foregoing. In an exemplary embodiment, the 3' UTR is derived from *Xenopus* beta globin. In another exemplary embodiment, the 3' UTR is derived from *Xenopus* beta globin and contains one or more UNA monomers. In a further exemplary embodiment, the 3' UTR comprises or consists of a sequence set forth in SEQ ID NOs: 5 and 33-37. In yet another exemplary embodiment, the 3' UTR is a fragment of a sequence set forth in SEQ ID NOs: 5 and 33-37, such as a fragment of at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 contiguous nucleotides of SEQ ID NO: 5 and 33-37.

[00209] In certain exemplary embodiments, the translatable oligomer encoding AGL comprises a 5' UTR sequence of SEQ ID NO: 3 and a 3' UTR sequence selected from SEQ ID NOs: 5 and 33-37. In some embodiments, the translatable oligomer encoding AGL further comprises a polyA tail shown in SEQ ID NO: 6, SEQ ID NO: 38, or SEQ ID NO: 39. In some embodiments, the mRNA coding sequence of AGL comprises a sequence selected from SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45.

Triple Stop Codon

[00210] In some embodiments, the translatable oligomer encoding AGL may comprise a sequence immediately downstream of the CDS that creates a triple stop codon. The triple stop codon may be incorporated to enhance the efficiency of translation. In some embodiments, the translatable oligomer may comprise the sequence AUAAGUGAA (SEQ ID NO: 40) immediately downstream of a PAH CDS described herein, as exemplified in SEQ ID NOs: 7-32 or SEQ ID NOs: 41-45.

Translation Initiation Sites

[00211] In some embodiments, the translatable oligomer encoding AGL may comprise a translation initiation site. Such sequences are known in the art and include the Kozak sequence. See, for example, Kozak, Marilyn (1988) *Mol. and Cell Biol.*, 8:2737–2744; Kozak, Marilyn (1991) *J. Biol. Chem.*, 266:19867–19870; Kozak, Marilyn (1990) *Proc Natl. Acad. Sci. USA*, 87:8301–8305; and Kozak, Marilyn (1989) *J. Cell Biol.*, 108:229–241; and the references cited therein. As is understood in the art, a Kozak sequence is a short consensus sequence centered around the translational initiation site of eukaryotic mRNAs that allows for efficient initiation of translation of the mRNA. The ribosomal translation machinery recognizes the AUG initiation codon in the context of the Kozak sequence.

[00212] In some embodiments, the translation initiation site, *e.g.*, a Kozak sequence, is inserted upstream of the coding sequence for AGL. In some embodiments, the translation initiation site is inserted downstream of a 5' UTR. In certain exemplary embodiments, the translation initiation site is inserted upstream of the coding sequence for AGL and downstream of a 5' UTR.

[00213] As is understood in the art, the length of the Kozak sequence may vary. Generally, increasing the length of the leader sequence enhances translation.

[00214] In some embodiments, the translatable oligomer encoding AGL comprises a Kozak sequence having the sequence of SEQ ID NO: 4. In certain exemplary embodiments, the translatable oligomer encoding AGL comprises a Kozak sequence having the sequence of SEQ ID NO: 4, wherein the Kozak sequence is immediately downstream of a 5' UTR and immediately upstream of the coding sequence for AGL.

Synthesis Methods

[00215] In various aspects, this invention provides methods for synthesis of translatable messenger molecules.

[00216] Translatable molecules of this invention can be synthesized and isolated using methods disclosed herein, as well as any pertinent techniques known in the art.

[00217] Some methods for preparing nucleic acids are given in, for example, Merino, *Chemical Synthesis of Nucleoside Analogues*, (2013); Gait, *Oligonucleotide synthesis: a*

practical approach (1984); Herdewijn, Oligonucleotide Synthesis, Methods in Molecular Biology, Vol. 288 (2005).

[00218] In some embodiments, a translatable molecule can be made by *in vitro* transcription (IVT) reaction. A mix of nucleoside triphosphates (NTP) can be polymerized using T7 reagents, for example, to yield RNA from a DNA template. The DNA template can be degraded with RNase-free DNase, and the RNA column-separated.

[00219] In some embodiments, a ligase can be used to link a synthetic oligomer to the 3' end of an RNA molecule or an RNA transcript to form a translatable molecule. The synthetic oligomer that is ligated to the 3' end can provide the functionality of a polyA tail, and advantageously provide resistance to its removal by 3'-exoribonucleases. The ligated product translatable molecule can have increased specific activity and provide increased levels of ectopic protein expression.

[00220] In certain embodiments, the ligated product of the translatable molecules of this invention can be made with an RNA transcript that has native specificity. The ligated product can be a synthetic molecule that retains the structure of the RNA transcript at the 5' end to ensure compatibility with the native specificity.

[00221] In further embodiments, the ligated product of the translatable molecules of this invention can be made with an exogenous RNA transcript or non-natural RNA. The ligated product can be a synthetic molecule that retains the structure of the RNA.

[00222] Without wishing to be bound by theory, the canonical mRNA degradation pathway in cells includes the steps: (i) the polyA tail is gradually cut back to a stub by 3' exonucleases, shutting down the looping interaction required for efficient translation and leaving the cap open to attack; (ii) decapping complexes remove the 5' cap; (iii) the unprotected and translationally incompetent residuum of the transcript is degraded by 5' and 3' exonuclease activity.

[00223] Embodiments of this invention involve new translatable structures which can have increased translational activity over a native transcript. Among other things, translatable molecules herein may prevent exonucleases from trimming back the polyA tail in the process of de-adenylation.

[00224] Embodiments of this invention provide structures, compositions and methods for translatable molecules. Embodiments of this invention can provide translatable molecules containing one or more UNA monomers and having increased functional half-life.

[00225] It has been found that ligation of a synthetic oligomer to the 3' end of an mRNA transcript can surprisingly be accomplished with high conversion of the mRNA transcript to the ligation product.

[00226] As used herein, the terms polyA tail and polyA oligomer refer to an oligomer of monomers, wherein the monomers can include nucleotides based on adenine, UNA monomers, naturally-occurring nucleotides, modified nucleotides, or nucleotide analogues.

[00227] Oligomers for ligation to the 3' end of an RNA may be from 2 to 120 monomers in length, or from 3 to 120 monomers in length, or from 4 to 120 monomers in length, or from 5 to 120 monomers in length, or longer. In an exemplary embodiment, the oligomer for ligation is about 30 monomers in length.

Lipid-Based Formulations

[00228] Lipid-based formulations have been increasingly recognized as one of the most promising delivery systems for RNA due to their biocompatibility and their ease of large-scale production. Cationic lipids have been widely studied as synthetic materials for delivery of RNA. After mixing together, nucleic acids are condensed by cationic lipids to form lipid/nucleic acid complexes known as lipoplexes. These lipid complexes are able to protect genetic material from the action of nucleases and to deliver it into cells by interacting with the negatively charged cell membrane. Lipoplexes can be prepared by directly mixing positively charged lipids at physiological pH with negatively charged nucleic acids.

[00229] Conventional liposomes consist of a lipid bilayer that can be composed of cationic, anionic, or neutral (phospho)lipids and cholesterol, which encloses an aqueous core. Both the lipid bilayer and the aqueous space can incorporate hydrophobic or hydrophilic compounds, respectively. Liposome characteristics and behaviour in vivo can be modified by addition of a hydrophilic polymer coating, e.g. polyethylene glycol (PEG), to the liposome surface to confer steric stabilization. Furthermore, liposomes can be used for specific targeting by attaching ligands (e.g., antibodies, peptides, and carbohydrates) to its surface or to the terminal end of the attached PEG chains (Front Pharmacol. 2015 Dec 1;6:286).

[00230] Liposomes are colloidal lipid-based and surfactant-based delivery systems composed of a phospholipid bilayer surrounding an aqueous compartment. They may present as spherical vesicles and can range in size from 20 nm to a few microns. Cationic lipid-based liposomes are able to complex with negatively charged nucleic acids via electrostatic

interactions, resulting in complexes that offer biocompatibility, low toxicity, and the possibility of the large-scale production required for in vivo clinical applications. Liposomes can fuse with the plasma membrane for uptake; once inside the cell, the liposomes are processed via the endocytic pathway and the genetic material is then released from the endosome/carrier into the cytoplasm. Liposomes have long been perceived as drug delivery vehicles because of their superior biocompatibility, given that liposomes are basically analogs of biological membranes, and can be prepared from both natural and synthetic phospholipids (Int J Nanomedicine. 2014; 9: 1833-1843).

[00231] Cationic liposomes have been traditionally the most commonly used non-viral delivery systems for oligonucleotides, including plasmid DNA, antisense oligos, and siRNA/small hairpin RNA (shRNA). Cationic lipids, such as DOTAP, (1,2-dioleoyl-3-trimethylammonium-propane) and DOTMA (N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium methyl sulfate) can form complexes or lipoplexes with negatively charged nucleic acids to form nanoparticles by electrostatic interaction, providing high in vitro transfection efficiency. Furthermore, neutral lipid-based nanoliposomes for RNA delivery as e.g. neutral 1,2-dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC)-based nanoliposomes were developed. (Adv Drug Deliv Rev. 2014 Feb; 66: 110-116.)

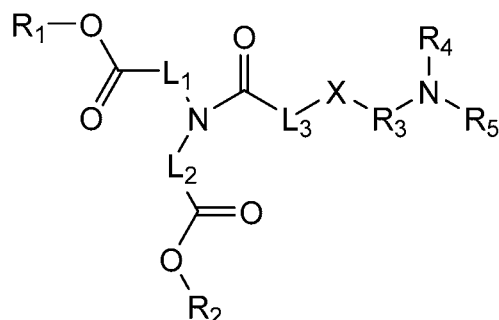
According to some embodiments, the expressible polynucleotides and heterologous mRNA constructs described herein are lipid formulated. The lipid formulation is preferably selected from, but not limited to, liposomes, lipoplexes, copolymers, such as PLGA, and lipid nanoparticles. In one preferred embodiment, a lipid nanoparticle (LNP) comprises:

- (a) a nucleic acid,
- (b) a cationic lipid,
- (c) an aggregation reducing agent (such as polyethylene glycol (PEG) lipid or PEG-modified lipid),
- (d) optionally a non-cationic lipid (such as a neutral lipid), and
- (e) optionally, a sterol.

[00232] In one embodiment, the lipid nanoparticle formulation consists of (i) at least one cationic lipid; (ii) a neutral lipid; (iii) a sterol, e.g., cholesterol; and (iv) a PEG-lipid, in a molar ratio of about 20-60% cationic lipid: 5-25% neutral lipid: 25-55% sterol; 0.5-15% PEG-lipid.

Thiocarbamate and Carbamate-Containing Lipid Formulations

[00233] Some examples of lipids and lipid compositions for delivery of an active molecule of this invention are given in WO/2015/074085 and USSN 15/387,067, each of which is hereby incorporated by reference in its entirety. In certain embodiments, the lipid is a compound of the following formula:



wherein

R₁ and R₂ both consist of a linear alkyl consisting of 1 to 14 carbons, or an alkenyl or alkynyl consisting of 2 to 14 carbons;

L₁ and L₂ both consist of a linear alkylene or alkenylene consisting of 5 to 18 carbons, or forming a heterocycle with N;

X is S;

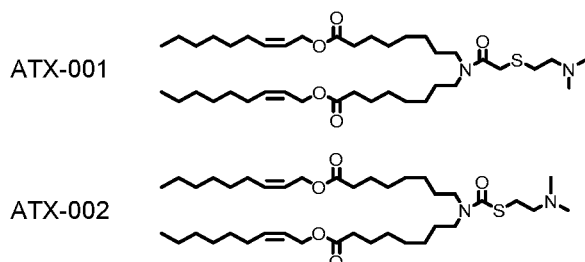
L₃ consists of a bond or a linear alkylene consisting of 1 to 6 carbons, or forming a heterocycle with N;

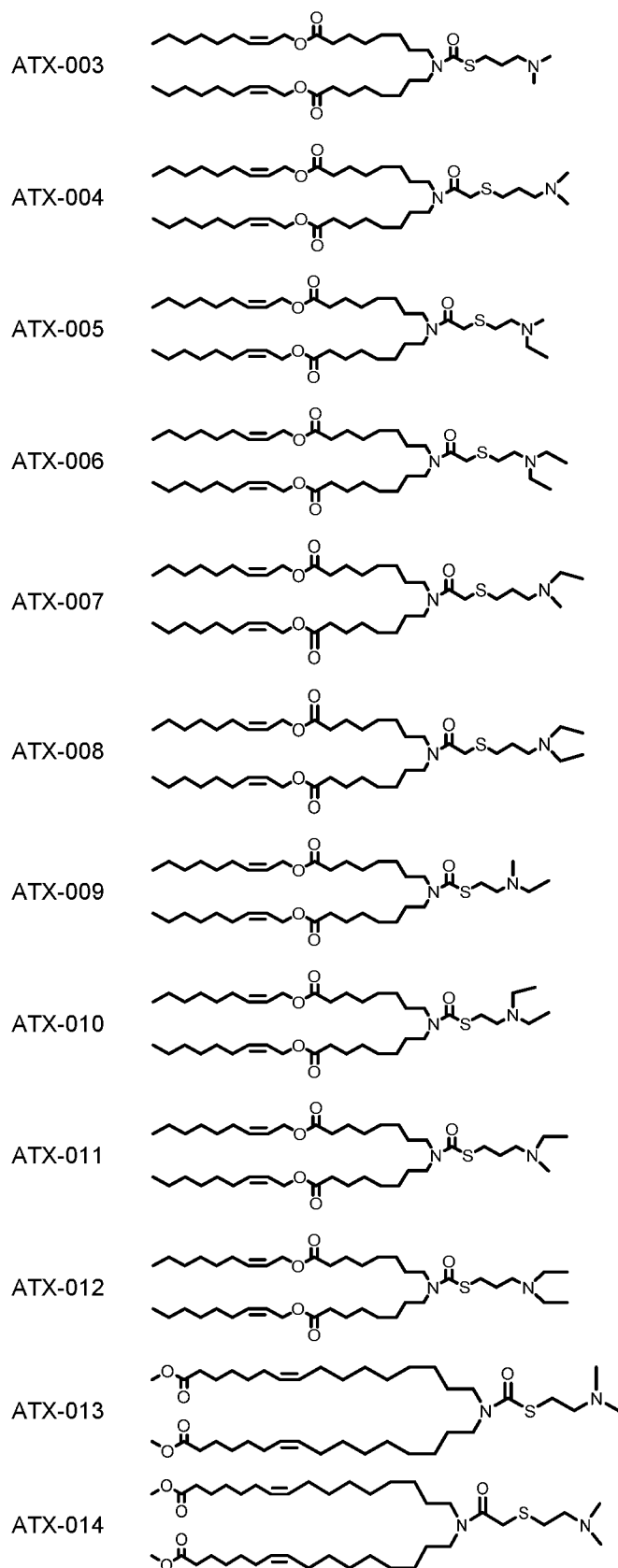
R₃ consists of a linear or branched alkylene consisting of 1 to 6 carbons; and

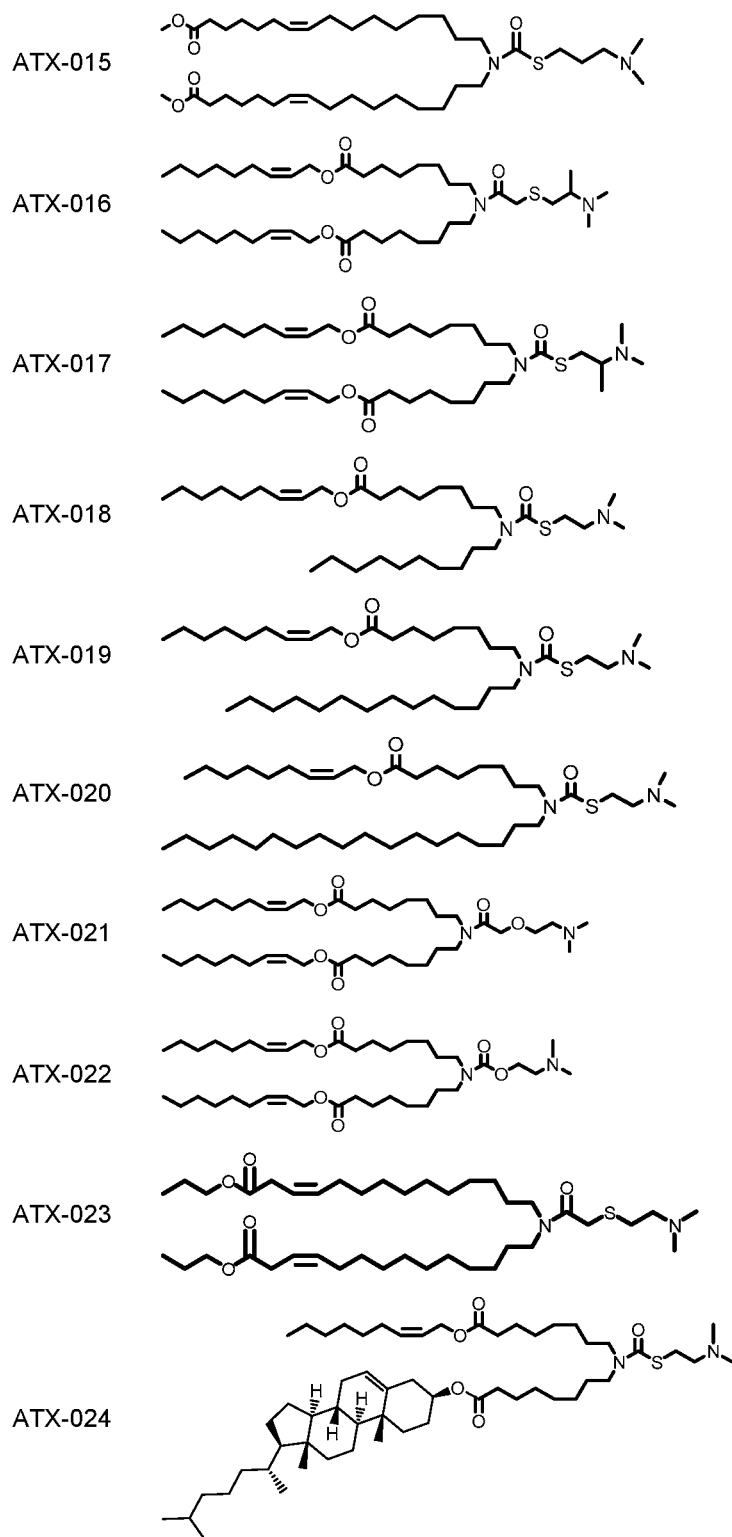
R₄ and R₅ are the same or different, each consisting of a hydrogen or a linear or branched alkyl consisting of 1 to 6 carbons;

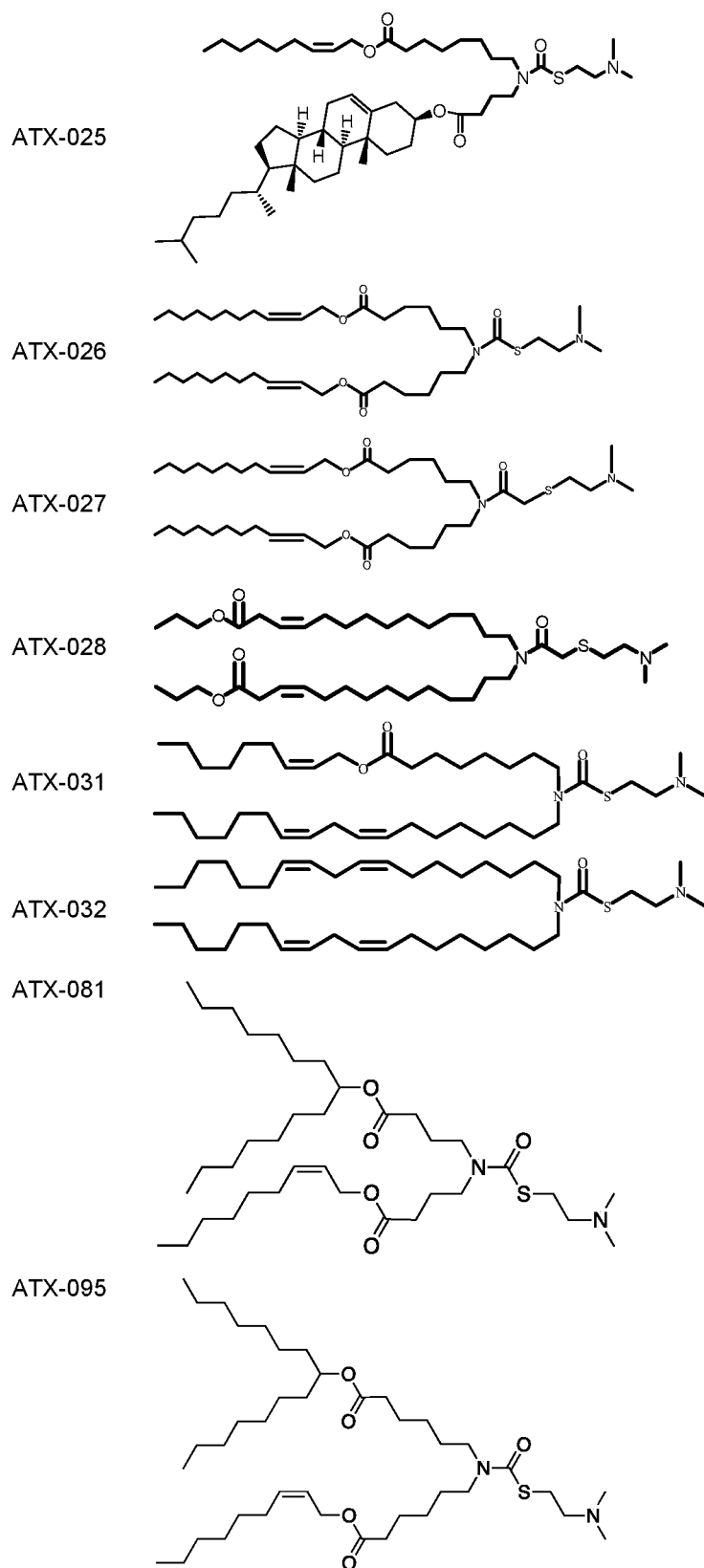
or a pharmaceutically acceptable salt thereof.

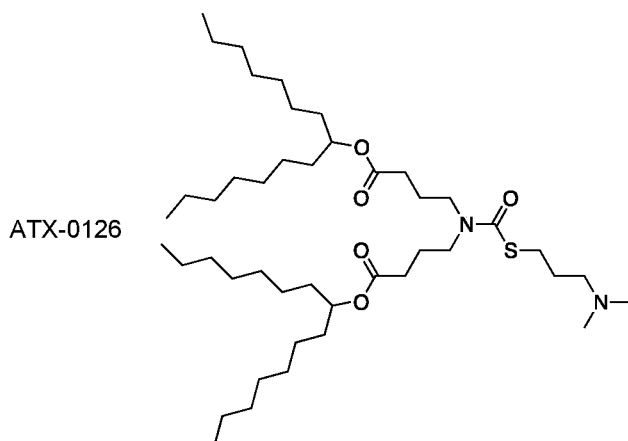
[00234] The lipid formulation may contain one or more ionizable cationic lipids selected from among the following:











Cationic Lipids

[00235] The lipid nanoparticle preferably includes a cationic lipid suitable for forming a lipid nanoparticle. Preferably, the cationic lipid carries a net positive charge at about physiological pH.

[00236] The cationic lipid may be, for example, N,N-dioleoyl-N,N-dimethylammonium chloride (DODAC), N,N-distearyl-N,N-dimethylammonium bromide (DDAB), 1,2-dioleoyltrimethylammoniumpropane chloride (DOTAP) (also known as N-(2,3-dioleoyloxy)propyl)-N,N,N-trimethylammonium chloride and 1,2-Dioleoyloxy-3-trimethylaminopropane chloride salt), N-(1-(2,3-dioleoyloxy)propyl)-N,N,N-trimethylammonium chloride (DOTMA), N,N-dimethyl-2,3-dioleoyloxy)propylamine (DODMA), 1,2-DiLinoleoyloxy-N,N-dimethylaminopropane (DLinDMA), 1,2-Dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA), 1,2-di-γ-linolenyloxy-N,N-dimethylaminopropane (γ-DLenDMA), 1,2-Dilinoleylcarbamoyloxy-3-dimethylaminopropane (DLin-C-DAP), 1,2-Dilinoleoyloxy-3-(dimethylamino)acetoxyp propane (DLin-DAC), 1,2-Dilinoleoyloxy-3-morpholinopropane (DLin-MA), 1,2-Dilinoleoyl-3-dimethylaminopropane (DLinDAP), 1,2-Dilinoleylthio-3-dimethylaminopropane (DLin-S-DMA), 1-Linoleoyl-2-linoleoyloxy-3-dimethylaminopropane (DLin-2-DMAP), 1,2-Dilinoleoyloxy-3-trimethylaminopropane chloride salt (DLin-TMA.Cl), 1,2-Dilinoleoyl-3-trimethylaminopropane chloride salt (DLin-TAP.Cl), 1,2-Dilinoleoyloxy-3-(N-methylpiperazino)propane (DLin-MPZ), or 3-(N,N-Dilinoleylamino)-1,2-propanediol (DLinAP), 3-(N,N-Dioleylamino)-1,2-propanedio (DOAP), 1,2-Dilinoleyloxo-3-(2-N,N-dimethylamino)ethoxypropane (DLin-EG-DMA), 2,2-Dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA) or analogs thereof, (3aR,5s,6aS)-N,N-dimethyl-2,2-di((9Z,12Z)-octadeca-9,12-dienyl)tetrahydro-3aH-

cyclopenta[d][1,3]dioxol-5-amine, (6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (MC3), 1,1'-(2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethylazanediy)didodecan-2-ol (C12-200), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-K-C2-DMA), 2,2-dilinoleyl-4-dimethylaminomethyl-1,3-dioxolane (DLin-K-DMA), (6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (DLin-M-C3-DMA), 3-(((6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yloxy)-N,N-dimethylpropan-1-amine (MC3 Ether), 4-(((6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yloxy)-N,N-dimethylbutan-1-amine (MC4 Ether), or any combination of any of the foregoing. Other cationic lipids include, but are not limited to, N,N-distearyl-N,N-dimethylammonium bromide (DDAB), 3P-(N-(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Choi), N-(1-(2,3-dioleoyloxy)propyl)-N-2-(sperminecarboxamido)ethyl)-N,N-dimethylammonium trifluoroacetate (DOSPA), dioctadecylamidoglycyl carboxyspermine (DOGS), 1,2-dioleoyl-sn-3-phosphoethanolamine (DOPE), 1,2-dioleoyl-3-dimethylammonium propane (DODAP), N-(1,2-dimyristyloxyprop-3-yl)-N,N-dimethyl-N-hydroxyethyl ammonium bromide (DMRIE), and 2,2-Dilinoleyl-4-dimethylaminoethyl-1,3-dioxolane (XTC). Additionally, commercial preparations of cationic lipids can be used, such as, e.g., LIPOFECTIN (including DOTMA and DOPE, available from GIBCO/BRL), and Lipofectamine (comprising DOSPA and DOPE, available from GIBCO/BRL).

[00237] Other suitable cationic lipids are disclosed in International Publication Nos. WO 09/086558, WO 09/127060, WO 10/048536, WO 10/054406, WO 10/088537, WO 10/129709, and WO 2011/153493; U.S. Patent Publication Nos. 2011/0256175, 2012/0128760, and 2012/0027803; U.S. Patent Nos. 8,158,601; and Love et al, PNAS, 107(5), 1864-69, 2010. Other suitable amino lipids include those having alternative fatty acid groups and other dialkylamino groups, including those, in which the alkyl substituents are different (e.g., N-ethyl-, N-methylamino-, and N-propyl-N-ethylamino-). In general, amino lipids having less saturated acyl chains are more easily sized, particularly when the complexes must be sized below about 0.3 microns, for purposes of filter sterilization. Amino lipids containing unsaturated fatty acids with carbon chain lengths in the range of C14 to C22 may be used. Other scaffolds can also be used to separate the amino group and the fatty acid or fatty alkyl portion of the amino lipid.

[00238] In a further preferred embodiment, the LNP comprises the cationic lipid with formula (III) according to the patent application PCT/EP2017/064066. In this context, the disclosure of PCT/EP2017/064066 is also incorporated herein by reference.

[00239] In certain embodiments, amino or cationic lipids of the invention have at least one protonatable or deprotonatable group, such that the lipid is positively charged at a pH at or below physiological pH (e.g. pH 7.4), and neutral at a second pH, preferably at or above physiological pH. It will, of course, be understood that the addition or removal of protons as a function of pH is an equilibrium process, and that the reference to a charged or a neutral lipid refers to the nature of the predominant species and does not require that all of the lipid be present in the charged or neutral form. Lipids that have more than one protonatable or deprotonatable group, or which are zwitterionic, are not excluded from use in the invention. In certain embodiments, the protonatable lipids have a pKa of the protonatable group in the range of about 4 to about 11, e.g., a pKa of about 5 to about 7.

[00240] The cationic lipid can comprise from about 20 mol % to about 70 or 75 mol % or from about 45 to about 65 mol % or about 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, or about 70 mol % of the total lipid present in the particle. In another embodiment, the lipid nanoparticles include from about 25% to about 75% on a molar basis of cationic lipid, e.g., from about 20 to about 70%, from about 35 to about 65%, from about 45 to about 65%, about 60%, about 57.5%, about 57.1%, about 50% or about 40% on a molar basis (based upon 100% total moles of lipid in the lipid nanoparticle). In one embodiment, the ratio of cationic lipid to nucleic acid is from about 3 to about 15, such as from about 5 to about 13 or from about 7 to about 11.

Pharmaceutical Compositions

[00241] In some aspects, this invention provides pharmaceutical compositions containing a translatable compound and a pharmaceutically acceptable carrier.

[00242] A pharmaceutical composition can be capable of local or systemic administration. In some aspects, a pharmaceutical composition can be capable of any modality of administration. In certain aspects, the administration can be by any route, including intravenous, subcutaneous, pulmonary, intramuscular, intraperitoneal, dermal, oral, inhalation or nasal administration.

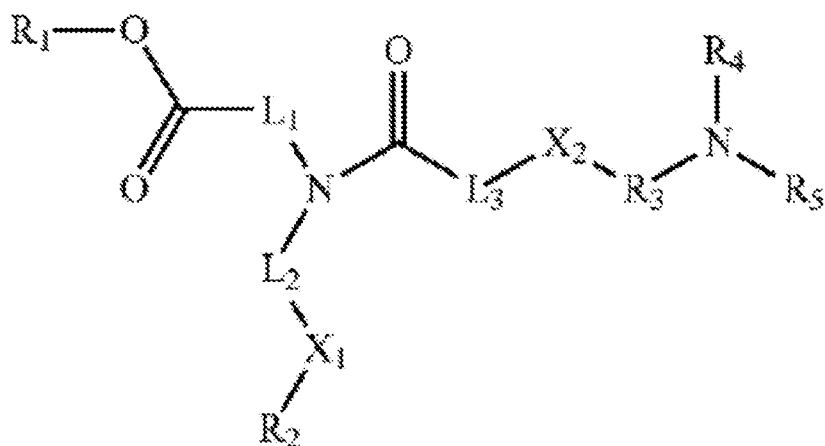
[00243] Embodiments of this invention include pharmaceutical compositions containing a translatable compound in a lipid formulation.

[00244] In some embodiments, a pharmaceutical composition may comprise one or more lipids selected from cationic lipids, anionic lipids, sterols, pegylated lipids, and any combination of the foregoing. In some embodiments, the pharmaceutical composition containing a translatable compound comprises a cationic lipid, a phospholipid, cholesterol, and a pegylated lipid.

[00245] In certain embodiments, a pharmaceutical composition can be substantially free of liposomes.

[00246] In further embodiments, a pharmaceutical composition can include nanoparticles.

[00247] Some examples of lipids and lipid compositions for delivery of an active molecule of this invention are given in WO/2015/074085, which is hereby incorporated by reference in its entirety. In certain embodiments, the lipid is a cationic lipid. In some embodiment, the cationic lipid comprises a compound of formula II:



Formula II,

in which R₁ and R₂ are the same or different, each a linear or branched alkyl, alkenyl, or alkynyl, L₁ and L₂ are the same or different, each a linear alkyl having at least five carbon atoms, or form a heterocycle with the N, X₁ is a bond, or is --CO--O-- whereby L₂--CO--O--R₂ is formed X₂ is S or O, L₃ is a bond or a lower alkyl, R₃ is a lower alkyl, R₄ and R₅ are the same or different, each a lower alkyl. What is also described herein is the compound of formula II, in which L₃ is absent, R₁ and R₂ each consists of at least seven carbon atoms, R₃ is ethylene or n-propylene, R₄ and R₅ are methyl or ethyl, and L₁ and L₂ each consists of a linear alkyl having at least five carbon atoms. What is also described herein is the compound of formula II, in

which L_3 is absent, R_1 and R_2 each consists of at least seven carbon atoms, R_3 is ethylene or n-propylene, R_4 and R_5 are methyl or ethyl, and L_1 and L_2 each consists of a linear alkyl having at least five carbon atoms. What is also described herein is the compound of formula II, in which L_3 is absent, R_1 and R_2 each consists of an alkenyl of at least nine carbon atoms, R_3 is ethylene or n-propylene, R_4 and R_5 are methyl or ethyl, and L_1 and L_2 each consists of a linear alkyl having at least five carbon atoms. What is also described herein is the compound of formula II, in which L_3 is methylene, R_1 and R_2 each consists of at least seven carbon atoms, R_3 is ethylene or n-propylene, R_4 and R_5 are methyl or ethyl, and L_1 and L_2 each consists of a linear alkyl having at least five carbon atoms. What is also described herein is the compound of formula II, in which L_3 is methylene, R_1 and R_2 each consists of at least nine carbon atoms, R_3 is ethylene or n-propylene, R_4 and R_5 are each methyl, L_1 and L_2 each consists of a linear alkyl having at least seven carbon atoms. What is also described herein is the compound of formula II, in which L_3 is methylene, R_1 consists of an alkenyl having at least nine carbon atoms and R_2 consists of an alkenyl having at least seven carbon atoms, R_3 is n-propylene, R_4 and R_5 are each methyl, L_1 and L_2 each consists of a linear alkyl having at least seven carbon atoms. What is also described herein is the compound of formula II, in which L_3 is methylene, R_1 and R_2 each consists of an alkenyl having at least nine carbon atoms, R_3 is ethylene, R_4 and R_5 are each methyl, L_1 and L_2 each consists of a linear alkyl having at least seven carbon atoms.

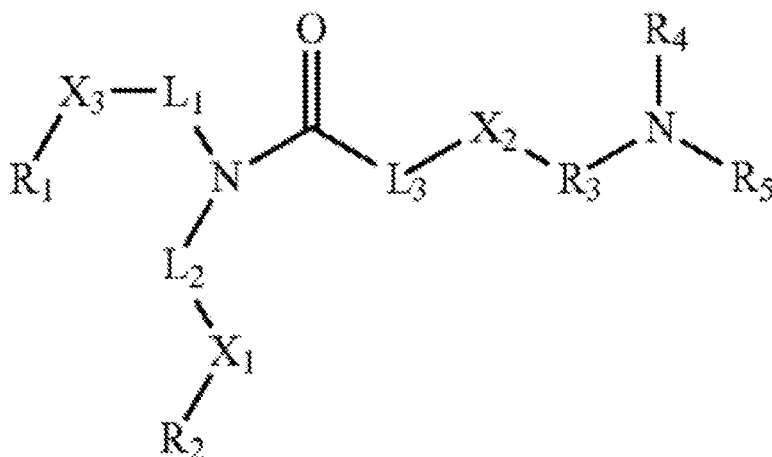
[00248] In exemplary embodiments, the cationic lipid comprises a compound of selected from the group consisting of ATX-001, ATX-002, ATX-003, ATX-004, ATX-005, ATX-006, ATX-007, ATX-008, ATX-009, ATX-010, ATX-011, ATX-012, ATX-013, ATX-014, ATX-015, ATX-016, ATX-017, ATX-018, ATX-019, ATX-020, ATX-021, ATX-022, ATX-023, ATX-024, ATX-025, ATX-026, ATX-027, ATX-028, ATX-029, ATX-030, ATX-031, ATX-032, ATX-081, ATX-095, and ATX-126, or a pharmaceutically acceptable salt thereof.

[00249] In certain exemplary embodiments, the cationic lipid comprises ATX-002, ATX-081, ATX-095, or ATX-126.

[00250] In some embodiments, the cationic lipid or a pharmaceutically acceptable salt thereof, may be presented in a lipid composition, comprising a nanoparticle or a bilayer of lipid molecules. The lipid bilayer preferably further comprises a neutral lipid or a polymer. The lipid composition preferably comprises a liquid medium. The composition preferably further encapsulates a translatable compound of the present invention. The lipid composition

preferably further comprises a translatable compound of the present invention and a neutral lipid or a polymer. The lipid composition preferably encapsulates the translatable compound.

[00251] In further embodiments, the cationic lipid comprises a compound of formula III:



Formula III,

wherein R_1 and R_2 are the same or different, each a linear or branched alkyl consisting of 1 to 9 carbons, an alkenyl or alkynyl consisting of 2 to 11 carbons, or cholesteryl, L_1 and L_2 are the same or different, each a linear alkylene or alkenylene consisting of 5 to 18 carbons, X_1 is --CO--O-- whereby $-L_2\text{--CO--O--}R_2$ is formed, X_2 is S or O, X_3 is --CO--O-- whereby $-L_1\text{--CO--O--}R_1$ is formed, L_3 is a bond, R_3 is a linear or branched alkylene consisting of 1 to 6 carbons, and R_4 and R_5 are the same or different, each hydrogen or a linear or branched alkyl consisting of 1 to 6 carbons; or a pharmaceutically acceptable salt thereof. In one embodiment, X_2 is S. In another embodiment, R_3 is selected from ethylene, n-propylene, or isobutylene. In yet another embodiment, R_4 and R_5 are separately methyl, ethyl, or isopropyl. In yet another embodiment, L_1 and L_2 are the same. In yet another embodiment, L_1 and L_2 differ. In yet another embodiment, L_1 or L_2 consists of a linear alkylene having seven carbons. In yet another embodiment, L_1 or L_2 consists of a linear alkylene having nine carbons. In yet another embodiment, R_1 and R_2 are the same. In yet another embodiment, R_1 and R_2 differ. In yet another embodiment, R_1 and R_2 each consists of an alkenyl. In yet another embodiment, R_1 and R_2 each consists of an alkyl. In yet another embodiment, the alkenyl consists of a single double bond. In yet another embodiment, R_1 or R_2 consists of nine carbons. In yet another embodiment, R_1 or R_2 consists of eleven carbons. In yet another embodiment, R_1 or R_2 consists of seven carbons. In yet another embodiment, L_3 is a bond, R_3 is ethylene, X_2 is S, and R_4 and

R₅ are each methyl. In yet another embodiment, L₃ is a bond, R₃ is n-propylene, X₂ is S, R₄ and R₅ are each methyl. In yet another embodiment, L₃ is a bond, R₃ is ethylene, X₂ is S, and R₄ and R₅ are each ethyl.

[00252] As would be appreciated by the skilled artisan, the compounds of formulas II and III form salts that are also within the scope of this disclosure. Reference to a compound of formulas II and III herein is understood to include reference to salts thereof, unless otherwise indicated. The term "salt(s)", as employed herein, denotes acidic salts formed with inorganic and/or organic acids, as well as basic salts formed with inorganic and/or organic bases. In addition, when a compound of formula II or III contains both a basic moiety, such as, but not limited to, a pyridine or imidazole, and an acidic moiety, such as, but not limited to, a carboxylic acid, zwitterions ("inner salts") may be formed and are included within the term "salt(s)" as used herein. The salts can be pharmaceutically acceptable (i.e., non-toxic, physiologically acceptable) salts, although other salts are also useful. Salts of the compounds of the formula II or III may be formed, for example, by reacting a compound of formula II or III with an amount of acid or base, such as an equivalent amount, in a medium such as one in which the salt precipitates or in an aqueous medium followed by lyophilization.

[00253] Exemplary acid addition salts include acetates, adipates, alginates, ascorbates, aspartates, benzoates, benzenesulfonates, bisulfates, borates, butyrates, citrates, camphorates, camphorsulfonates, cyclopentanepropionates, digluconates, dodecylsulfates, ethanesulfonates, fumarates, glucoheptanoates, glycerophosphates, hemisulfates, heptanoates, hexanoates, hydrochlorides, hydrobromides, hydroiodides, 2-hydroxyethanesulfonates, lactates, maleates, methanesulfonates, 2-naphthalenesulfonates, nicotines, nitrates, oxalates, pectinates, persulfates, 3-phenylpropionates, phosphates, picrates, pivalates, propionates, salicylates, succinates, sulfates, sulfonates (such as those mentioned herein), tartarates, thiocyanates, toluenesulfonates (also known as tosylates) undecanoates, and the like. Additionally, acids which are generally considered suitable for the formation of pharmaceutically useful salts from basic pharmaceutical compounds are discussed, for example, by S. Berge et al, J. Pharmaceutical Sciences (1977) 66(1)1-19; P. Gould, International J. Pharmaceutics (1986) 33 201-217; Anderson et al, The Practice of Medicinal Chemistry (1996), Academic Press, New York; and in The Orange Book (Food & Drug Administration, Washington, D.C. on their website). These disclosures are incorporated by reference herein.

[00254] Exemplary basic salts include ammonium salts, alkali metal salts such as sodium, lithium, and potassium salts, alkaline earth metal salts such as calcium and magnesium

salts, salts with organic bases (for example, organic amines) such as benzathines, dicyclohexylamines, hydrabamines (formed with N,N-bis(dehydroabietyl)ethylenediamine), N-methyl-D-glucamines, N-methyl-D-glucamides, t-butyl amines, and salts with amino acids such as arginine, lysine, and the like. Basic nitrogen-containing groups may be quarternized with agents such as lower alkyl halides (e.g., methyl, ethyl, propyl, and butyl chlorides, bromides, and iodides), dialkyl sulfates (e.g., dimethyl, diethyl, dibutyl, and diamyl sulfates), long chain halides (e.g., decyl, lauryl, myristyl, and stearyl chlorides, bromides, and iodides), arylalkyl halides (e.g., benzyl and phenethyl bromides), and others.

[00255] All such acid and base salts are intended to be pharmaceutically acceptable salts within the scope of the disclosure and all acid and base salts are considered equivalent to the free forms of the corresponding compounds for purposes of the disclosure. Compounds of formula II or III can exist in unsolvated and solvated forms, including hydrated forms. In general, the solvated forms, with pharmaceutically acceptable solvents such as water, ethanol, and the like, are equivalent to the unsolvated forms for the purposes of this disclosure. Compounds of formula II or III and salts, solvates thereof, may exist in their tautomeric form (for example, as an amide or imino ether). All such tautomeric forms are contemplated herein as part of the present disclosure.

[00256] The cationic lipid compounds described herein may be combined with a translatable compound of the invention to form microparticles, nanoparticles, liposomes, or micelles. The translatable compound of the invention to be delivered by the particles, liposomes, or micelles may be in the form of a gas, liquid, or solid. The cationic lipid compound and the translatable compound may be combined with other cationic lipid compounds, polymers (synthetic or natural), surfactants, cholesterol, carbohydrates, proteins, lipids, etc. to form the particles. These particles may then optionally be combined with a pharmaceutical excipient to form a pharmaceutical composition.

[00257] In certain embodiments, the cationic lipid compounds are relatively non-cytotoxic. The cationic lipid compounds may be biocompatible and biodegradable. The cationic lipid may have a pKa in the range of approximately 5.5 to approximately 7.5, more preferably between approximately 6.0 and approximately 7.0. It may be designed to have a

desired pKa between approximately 3.0 and approximately 9.0, or between approximately 5.0 and approximately 8.0.

[00258] A composition containing a cationic lipid compound may be 30-70% cationic lipid compound, 0-60% cholesterol, 0-30% phospholipid and 1-10% polyethylene glycol (PEG). Preferably, the composition is 30-40% cationic lipid compound, 40-50% cholesterol, and 10-20% PEG. In other preferred embodiments, the composition is 50-75% cationic lipid compound, 20-40% cholesterol, and 5 to 10% phospholipid, and 1-10% PEG. The composition may contain 60-70% cationic lipid compound, 25-35% cholesterol, and 5-10% PEG. The composition may contain up to 90% cationic lipid compound and 2 to 15% helper lipid. The formulation may be a lipid particle formulation, for example containing 8-30% compound, 5-30% helper lipid, and 0-20% cholesterol; 4-25% cationic lipid, 4-25% helper lipid, 2 to 25% cholesterol, 10 to 35% cholesterol-PEG, and 5% cholesterol-amine; or 2-30% cationic lipid, 2-30% helper lipid, 1 to 15% cholesterol, 2 to 35% cholesterol-PEG, and 1-20% cholesterol-amine; or up to 90% cationic lipid and 2-10% helper lipids, or even 100% cationic lipid.

[00259] In some embodiments, the one or more cholesterol-based lipids are selected from cholesterol, PEGylated cholesterol and DC-Chol (N,N-dimethyl-N-ethylcarboxamidocholesterol), and 1,4-bis(3-N-oleylamino-propyl)piperazine. In an exemplary embodiment, the cholesterol-based lipid is cholesterol.

[00260] In some embodiments, the one or more pegylated lipids, *i.e.*, PEG-modified lipids. In some embodiments, the one or more PEG-modified lipids comprise a poly(ethylene) glycol chain of up to 5 kDa in length covalently attached to a lipid with alkyl chain(s) of C₆-C₂₀ length. In some embodiments, a PEG-modified lipid is a derivatized ceramide such as N-Octanoyl-Sphingosine-1-[Succinyl(Methoxy Polyethylene Glycol)-2000]. In some embodiments, a PEG-modified or PEGylated lipid is PEGylated cholesterol or Dimyristoylglycerol (DMG)-PEG-2K. In an exemplary embodiment, the PEG-modified lipid is PEGylated cholesterol.

[00261] In additional embodiments, a pharmaceutical composition can contain an oligomeric compound within a viral or bacterial vector.

[00262] A pharmaceutical composition of this disclosure may include carriers, diluents or excipients as are known in the art. Examples of pharmaceutical compositions and methods are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co.

(A.R. Gennaro ed. 1985), and Remington, The Science and Practice of Pharmacy, 21st Edition (2005).

[00263] Examples of excipients for a pharmaceutical composition include antioxidants, suspending agents, dispersing agents, preservatives, buffering agents, tonicity agents, and surfactants.

[00264] An effective dose of an agent or pharmaceutical formulation of this invention can be an amount that is sufficient to cause translation of a translatable molecule in a cell.

[00265] A therapeutically effective dose can be an amount of an agent or formulation that is sufficient to cause a therapeutic effect. A therapeutically effective dose can be administered in one or more separate administrations, and by different routes. As will be appreciated in the art, a therapeutically effective dose or a therapeutically effective amount is largely determined based on the total amount of the therapeutic agent contained in the pharmaceutical compositions of the present invention. Generally, a therapeutically effective amount is sufficient to achieve a meaningful benefit to the subject (*e.g.*, treating, modulating, curing, preventing and/or ameliorating GSD III). For example, a therapeutically effective amount may be an amount sufficient to achieve a desired therapeutic and/or prophylactic effect. Generally, the amount of a therapeutic agent (*e.g.*, a translatable oligomer encoding AGL) administered to a subject in need thereof will depend upon the characteristics of the subject. Such characteristics include the condition, disease severity, general health, age, sex and body weight of the subject. One of ordinary skill in the art will be readily able to determine appropriate dosages depending on these and other related factors. In addition, both objective and subjective assays may optionally be employed to identify optimal dosage ranges.

[00266] Methods provided herein contemplate single as well as multiple administrations of a therapeutically effective amount of the translatable compound (*e.g.*, a translatable oligomer encoding AGL) described herein. Pharmaceutical compositions comprising a translatable compound encoding AGL can be administered at regular intervals, depending on the nature, severity and extent of the subject's condition (*e.g.*, the severity of a subject's GSD III disease state and the associated symptoms of GSD III, and/or the subject's AGL activity levels). In some embodiments, a therapeutically effective amount of the translatable compound (*e.g.*, a translatable oligomer encoding AGL) of the present invention may be administered periodically at regular intervals (*e.g.*, once every year, once every six months, once every four months, once every three months, once every two months, once a month), biweekly, weekly,

daily, twice a day, three times a day, four times a day, five times a day, six times a day, or continuously.

[00267] In some embodiments, the pharmaceutical compositions of the present invention are formulated such that they are suitable for extended-release of the translatable compound encoding AGL contained therein. Such extended-release compositions may be conveniently administered to a subject at extended dosing intervals. For instance, in one embodiment, the pharmaceutical compositions of the present invention are administered to a subject twice a day, daily or every other day. In some embodiments, the pharmaceutical compositions of the present invention are administered to a subject twice a week, once a week, every 10 days, every two weeks, every 28 days, every month, every six weeks, every eight weeks, every other month, every three months, every four months, every six months, every nine months or once a year. Also contemplated herein are pharmaceutical compositions which are formulated for depot administration (*e.g.*, subcutaneously, intramuscularly) to either deliver or release a translatable compound encoding AGL over extended periods of time. Preferably, the extended-release means employed are combined with modifications made to the translatable compound encoding AGL to enhance stability.

[00268] In some embodiments, a therapeutically effective dose, upon administration, can result in serum or plasma levels of AGL of 1-1000 pg/ml, or 1-1000 ng/ml, or 1-1000 µg/ml, or more.

[00269] In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of the invention can result in increased liver AGL protein levels in a treated subject. In some embodiments, administering a composition comprising a translatable molecule of the invention results in a 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95% increase in liver AGL protein levels relative to a baseline AGL protein level in the subject prior to treatment. In certain embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of the invention will result an increase in liver AGL levels relative to baseline liver AGL levels in the subject prior to treatment. In some embodiments, the increase in liver AGL levels relative to baseline liver AGL levels will be at least 5%, 10%, 20%, 30%, 40%, 50%, 100%, 200%, or more.

[00270] In some embodiments, a therapeutically effective dose, when administered regularly, results in increased expression of AGL in the liver as compared to baseline levels

prior to treatment. In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of the invention results in the expression of a AGL protein level at or above about 10 ng/mg, about 20 ng/mg, about 50 ng/mg, about 100 ng/mg, about 150 ng/mg, about 200 ng/mg, about 250 ng/mg, about 300 ng/mg, about 350 ng/mg, about 400 ng/mg, about 450 ng/mg, about 500 ng/mg, about 600 ng/mg, about 700 ng/mg, about 800 ng/mg, about 900 ng/mg, about 1000 ng/mg, about 1200 ng/mg or about 1500 ng/mg of the total protein in the liver of a treated subject.

[00271] In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable oligomer encoding AGL will result in reduced levels of one or more of markers selected from alanine transaminase (ALT), aspartate transaminase (AST), alkaline phosphatase (ALP), creatine phosphokinase (CPK), glycogen, and limit dextrin (*i.e.*, a low-molecular carbohydrate produced by the hydrolysis of glycogen).

[00272] In some embodiments, a therapeutically effective dose, when administered regularly, results in a reduction of ALT, AST, ALP, and/or CPK levels in a biological sample. In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of ALT, AST, ALP, and/or CPK levels in a biological sample (*e.g.*, a plasma or serum sample) by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% as compared to baseline ALT, AST, ALP, and/or CPK levels before treatment. In some embodiments, the biological sample is selected from plasma, serum, whole blood, urine, or cerebrospinal fluid.

[00273] In certain exemplary embodiments, a therapeutically effective dose, when administered regularly, results in a reduction of ALT levels, *e.g.*, as measured in units of ALT activity/liter (U/l), in a serum or plasma sample. In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of ALT levels in a biological sample (*e.g.*, a plasma or serum sample) by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% as compared to baseline ALT levels before treatment. In an exemplary

embodiment, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of ALT levels in a biological sample (*e.g.*, a plasma or serum sample) by at least about 50% as compared to baseline ALT levels before treatment. In a further exemplary embodiment, ALT levels are measured after fasting, *e.g.*, after 6, 8, 10, 12, 18, or 24 hours of fasting.

[00274] In other exemplary embodiments, a therapeutically effective dose, when administered regularly, results in a reduction of AST levels, *e.g.*, as measured in units of AST activity/liter (U/l), in a serum or plasma sample. In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of AST levels in a biological sample (*e.g.*, a plasma or serum sample) by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% as compared to baseline AST levels before treatment. In an exemplary embodiment, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of AST levels in a biological sample (*e.g.*, a plasma or serum sample) by at least about 50% as compared to baseline AST levels before treatment. In a further exemplary embodiment, AST levels are measured after fasting, *e.g.*, after 6, 8, 10, 12, 18, or 24 hours of fasting.

[00275] Measurements of ALT, AST, ALP, and/or CPK levels can be made using any method known in the art, *e.g.*, using a Fuji Dri-Chem Clinical Chemistry Analyzer FDC 3500 as described in Liu *et al.*, 2014, *Mol Genet and Metabolism* 111: 467-76.

[00276] In other exemplary embodiments, a therapeutically effective dose, when administered regularly, results in a reduction of glycogen levels in a biological sample. In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of glycogen accumulation in a biological sample (*e.g.*, a liver sample) by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% as compared to baseline glycogen levels before treatment. In some embodiments, the biological sample is a portion of an organ selected from

liver, heart, diaphragm, quadriceps, and gastrocnemius. In an exemplary embodiment, the biological sample is a liver section, *e.g.*, a section of hepatocytes.

[00277] In other exemplary embodiments, a therapeutically effective dose, when administered regularly, results in a reduction of limit dextrin levels in a biological sample. In some embodiments, administering a therapeutically effective dose of a composition comprising a translatable molecule of this invention results in a reduction of limit dextrin accumulation in a biological sample (*e.g.*, a liver sample) by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% as compared to baseline limit dextrin levels before treatment. In some embodiments, the biological sample is a portion of an organ selected from liver, heart, diaphragm, quadriceps, and gastrocnemius. In an exemplary embodiment, the biological sample is a liver section, *e.g.*, a section of hepatocytes. In a further exemplary embodiment, a therapeutically effective dose, when administered regularly, results in at least a 50%, 60%, 70%, or 80% reduction of limit dextrin levels in a liver sample as compared to baseline limit dextrin levels before treatment.

[00278] In further embodiments, a therapeutically effective dose, when administered regularly, delays the onset of liver fibrosis in a treated subject. In some embodiments, a therapeutically effective dose, when administered regularly, slows the development of liver fibrosis or reduces the amount of liver fibrosis in a subject afflicted with GSD III.

[00279] A therapeutically effective dose of an active agent (*e.g.*, a translatable oligomer encoding AGL) *in vivo* can be a dose of about 0.001 to about 500 mg/kg body weight. For instance, the therapeutically effective dose may be about 0.001-0.01 mg/kg body weight, or 0.01-0.1 mg/kg, or 0.1-1 mg/kg, or 1-10 mg/kg, or 10-100 mg/kg. In some embodiments, a translatable oligomer encoding AGL is provided at a dose ranging from about 0.1 to about 10 mg/kg body weight, *e.g.*, from about 0.5 to about 5 mg/kg, from about 1 to about 4.5 mg/kg, or from about 2 to about 4 mg/kg.

[00280] A therapeutically effective dose of an active agent (*e.g.*, a translatable oligomer encoding AGL) *in vivo* can be a dose of at least about 0.001 mg/kg body weight, or at least about 0.01 mg/kg, or at least about 0.1 mg/kg, or at least about 1 mg/kg, or at least about 2 mg/kg, or at least about 3 mg/kg, or at least about 4 mg/kg, or at least about 5 mg/kg, at least

about 10 mg/kg, at least about 20 mg/kg, at least about 50 mg/kg, or more. In some embodiments, a translatable oligomer encoding AGL is provided at a dose of about 0.1 mg/kg, about 0.5 mg/kg, about 1 mg/kg, about 1.5 mg/kg, about 2 mg/kg, about 2.5 mg/kg, about 3 mg/kg, about 3.5 mg/kg, about 4 mg/kg, about 5 mg/kg, or about 6, 7, 8, 9, 10, 15, 20, 25, 50, 75, or 100 mg/kg.

[00281] Nucleobase sequences shown herein are from left to right, 5' to 3', unless stated otherwise.

Transfections

[00282] In some experiments, translatable messenger molecules were transfected into Hepa1-6 or AML12 cells in 96 well plates. The MessengerMAX transfection reagent (Thermo Fisher Scientific) was used by manufacture instruction for all transfections. Other suitable cell lines include HEK293 and Hep3B cells.

[00283] An example transfection protocol *in vitro* was as follows:

[00284] Plate hepatocyte Hepa1-6 cells 5000 cells per well in 96 well plate at least 8 hours before transfection.

[00285] Replace 90 μ L DMEM medium containing 10% FBS and Non-essential amino acid) adding 90 μ L into each well of 96 well plate immediately before beginning the transfection experiment.

[00286] Prepare MessengerMAX transfection reagent (Thermo Fisher Scientific) translatable molecule complex according to manufacturer's instruction.

[00287] Transfer 10 μ L of the complex into a well containing the cells in the 96-well plate.

[00288] Collect the medium after desired time points and add 100 μ L fresh medium into each well. Medium will be kept at -80°C until an ELISA assay for AGL is performed using the standard manufacturer protocol.

[00289] An example of a transfection protocol *in vivo* was as follows:

[00290] The translatable molecule is formulated with nanoparticles.

[00291] Inject the nanoparticle-formulated translatable molecule (1 mg/kg) into BL57BL/c mice (4-6 week-old) via standard i.v. injection in the lateral tail vein.

[00292] Collect approximately 50 μ L of blood in a Heparin-coated microcentrifuge tube at a suitable time post-injection.

[00293] Centrifuge at 3,000 X g for 10 minutes at 4°C.

[00294] Transfer the supernatant (plasma) into a fresh microcentrifuge tube. Plasma will be kept at -80°C until an ELISA assay for AGL is performed using the standard manufacturer protocol.

Nanoparticle Formulations

[00295] Lipid nanoparticles can be prepared containing an mRNA, using appropriate volumes of lipids in an ethanol/aqueous buffer containing the mRNA. A Nanossemblr microfluidic device can be used for this purpose, followed by downstream processing. For example, to prepare nanoparticles, a desired amount of targeted mRNA can be dissolved into 5 mM Citric Acid buffer (pH 3.5). The lipids can be dissolved at the adequate molar ratio, in ethanol. The molar percentage ratio for the constituent lipids can be, for example, 50% ionizable lipid, 7% DSPC (1,2-distearoyl-sn-glycero-3-phosphocholine; Avanti Polar Lipids), 40% cholesterol (Avanti Polar Lipids), and 3% DMG-PEG (1,2-Dimyristoyl-sn-glycerol, methoxypolyethylene glycol, PEG chain molecular weight: 2000; NOF America Corporation). Next, the lipid and mRNA solutions can be combined in the microfluidic device (Precision NanoSystems) at a flow ratio of 1:3 (ethanol:aqueous phase). The total combined flow rate can be 12 mL/min. Lipid nanoparticles can be formed and subsequently purified by overnight dialysis using a phosphate buffer in a dialysis device (Float-a-lyzer, Spectrum Labs), followed by concentration using Amicon Ultra-15 centrifugal filters (Merck Millipore). The particle size can be determined by dynamic light scattering (ZEN3600, Malvern Instruments). An “encapsulation” efficiency can be calculated by determining the un-encapsulated mRNA content measured by the fluorescence upon the addition of RiboGreen (Molecular Probes) to the LNP slurry (F_i); then, the value was compared to the total mRNA content that is obtained upon lysis of the LNPs by 1% Triton X-100 (F_t), where percentage of “encapsulation” = $(F_t - F_i)/F_t \times 100$. Encapsulation can refer to inclusion of the mRNA in the nanoparticle, regardless of form.

In-Cell Western

[00296] 96-well collagen plates were used to seed the cells at the appropriate density in DMEM/FBS culture media. At the optimal confluence, cells were transfected with the targeted mRNAs diluted in the transfection reagent mix (MessengerMax and Opti-MEM). Cells were placed in the CO₂ incubator and let them grow. At the desire timepoint, media was removed and cells were fixed in 4% fresh PFA for 20min. After that, fixative was removed and cells were permeabilized in TBST for 5 minutes several times. When permeabilization washes are complete, cells were incubated with the blocking buffer for 45 min. Primary antibody was then added and incubated for 1h at room temperature. Following that, cells were washed several times in TBST, and then incubated for 1h with the secondary antibody diluted in blocking buffer and containing the CellTag 700 stain. To finalize, cells were washed several times in TBST followed by a last wash in TBS. Then, plate was imaged using the Licor detection system and data was normalized to the total number of cells labeled by the CellTag 700.

Generating Tail PCR Products

[00297] Plasmid DNA (10 ng) containing each mRNA expression construct can be used to generate the poly A tail 120 PCR products in a 50 µl PCR reaction with 2X KAPA HiFi PCR mix (KR0370) as per the manufacturer's instructions. The product can be then checked on a 2% gel from Thermo Fisher Scientific and approximately quantified based on the intensity of the low molecular weight ladder (Thermo Fisher Scientific, 10068-013), and cleaned with the Qiagen PCR purification kit and resuspended in 50ul water.

In vitro Transcription (IVT) for Synthesis

[00298] The following protocol is for a 200 µl IVT reaction using NEB HiScribe T7 RNA polymerase reagents, which should yield about 1 mg of RNA. 2.5X NTP mix was prepared as required by thawing individual 100mM NTP stocks (ATP, GTP, CTP, and UTP nucleotides, or chemically modified counterparts) and pooling them together. For the IVT reaction, about 2-4 µg of the template was used for a 200 µl reaction. The 10X IVT reaction buffer, the 2.5X dNTP mix, the template DNA and the T7 RNA polymerase are mixed well by pipetting and incubated at 37°C for 4 hours. To degrade the DNA template, the IVT reaction

is diluted with 700ul of nuclease-free water and then 10X DNase I buffer and 20ul of the RNase-free DNase I are added to the IVT mix and incubated at 37°C for 15 minutes. The diluted (to 1 ml) and DNase treated reaction is then purified by a Qiagen RNeasy Maxi columns as per the manufacturer's instructions with a final elution in RNase-free water. The purified RNA is then quantified by UV absorbance where the A260/A280 should be about 1.8-2.2, depending on the resuspension buffer used.

Enzymatic Capping of IVT RNA

[00299] For enzymatic capping, a 50X scaled-up version of NEB's one-step capping and 2'-O-methylation reaction can be used, that is suitable for treating up to 1mg of IVT transcripts. A 10 µg RNA in a 20 µl reaction is recommended, based on the assumption that transcript length would be as short as 100 nt. However, a higher substrate-to-reaction volume is acceptable for transcripts, which can be generally longer (about 300-600 nt) in length. Before initiating the capping reaction, the RNA is denatured at 65°C for 5 minutes and then snap chilled to relieve any secondary conformations. For the total 1 ml capping reaction, 1 mg denatured RNA in 700 µl of nuclease-free water is used along with 100 µl (10X) capping buffer, 50 µl (10 mM) GTP, 50 µl (4 mM) SAM, 50 µl of (10 U/µl) Vaccinia capping enzyme and 50 µl of mRNA cap 2'-O-methyltransferase at (50 U/µl) are combined and incubated at 37°C for 1 hour. The resulting capped mRNA is eluted using RNase free water, re-purified on an RNeasy column, quantified by nanodrop. The mRNA is also visualized on the gel by running 500 ng of the purified product per lane in a denaturing gel after denaturation and snap-chill to remove secondary structures.

EXAMPLES

[00300] Example 1: Reference translatable molecule 534.

[00301] In this example, a reference translatable molecule 534 was made and used for expressing human WT amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL). The translatable molecule comprised a 5' cap (7mGpppG), a 5' UTR of TEV, a Kozak sequence, a WT AGL CDS (SEQ ID NO: 1), a 3'UTR of Xenopus beta-globin, and a Poly(A) tail region consisting of 114 As (*i.e.*, "Poly(A) 114 tail region"). The reference translatable molecule further comprised the sequence of SEQ ID NO: 40 immediately downstream of the AGL CDS.

This reference translatable molecule was synthesized with N¹-methylpseudouridine in place of uridine.

[00302] Details of the structure of this reference translatable molecule are as follows: Tobacco Etch Virus (TEV) 5' UTR of SEQ ID NO: 3, a Kozak Sequence of SEQ ID NO: 4, a Xenopus beta-globin (XBG) 3' UTR of SEQ ID NO: 5, and a Poly(A) 114 Tail of SEQ ID NO: 6.

[00303] Translatable molecules in the examples below can be synthesized with the 5' cap being a m7GpppGm cap. The translatable molecules in the examples below can contain a 5'-UTR (*e.g.*, a 5' UTR of TEV (SEQ ID NO: 3)), a translation initiation sequence (*e.g.*, a Kozak sequence of SEQ ID NO: 4), a sequence of SEQ ID NO: 40, a 3' UTR (*e.g.*, a 3' UTR of Xenopus beta-globin (SEQ ID NO: 5)), and a poly(A) tail (*e.g.*, a polyA tail of SEQ ID NO: 6, SEQ ID NO: 38, or SEQ ID NO: 39).

[00304] Example 2: Translatable molecules encoding AGL.

[00305] In this example, translatable molecules 522-533, 546, 730-740, and 1783-1784 were made and used for expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) with advantageously increased efficiency of translation. These translatable molecules expressing human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (AGL) exhibited activity suitable for use in methods for ameliorating or treating GSD III. These translatable molecules comprised a 5' cap (7mGpppG), a 5' UTR of TEV, a Kozak sequence, an AGL CDS, and a 3'UTR of Xenopus beta-globin. Translatable molecules 522-533, 546, and 1783-1784 further comprise a Poly(A) 114 tail region. Translatable molecule 730 further comprises a Poly(A) 100 tail region, while translatable molecules 731-740 further comprise a Poly(A) 110 tail region. Translatable molecules 522-533, 546, 730-740, and 1783-1784 further comprise the sequence of SEQ ID NO: 40 immediately downstream of the AGL CDS. Two additional translatable molecules – 2258 and 2259 – were developed that are identical to 546 and 1783, respectively, except that they contain a Poly(A) 100 tail region as opposed to a Poly(A) 114 tail region. The coding sequence of 1783 and 2259 may optionally be modified to contain 12 nucleotide differences, as reflected in SEQ ID NO: 45. The translatable molecules described in this example were synthesized with N¹-methylpseudouridine in place of uridine

[00306] The AGL CDS in each of the translatable molecules is comprised of the following sequences:

Molecule	AGL CDS
522	SEQ ID NO: 7
523	SEQ ID NO: 8
524	SEQ ID NO: 9
525	SEQ ID NO: 10
526	SEQ ID NO: 11
527	SEQ ID NO: 12
528	SEQ ID NO: 13
529	SEQ ID NO: 14
530	SEQ ID NO: 15
531	SEQ ID NO: 16
532	SEQ ID NO: 17
533	SEQ ID NO: 18
546 & 2258	SEQ ID NO: 19
730	SEQ ID NO: 20
731	SEQ ID NO: 21
732	SEQ ID NO: 22
733	SEQ ID NO: 23
734	SEQ ID NO: 24
735	SEQ ID NO: 25
736	SEQ ID NO: 26
737	SEQ ID NO: 27
738	SEQ ID NO: 28
739	SEQ ID NO: 29
740	SEQ ID NO: 30
1783 & 2259	SEQ ID NO: 31
1784	SEQ ID NO: 32

[00307] The translatable molecules of this example were translated in AML12 and C2C12 cells to produce human AGL.

[00308] Example 3: Translation enhancer based on *Xenopus* beta-globin 3'UTR.

[00309] In this example, the structures of 3' UTR sequences for use in enhancing translational efficiency of a translatable molecule are shown.

[00310] The base sequences shown in SEQ ID NOs: 33-37 are the portion of the translatable molecule that may correspond in functionality to the 3'-UTR of *Xenopus* beta-globin. The complete translatable molecule comprises a 5' cap (m7GpppGm), 5'-UTR, and coding region (CDS) upstream of the sequence below, and a polyA tail downstream of the sequence below, each of which corresponds to the structure of a native human mRNA. As shown above, a Kozak sequence may optionally be used. Thus, a translatable molecule incorporating the fragment below can have enhanced translational efficiency. The *Xenopus* beta-globin gene sequence is shown in accession no. NM_001096347.1

[00311] Example 4: AGL Expression in Human Primary Hepatocytes

[00312] In this example, human primary hepatocytes were transfected with codon-optimized mRNA molecules 546, 1783, and 1784. AGL protein expression was measured by In-Cell Western™ at 6, 24, 48, and 72 hours post-transfection. The expression of the mRNA sequences as compared with an untreated control (“unt”) is shown in FIG. 5.

[00313] Example 5: In vivo Analysis of AGL Protein Expression in Wild-Type Mice

[00314] In this example, wild-type C57BL/6 mice were injected with human AGL mRNA formulated with lipid nanoparticles. Mice were sacrificed 6 hours post-injection. Liver biopsy samples were taken from mice, and human and mouse AGL protein expression in liver homogenates was analyzed. FIG 6. shows ectopic expression of human AGL protein from various mRNA molecules. FIG 7. shows mouse AGL protein levels, indicating similar levels of endogenous expression of mouse AGL protein across the treated mice. Translatable molecule 546.7 – as shown in FIGs. 6 and 7 – has an identical nucleobase sequence to translatable molecule 546, but was synthesized with 5-methoxyuridine in place of uridine instead of N¹-methylpseudouridine.

[00315] Example 6: mRNA Treatment Reduces Glycogen Accumulation in GSD3 Mice

[00316] In this example, AGL knockout mice were treated with vehicle or translatable molecule 546 formulated with ATX2 lipid nanoparticles. Livers from knockout mice treated with vehicle (“VEH”) showed marked to severe vacuolation of hepatocytes and moderate to marked increases in glycogen accumulation within hepatocytes (FIG. 8). In contrast, livers from knockout mice treated with translatable molecule 546 formulated with ATX2 lipid nanoparticles had only mild to moderate vacuolation of hepatocytes and only mild to moderate increases in glycogen accumulation within hepatocytes (FIG. 9). Based upon the histopathology results shown in this example, there appears to be a reduction in the severity of hepatocellular vacuolization and glycogen accumulation in livers from knockout mice treated with mRNA compared with KO mice treated with vehicle.

[00317] Example 7: Additional Translatable Molecules Expressing AGL

[00318] Four additional translatable molecules – 1970, 1987, SD1, and SD2 – were designed with additionally modified codon-optimized human AGL coding sequences shown in SEQ ID NOs: 41, 42, 43, and 44, respectively. Translatable molecules 1970, 1987, SD1, and SD2 further comprise a 5' cap (7mGpppG), a 5' UTR of TEV, a Kozak sequence, a sequence

of SEQ ID NO: 40 immediately downstream of the coding sequence, a 3' UTR of *Xenopus* beta-globin, and a Poly(A) tail region (*e.g.*, a poly(A) 100, 110, or 114 tail region). These translatable molecules may optionally be synthesized with either N¹-methylpseudouridine or 5-methoxyuridine in place of uridine.

[00319] All publications, patents and literature specifically mentioned herein are incorporated by reference for all purposes.

[00320] It is understood that this invention is not limited to the particular methodology, protocols, materials, and reagents described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which will be encompassed by the appended claims.

[00321] It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. As well, the terms "a" (or "an"), "one or more" and "at least one" can be used interchangeably herein. It is also to be noted that the terms "comprises," "comprising", "containing," "including", and "having" can be used interchangeably.

[00322] Without further elaboration, it is believed that one skilled in the art can, based on the above description, utilize the present invention to its fullest extent. The following specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

[00323] All of the features disclosed in this specification may be combined in any combination. Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose.

WHAT IS CLAIMED IS:

1. A polynucleotide comprising a nucleobase sequence encoding a human amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase (hAGL) of SEQ ID NO: 2, wherein the nucleobase sequence encoding hAGL is at least 99% identical to a sequence selected from SEQ ID NOs: 7-32 and SEQ ID NOs: 41-45.
2. The polynucleotide of claim 1, wherein the nucleobase sequence encoding hAGL is at least 99% identical to SEQ ID NO: 41.
3. The polynucleotide of claim 1 or 2, wherein at least one uridine residue is replaced with an N¹-methylpseudouridine residue.
4. The polynucleotide of any one of claims 1-3, wherein all uridine residues are replaced with N¹-methylpseudouridine residues.
5. The polynucleotide of any one of claims 1-4, wherein the polynucleotide comprises a 5' cap.
6. The polynucleotide of any one of claims 1-5, wherein the polynucleotide comprises a 5' untranslated region (5' UTR).
7. The polynucleotide of claim 6, wherein the 5' UTR is derived from a tobacco etch virus (TEV).
8. The polynucleotide of claim 6 or 7, wherein the 5' UTR comprises SEQ ID NO: 3.
9. The polynucleotide of any one of claims 1-8, wherein the polynucleotide comprises a 3' untranslated region (3' UTR).
10. The polynucleotide of claim 9, wherein the 3' UTR is derived from a Xenopus beta globin.
11. The polynucleotide of claim 9 or 10, wherein the 3' UTR comprises SEQ ID NO: 5.

12. The polynucleotide of any one of claims 1-11, wherein the polynucleotide comprises a 3' polyA tail.
13. The polynucleotide of claim 12, wherein the 3' polyA tail is 60 to 220 adenosine nucleotides in length.
14. The polynucleotide of claim 13, wherein the 3' polyA tail is about 100 nucleotides in length.
15. A composition comprising one or more polynucleotides of any one of claims 1-14 and a pharmaceutically acceptable carrier.
16. The composition of claim 15, wherein the carrier is selected from a transfection reagent, a nanoparticle, and a liposome.
17. The composition of claim 16, wherein the carrier is a nanoparticle.
18. The composition of claim 17, wherein the nanoparticle comprises a cationic lipid selected from ATX-002, ATX-081, ATX-095, and ATX-126.
19. A method for ameliorating, preventing, delaying onset, or treating a disease or disorder associated with reduced activity of hAGL in a subject in need thereof, the method comprising administering to the subject the composition of any one of claims 15-18.
20. The method of claim 19, wherein the disease is glycogen storage disease type III.
21. The method of any one of claims 19-20, wherein the administration is intravenous, subcutaneous, pulmonary, intramuscular, intraperitoneal, dermal, oral, nasal, or inhalation.
22. The method of any one of claims 19-21, wherein the administration is once daily, weekly, biweekly, or monthly.
23. The method of any one of claims 19-22, wherein the administration comprises an effective dose of from 0.01 to 10 mg/kg.

24. The method of claim 23, wherein the administration comprises an effective dose of 0.1 mg/kg.
25. A polynucleotide comprising a nucleobase sequence of SEQ ID NO: 19
26. A polynucleotide comprising a nucleobase sequence of SEQ ID NO: 31.
27. A polynucleotide comprising a nucleobase sequence of SEQ ID NO: 45.
28. Use of the polynucleotide of any one of claims 1-14 or 25-27, or the composition of any one of claims 15-18 in the manufacture of a medicament for ameliorating, preventing, delaying onset, or treating a disease or disorder associated with reduced activity of hAGL in a subject in need thereof.

FIG. 1

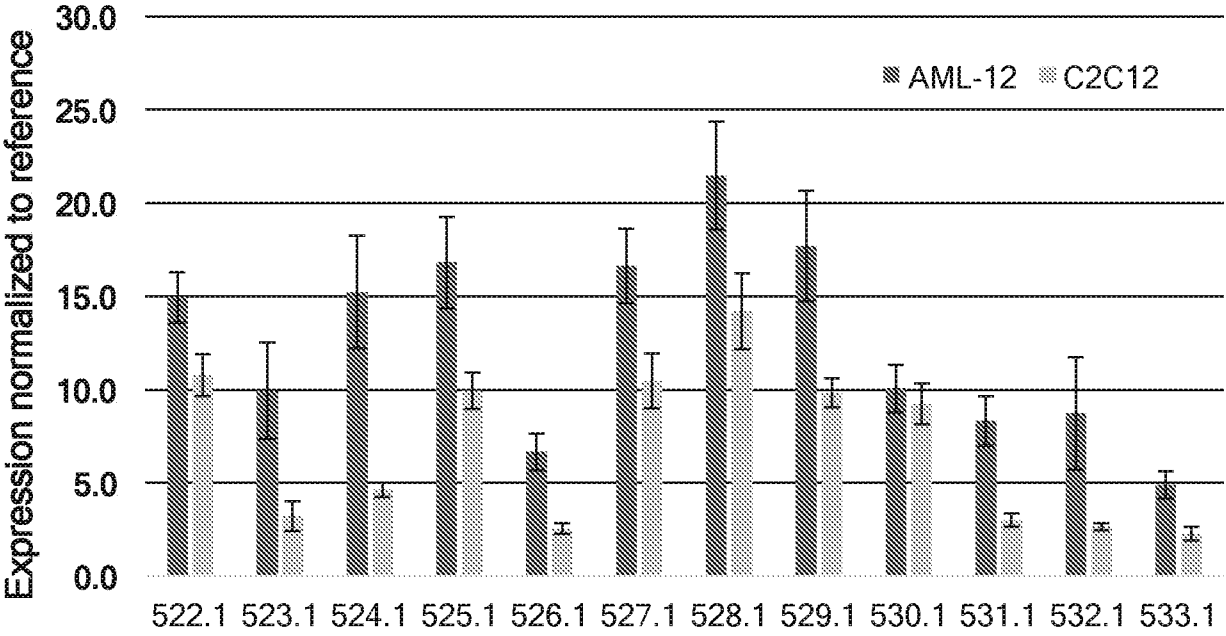


FIG. 2

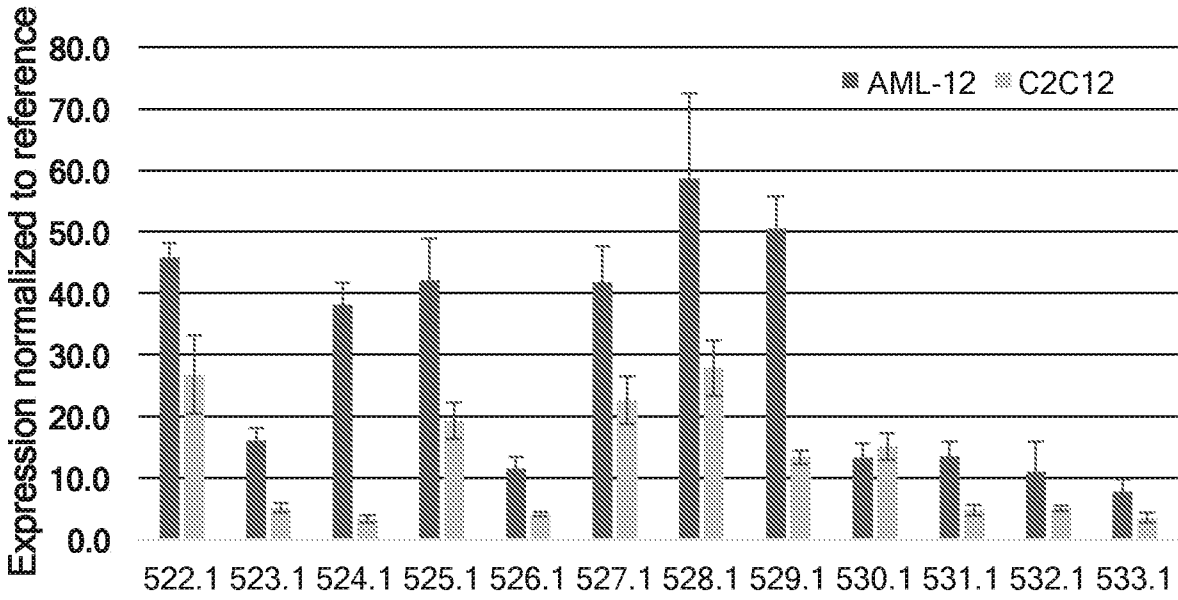


Fig. 3 AGL expression in WT mice

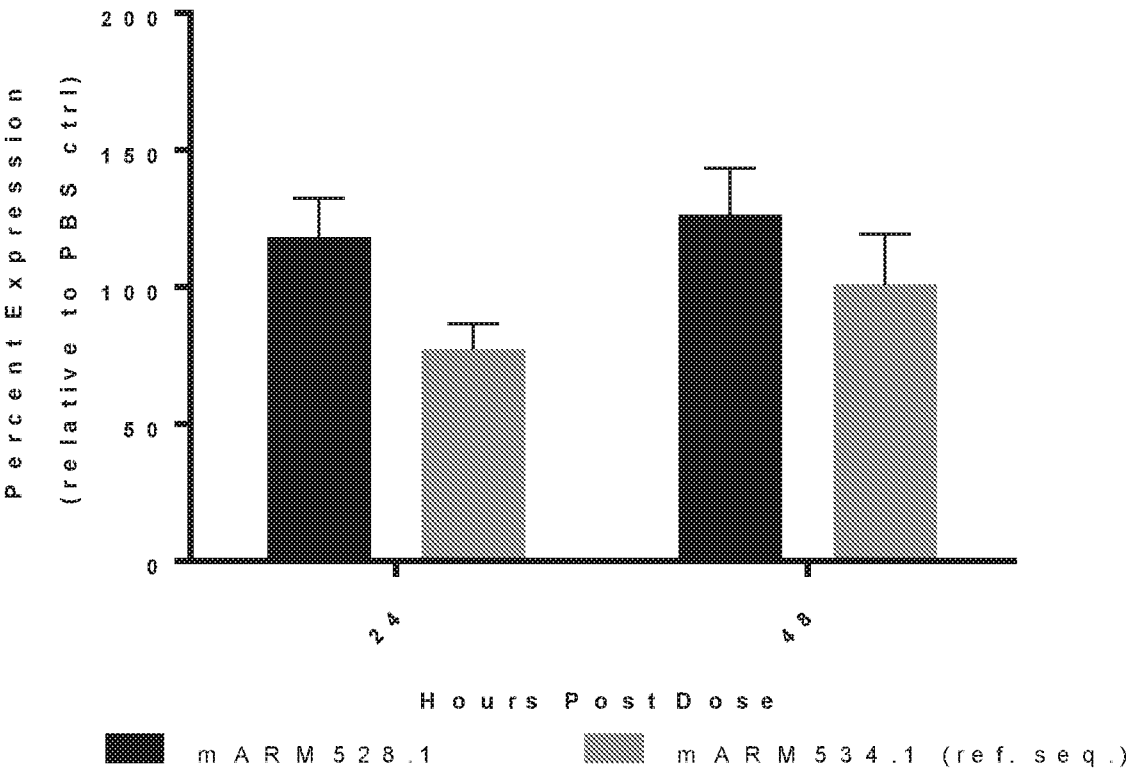


Fig. 4 AGL expression in WT mouse liver

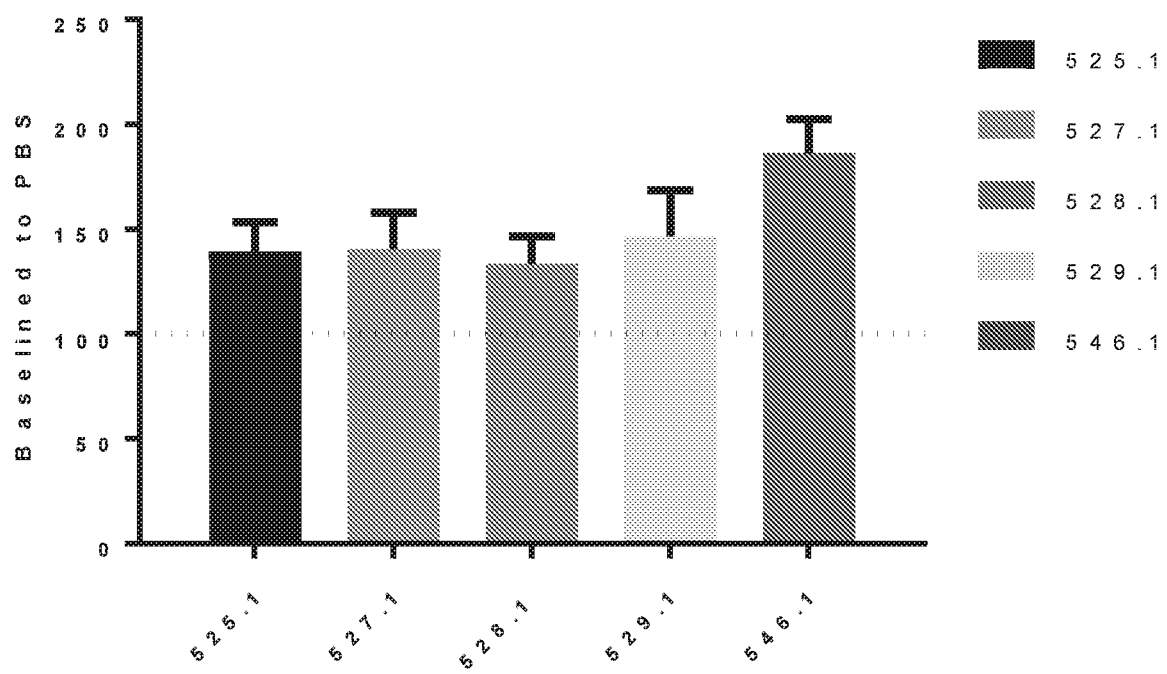


FIG. 5

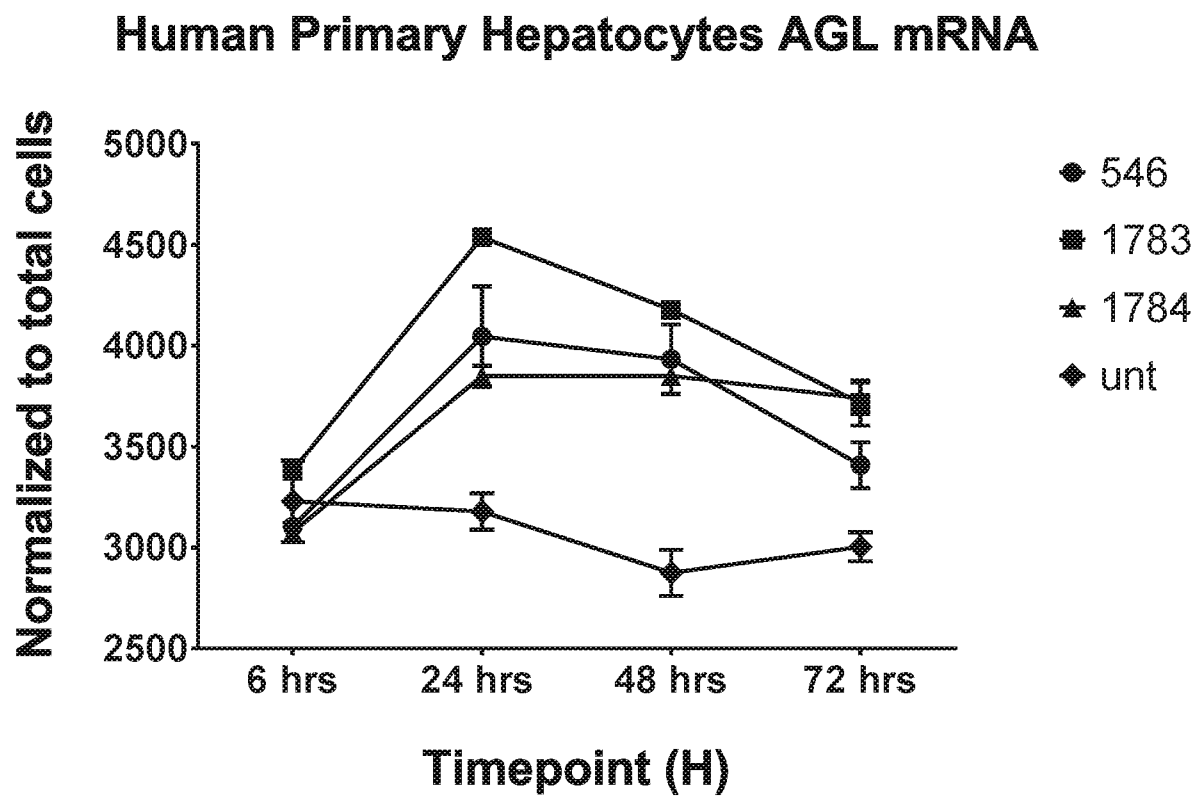


FIG. 6

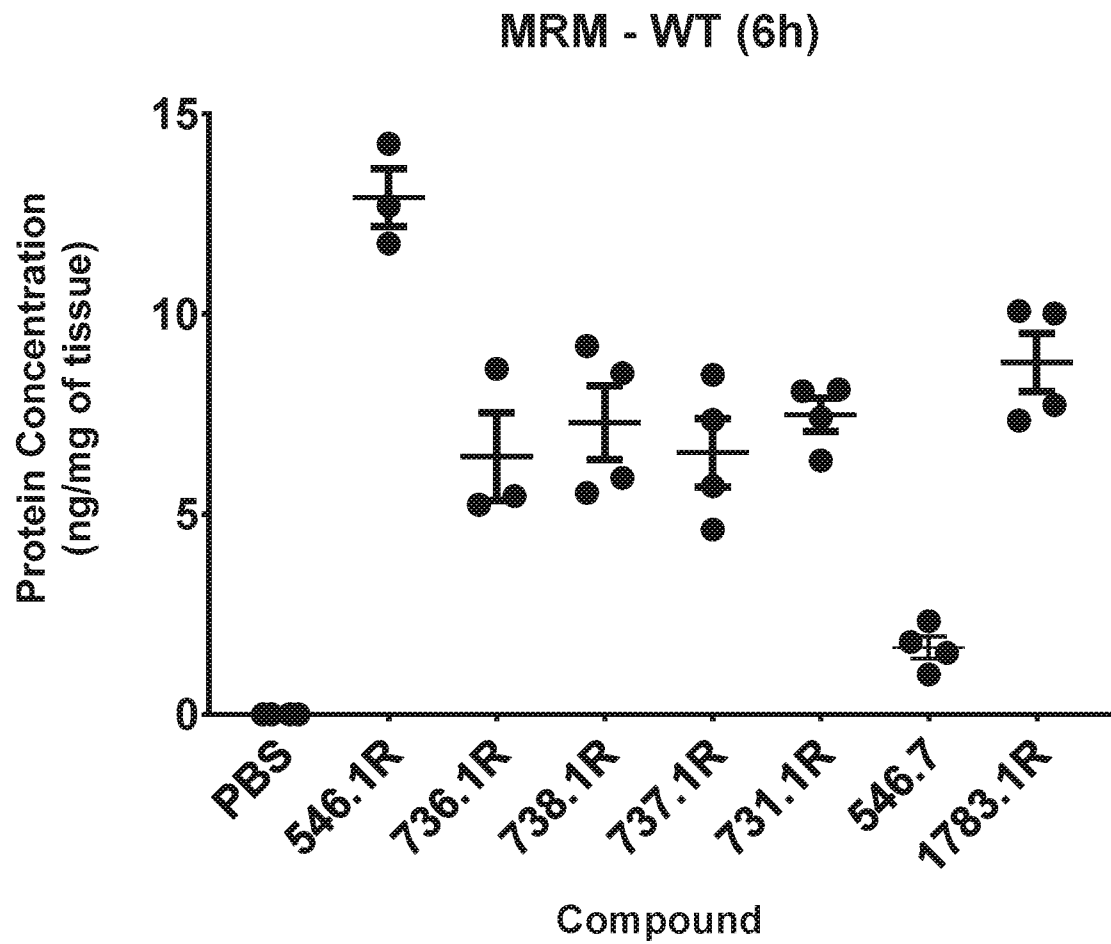


FIG. 7

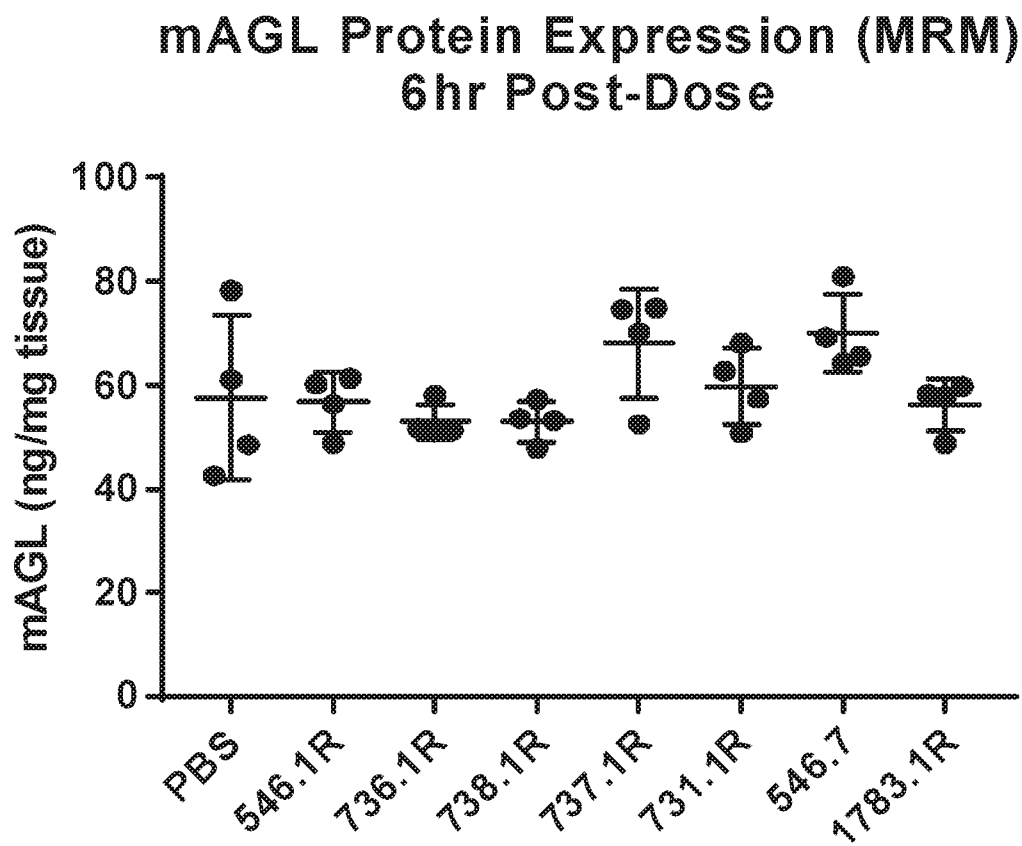


FIG. 8

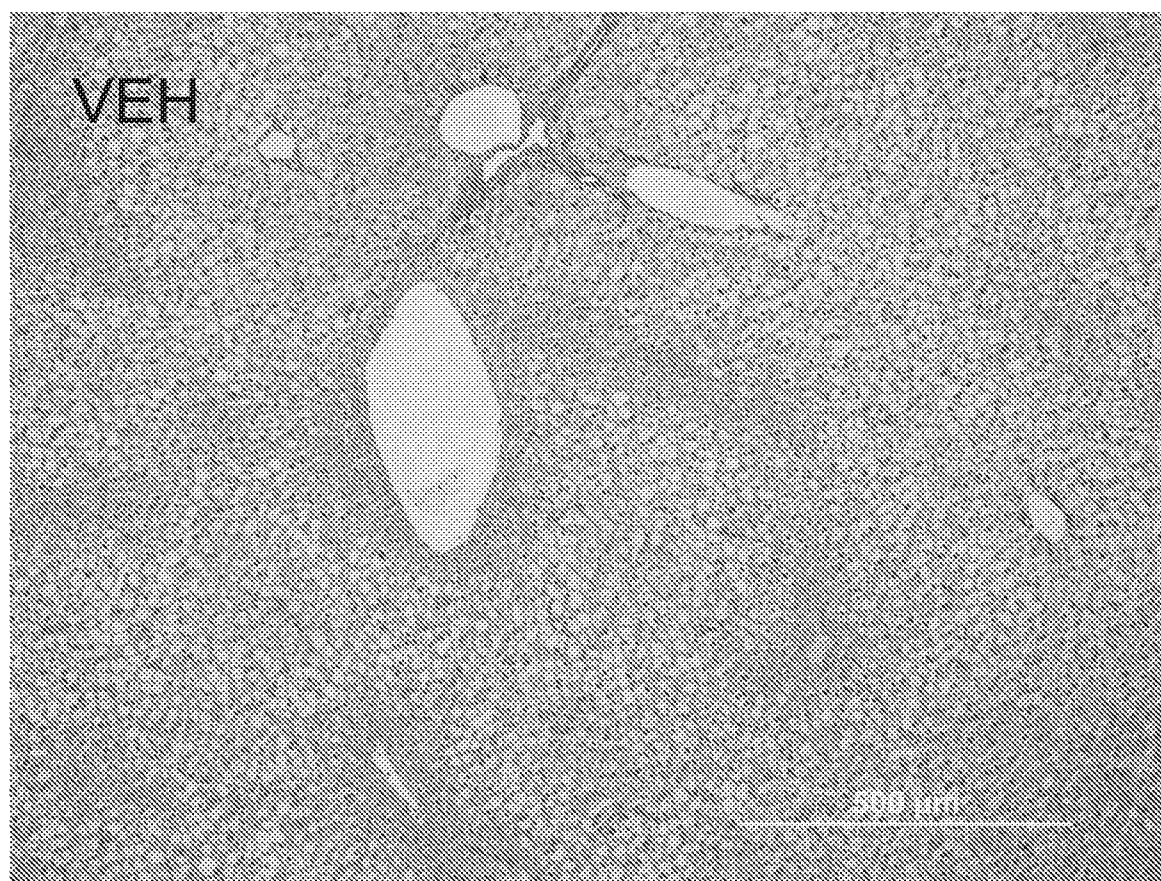
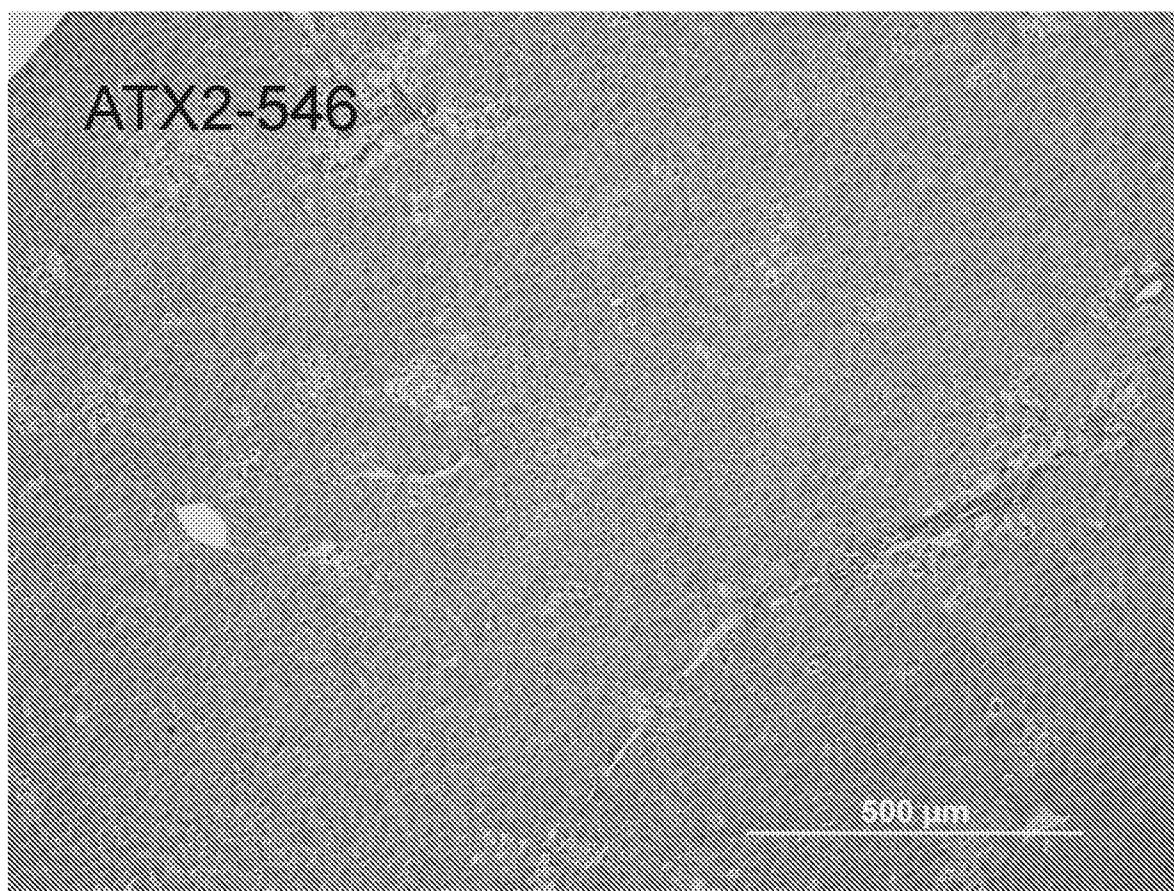


FIG. 9



ULPI_041_01WO_SeqList_ST25.txt
SEQUENCE LISTING

<110> Ultragenyx Pharmaceutical Inc.
Tachikawa, Kiyoshi
Perez-Garcia, Carlos Gustavo
Chivukula, Padmanabh
Bhaskaran, Hari Prakash
Cobaugh, Christian W.
Daugherty, Sean Christopher

<120> THERAPEUTICS FOR GLYCOGEN STORAGE DISEASE TYPE III

<130> ULPI-041/01WO 315613-2474

<150> US 62/513,350

<151> 2017-05-31

<160> 45

<170> PatentIn version 3.5

<210> 1

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Consensus Human AGL Coding Sequence

<400> 1

augggacaca guaaacagau ucgaauuuua cuucugaacg aaauggagaa acuggaaaag	60
acccucuuca gacuugaaca aggguaugag cuacaguucc gauuaggccc aacuuuacag	120
ggaaaagcag uuaccgugua uacaaaauac ccauuuccug gagaaacauu uaauagagaa	180
aaauuccguu cucuggauug ggaaaaucca acagaaagag aagaugauuc ugauaaaauac	240
uguaaaacua aucugcaaca aucugguuca uuucaguauu auuuccuua aggaaaugag	300
aaaaguggug gagguuacau aguuguggac ccacuuuuac guguuggugc ugauaaucan	360
gugcuacccu uggacugugu uacucuucag acauuuuuag cuaaguguuu gggaccuuuu	420
gaugaauggg aaagcagacu uagggguugca aaagaucag gcuacaacau gauucauuuu	480
accccauugc agacucuugg acuaucuagg ucaugcuacu cccuugccaa ucaguugaa	540

ULPI_041_01WO_SeqList_ST25.txt

uuaaauccug acuuuucaag accuaauaga aaguauaccu ggaaugaugu uggacagcua	600
guggaaaaau uaaaaaagga auggaauuu auuuguauua cugauguugu cuacaaucau	660
acugcugcua auaguaaaug gauccaggaa cauccagaau gugccuauaa ucuugugaau	720
ucuccacacu uaaaaccugc cugggucuaa gacagagcac uuuggcguuu cuccugugau	780
guugcagaag ggaaauacaa agaaaaggga auaccugcuu ugauugaaaa ugauccacau	840
augaauucca uccgaaaaau aauuugggag gauauuuuuc caaagcuuaa acucugggaa	900
uuuuuccaag uagaugucua caaagcgguu gagcauuua gaagacuucu uacacaagaa	960
aaauaggcgag uaaccaaguc ugauccaaac caacaccuaa cgauuauuca agauccugaa	1020
uacagacggu uuggcuguac uguagauaug aacauugcac uaacgacuuu cauaccacau	1080
gacaaggggc cagcagcaau ugaagaugc uguaaauuggu uucauaaaag aauggaggaa	1140
uuaaaauucag agaagcaucg acucauuuac uaucaucagg aacaggcagu uaaaugccuu	1200
uugggaaaug uguuuuaua acgacuggcu ggccaugguc caaaacuagg accugucacu	1260
agaaagcauc cuuuaguua cagguauuuu acuuucccau uugaagagau agacuucucc	1320
auggaagaau cuaugauua ucugccaaau aaagcuuguu uucugauggc acacaaugga	1380
uggguaaugg gagaugauc ucuucgaaac uuugcugaac cggguucaga aguuuaccua	1440
aggagagaac uuauuugcug gggagacagu guuaaaauac gcuaugggaa uaaaccagag	1500
gacuguccuu aucucugggc acacaugaaa aaauacacug aaauaacugc aacuuauuuc	1560
cagggaguac gucuugauaa cugccacuca acaccucuuc acguagcuga guacauguug	1620
gaugcugcua ggaauuugca acccauuua uauguaguag cugaacuguu cacaggaagu	1680
gaagaucugg acaaugucuu uguuacuaga cugggcauaa guuccuuaa aagagaggca	1740
augagugcau auaauaguca ugaagagggc agauuaguuu accgauaugg aggagaaccu	1800
guuggauccu uuguucagcc cuguuugagg ccuuuaaugc cagcuauugc acaugcccug	1860
uuuauggaua uuacgcauga uaaugagugu ccuaauuguc auagaucagc guaugaugcu	1920
cuaccaagua cuacaauugu uucuauggca uguugugcua guggaaguac aagaggcuau	1980

ULPI_041_01W0_SeqList_ST25.txt

gaugaauuag ugccucauca gauuucagug guuucugaag aacgguuuuu cacuaagugg	2040
aauccugaag cauugccuuc aaacacaggu gaaguuaauu uccaaagcgg cauuaugca	2100
gccaggugug cuaucaguaa acuucaucag gagcuuggag ccaaggguuu uauucaggug	2160
uauugggau c aaguugauga agacauagug gcaguaacaa gacacucacc uagcauccau	2220
cagucuguug uggcuguau c uagaacugcu uucaggaau c ccaagacuuc auuuuacagc	2280
aaggaagugc cucaaauug c cauccuggc aaaaauugaag aaguaguucu ugaagcuaga	2340
acuauugaga gaaacacgaa accuuauagg aaggaugaga auucaaucaa uggaacacca	2400
gauaucacag uagaaauuag agaacauau cagcuuaug aaaguaaaau uguuaaacia	2460
gcuggaguug ccacaaaagg gcccaaugaa uauauucaag aaauagaauu ugaaaacuug	2520
ucuccaggaa guguuauau auucagagu agucuugauc cacaugcaca agucgcuguu	2580
ggaauucuuc gaaaucaucu gacacaauc aguccucacu uuaaaucugg cagccuagcu	2640
guugacaaug cagauccuau auuaaaaau cccuuugcuu cucuugccuc cagauuaacu	2700
uuggcugagc uaaaucagau cccuuaccga ugugaucag aagaaaagga agauggugga	2760
gggugcuau acauaccaa cuggucagcc cuuaaaauug caggucuua agguuuauug	2820
ucuguauugg cagaaauaag accaaagaau gacuuggggc auccuuuuug uaauauuuug	2880
agaucuggag auuggaugau ugacuauugc aguaaccggc uuauuuacg aucaggaacu	2940
auugcugaag uugguauuug guugcaggcu auguucuuc accugaagca gaucacacgu	3000
uaccuuaucc cauguuacuu ugaugcuaua uuaauuggug cauauaccac ucuucuggau	3060
acagcaugga agcagauguc aagcuuuguu cagaauuguu caaccuuugu gaaacaccuu	3120
ucauuggguu caguucaacu guguggagua ggaaaauucc cuucccugcc aaauuuuua	3180
ccugcccuaa uggauguacc uuauagguua aaugagauca caaaagaaaa ggagcaaugu	3240
uguguuuuc uagcugcagg cuuaccucau uuuucuucug guauuuuccg cugcugggga	3300
agggauacuu uuauugcacu uagagguaua cugcugauua cuggacgcu uguagaagcc	3360
aggaauauua uuuuagcauu ugcggguacc cugaggcaug gacucauucc uaaucuacug	3420

ULPI_041_01W0_SeqList_ST25.txt

ggugaaggaa uuuaugccag auacaauugu cgggaugcug ugugguggug gcugcagugu	3480
auccaggauu acuguaaaau gguuccaaau ggucuagaca uucucaagug cccaguuucc	3540
agaauguau cuacagauga uucugcuccu uugccugcug gcacacugga ucagccauug	3600
uuugaaguca uacaggaagc aaugcaaaaa cacaugcagg gcacacaguu ccgagaaagg	3660
aaugcugguc cccagauaga ucgaaacaug aaggacgaag guuuuaauau aacugcagga	3720
guugaugaag aaacaggauu uguuuauugga ggaaaucguu ucaauugugg cacauggaug	3780
gauaaaaugg gagaaaguga cagagcuaga aacagaggaa ucccagccac accaagagau	3840
gggucugcug uggaaaauugu gggccugagu aaucugcug uucgcugguu gcuggaauua	3900
uccaaaaaaaa auauuuuccc uuaucauga gucacagua aaagacaugg aaaggcuaua	3960
aagguaucau augaugagug gaacagaaaa auacaagaca acuuugaaaa gcuauuucan	4020
guuuccgaag acccuucaga uuuaaaugaa aagcauccaa aucugguuca caaacguggc	4080
auauacaaaag auaguuaugg agcuucaagu ccuuggugug acuaucagcu caggccuaau	4140
uuuaccuag caaugguugu ggccccugag cucuuuacua cagaaaaagc auggaaagcu	4200
uuggagauug cagaaaaaaaa auugcuuggu ccccuuggca ugaaaacuuu agauccagau	4260
gauaugguuu acuguggaau uuaugacaau gcauuagaca augacaacua caaucuugcu	4320
aaagguuuca auuauaccca aggaccugag uggcuguggc cuauugggua uuuucucgu	4380
gcaaaaauuu auuuuuuccag auugaugggc ccggagacua cugcaaagac uauaguuug	4440
guuaaaaaag uucuuucccg acauuanguu caucuugaga gaucccuug gaaaggacuu	4500
ccagaacuga ccaaugagaa ugcccaguac uguccuuca gcugugaaac acaagccugg	4560
ucaauugcua cuauucuuga gacacuuuau gauuuauag	4599

<210> 2

<211> 1532

<212> PRT

<213> Artificial Sequence

<220>

<223> Consensus Human AGL Amino Acid Sequence

ULPI_041_01W0_SeqList_ST25.txt

<400> 2

Met Gly His Ser Lys Gln Ile Arg Ile Leu Leu Leu Asn Glu Met Glu
1 5 10 15

Lys Leu Glu Lys Thr Leu Phe Arg Leu Glu Gln Gly Tyr Glu Leu Gln
20 25 30

Phe Arg Leu Gly Pro Thr Leu Gln Gly Lys Ala Val Thr Val Tyr Thr
35 40 45

Asn Tyr Pro Phe Pro Gly Glu Thr Phe Asn Arg Glu Lys Phe Arg Ser
50 55 60

Leu Asp Trp Glu Asn Pro Thr Glu Arg Glu Asp Asp Ser Asp Lys Tyr
65 70 75 80

Cys Lys Leu Asn Leu Gln Gln Ser Gly Ser Phe Gln Tyr Tyr Phe Leu
85 90 95

Gln Gly Asn Glu Lys Ser Gly Gly Gly Tyr Ile Val Val Asp Pro Ile
100 105 110

Leu Arg Val Gly Ala Asp Asn His Val Leu Pro Leu Asp Cys Val Thr
115 120 125

Leu Gln Thr Phe Leu Ala Lys Cys Leu Gly Pro Phe Asp Glu Trp Glu
130 135 140

Ser Arg Leu Arg Val Ala Lys Glu Ser Gly Tyr Asn Met Ile His Phe
145 150 155 160

Thr Pro Leu Gln Thr Leu Gly Leu Ser Arg Ser Cys Tyr Ser Leu Ala
165 170 175

Asn Gln Leu Glu Leu Asn Pro Asp Phe Ser Arg Pro Asn Arg Lys Tyr

ULPI_041_01W0_SeqList_ST25.txt

180

185

190

Thr Trp Asn Asp Val Gly Gln Leu Val Glu Lys Leu Lys Lys Glu Trp
195 200 205

Asn Val Ile Cys Ile Thr Asp Val Val Tyr Asn His Thr Ala Ala Asn
210 215 220

Ser Lys Trp Ile Gln Glu His Pro Glu Cys Ala Tyr Asn Leu Val Asn
225 230 235 240

Ser Pro His Leu Lys Pro Ala Trp Val Leu Asp Arg Ala Leu Trp Arg
245 250 255

Phe Ser Cys Asp Val Ala Glu Gly Lys Tyr Lys Glu Lys Gly Ile Pro
260 265 270

Ala Leu Ile Glu Asn Asp His His Met Asn Ser Ile Arg Lys Ile Ile
275 280 285

Trp Glu Asp Ile Phe Pro Lys Leu Lys Leu Trp Glu Phe Phe Gln Val
290 295 300

Asp Val Asn Lys Ala Val Glu Gln Phe Arg Arg Leu Leu Thr Gln Glu
305 310 315 320

Asn Arg Arg Val Thr Lys Ser Asp Pro Asn Gln His Leu Thr Ile Ile
325 330 335

Gln Asp Pro Glu Tyr Arg Arg Phe Gly Cys Thr Val Asp Met Asn Ile
340 345 350

Ala Leu Thr Thr Phe Ile Pro His Asp Lys Gly Pro Ala Ala Ile Glu
355 360 365

Glu Cys Cys Asn Trp Phe His Lys Arg Met Glu Glu Leu Asn Ser Glu

ULPI_041_01W0_SeqList_ST25.txt

370

375

380

Lys His Arg Leu Ile Asn Tyr His Gln Glu Gln Ala Val Asn Cys Leu
385 390 395 400

Leu Gly Asn Val Phe Tyr Glu Arg Leu Ala Gly His Gly Pro Lys Leu
405 410 415

Gly Pro Val Thr Arg Lys His Pro Leu Val Thr Arg Tyr Phe Thr Phe
420 425 430

Pro Phe Glu Glu Ile Asp Phe Ser Met Glu Glu Ser Met Ile His Leu
435 440 445

Pro Asn Lys Ala Cys Phe Leu Met Ala His Asn Gly Trp Val Met Gly
450 455 460

Asp Asp Pro Leu Arg Asn Phe Ala Glu Pro Gly Ser Glu Val Tyr Leu
465 470 475 480

Arg Arg Glu Leu Ile Cys Trp Gly Asp Ser Val Lys Leu Arg Tyr Gly
485 490 495

Asn Lys Pro Glu Asp Cys Pro Tyr Leu Trp Ala His Met Lys Lys Tyr
500 505 510

Thr Glu Ile Thr Ala Thr Tyr Phe Gln Gly Val Arg Leu Asp Asn Cys
515 520 525

His Ser Thr Pro Leu His Val Ala Glu Tyr Met Leu Asp Ala Ala Arg
530 535 540

Asn Leu Gln Pro Asn Leu Tyr Val Val Ala Glu Leu Phe Thr Gly Ser
545 550 555 560

Glu Asp Leu Asp Asn Val Phe Val Thr Arg Leu Gly Ile Ser Ser Leu

ULPI_041_01W0_SeqList_ST25.txt

565

570

575

Ile Arg Glu Ala Met Ser Ala Tyr Asn Ser His Glu Glu Gly Arg Leu
580 585 590

Val Tyr Arg Tyr Gly Gly Glu Pro Val Gly Ser Phe Val Gln Pro Cys
595 600 605

Leu Arg Pro Leu Met Pro Ala Ile Ala His Ala Leu Phe Met Asp Ile
610 615 620

Thr His Asp Asn Glu Cys Pro Ile Val His Arg Ser Ala Tyr Asp Ala
625 630 635 640

Leu Pro Ser Thr Thr Ile Val Ser Met Ala Cys Cys Ala Ser Gly Ser
645 650 655

Thr Arg Gly Tyr Asp Glu Leu Val Pro His Gln Ile Ser Val Val Ser
660 665 670

Glu Glu Arg Phe Tyr Thr Lys Trp Asn Pro Glu Ala Leu Pro Ser Asn
675 680 685

Thr Gly Glu Val Asn Phe Gln Ser Gly Ile Ile Ala Ala Arg Cys Ala
690 695 700

Ile Ser Lys Leu His Gln Glu Leu Gly Ala Lys Gly Phe Ile Gln Val
705 710 715 720

Tyr Val Asp Gln Val Asp Glu Asp Ile Val Ala Val Thr Arg His Ser
725 730 735

Pro Ser Ile His Gln Ser Val Val Ala Val Ser Arg Thr Ala Phe Arg
740 745 750

Asn Pro Lys Thr Ser Phe Tyr Ser Lys Glu Val Pro Gln Met Cys Ile

ULPI_041_01W0_SeqList_ST25.txt

755

760

765

Pro Gly Lys Ile Glu Glu Val Val Leu Glu Ala Arg Thr Ile Glu Arg
770 775 780

Asn Thr Lys Pro Tyr Arg Lys Asp Glu Asn Ser Ile Asn Gly Thr Pro
785 790 795 800

Asp Ile Thr Val Glu Ile Arg Glu His Ile Gln Leu Asn Glu Ser Lys
805 810 815

Ile Val Lys Gln Ala Gly Val Ala Thr Lys Gly Pro Asn Glu Tyr Ile
820 825 830

Gln Glu Ile Glu Phe Glu Asn Leu Ser Pro Gly Ser Val Ile Ile Phe
835 840 845

Arg Val Ser Leu Asp Pro His Ala Gln Val Ala Val Gly Ile Leu Arg
850 855 860

Asn His Leu Thr Gln Phe Ser Pro His Phe Lys Ser Gly Ser Leu Ala
865 870 875 880

Val Asp Asn Ala Asp Pro Ile Leu Lys Ile Pro Phe Ala Ser Leu Ala
885 890 895

Ser Arg Leu Thr Leu Ala Glu Leu Asn Gln Ile Leu Tyr Arg Cys Glu
900 905 910

Ser Glu Glu Lys Glu Asp Gly Gly Gly Cys Tyr Asp Ile Pro Asn Trp
915 920 925

Ser Ala Leu Lys Tyr Ala Gly Leu Gln Gly Leu Met Ser Val Leu Ala
930 935 940

Glu Ile Arg Pro Lys Asn Asp Leu Gly His Pro Phe Cys Asn Asn Leu

ULPI_041_01W0_SeqList_ST25.txt

945 950 955 960
 Arg Ser Gly Asp Trp Met Ile Asp Tyr Val Ser Asn Arg Leu Ile Ser
 965 970 975
 Arg Ser Gly Thr Ile Ala Glu Val Gly Lys Trp Leu Gln Ala Met Phe
 980 985 990
 Phe Tyr Leu Lys Gln Ile Pro Arg Tyr Leu Ile Pro Cys Tyr Phe Asp
 995 1000 1005
 Ala Ile Leu Ile Gly Ala Tyr Thr Thr Leu Leu Asp Thr Ala Trp
 1010 1015 1020
 Lys Gln Met Ser Ser Phe Val Gln Asn Gly Ser Thr Phe Val Lys
 1025 1030 1035
 His Leu Ser Leu Gly Ser Val Gln Leu Cys Gly Val Gly Lys Phe
 1040 1045 1050
 Pro Ser Leu Pro Ile Leu Ser Pro Ala Leu Met Asp Val Pro Tyr
 1055 1060 1065
 Arg Leu Asn Glu Ile Thr Lys Glu Lys Glu Gln Cys Cys Val Ser
 1070 1075 1080
 Leu Ala Ala Gly Leu Pro His Phe Ser Ser Gly Ile Phe Arg Cys
 1085 1090 1095
 Trp Gly Arg Asp Thr Phe Ile Ala Leu Arg Gly Ile Leu Leu Ile
 1100 1105 1110
 Thr Gly Arg Tyr Val Glu Ala Arg Asn Ile Ile Leu Ala Phe Ala
 1115 1120 1125
 Gly Thr Leu Arg His Gly Leu Ile Pro Asn Leu Leu Gly Glu Gly

ULPI_041_01W0_SeqList_ST25.txt

1130

1135

1140

Ile Tyr Ala Arg Tyr Asn Cys Arg Asp Ala Val Trp Trp Trp Leu
1145 1150 1155

Gln Cys Ile Gln Asp Tyr Cys Lys Met Val Pro Asn Gly Leu Asp
1160 1165 1170

Ile Leu Lys Cys Pro Val Ser Arg Met Tyr Pro Thr Asp Asp Ser
1175 1180 1185

Ala Pro Leu Pro Ala Gly Thr Leu Asp Gln Pro Leu Phe Glu Val
1190 1195 1200

Ile Gln Glu Ala Met Gln Lys His Met Gln Gly Ile Gln Phe Arg
1205 1210 1215

Glu Arg Asn Ala Gly Pro Gln Ile Asp Arg Asn Met Lys Asp Glu
1220 1225 1230

Gly Phe Asn Ile Thr Ala Gly Val Asp Glu Glu Thr Gly Phe Val
1235 1240 1245

Tyr Gly Gly Asn Arg Phe Asn Cys Gly Thr Trp Met Asp Lys Met
1250 1255 1260

Gly Glu Ser Asp Arg Ala Arg Asn Arg Gly Ile Pro Ala Thr Pro
1265 1270 1275

Arg Asp Gly Ser Ala Val Glu Ile Val Gly Leu Ser Lys Ser Ala
1280 1285 1290

Val Arg Trp Leu Leu Glu Leu Ser Lys Lys Asn Ile Phe Pro Tyr
1295 1300 1305

His Glu Val Thr Val Lys Arg His Gly Lys Ala Ile Lys Val Ser

ULPI_041_01W0_SeqList_ST25.txt

1310

1315

1320

Tyr Asp Glu Trp Asn Arg Lys Ile Gln Asp Asn Phe Glu Lys Leu
1325 1330 1335

Phe His Val Ser Glu Asp Pro Ser Asp Leu Asn Glu Lys His Pro
1340 1345 1350

Asn Leu Val His Lys Arg Gly Ile Tyr Lys Asp Ser Tyr Gly Ala
1355 1360 1365

Ser Ser Pro Trp Cys Asp Tyr Gln Leu Arg Pro Asn Phe Thr Ile
1370 1375 1380

Ala Met Val Val Ala Pro Glu Leu Phe Thr Thr Glu Lys Ala Trp
1385 1390 1395

Lys Ala Leu Glu Ile Ala Glu Lys Lys Leu Leu Gly Pro Leu Gly
1400 1405 1410

Met Lys Thr Leu Asp Pro Asp Asp Met Val Tyr Cys Gly Ile Tyr
1415 1420 1425

Asp Asn Ala Leu Asp Asn Asp Asn Tyr Asn Leu Ala Lys Gly Phe
1430 1435 1440

Asn Tyr His Gln Gly Pro Glu Trp Leu Trp Pro Ile Gly Tyr Phe
1445 1450 1455

Leu Arg Ala Lys Leu Tyr Phe Ser Arg Leu Met Gly Pro Glu Thr
1460 1465 1470

Thr Ala Lys Thr Ile Val Leu Val Lys Asn Val Leu Ser Arg His
1475 1480 1485

Tyr Val His Leu Glu Arg Ser Pro Trp Lys Gly Leu Pro Glu Leu

ULPI_041_01W0_SeqList_ST25.txt

1490

1495

1500

Thr Asn Glu Asn Ala Gln Tyr Cys Pro Phe Ser Cys Glu Thr Gln
1505 1510 1515

Ala Trp Ser Ile Ala Thr Ile Leu Glu Thr Leu Tyr Asp Leu
1520 1525 1530

<210> 3

<211> 129

<212> RNA

<213> Tobacco Etch Virus

<400> 3

ucaacacaac auauacaaaa caaacgauc ucaagcauc aagcauucua cuucuaauugc 60

agcaauuuuaa aucauuucuu uuaaagcaaa agcaauuuuc ugaaaauuuu caccauuuac 120

gaacgauag 129

<210> 4

<211> 6

<212> RNA

<213> Unknown

<220>

<223> Kozak Sequence

<400> 4

gccacc 6

<210> 5

<211> 158

<212> RNA

<213> Xenopus sp.

<400> 5

cuagugacug acuaggaucu gguuaccacu aaaccagccu caagaacacc cgaauaggagu 60

cucuaagcua cauaauacca acuuacacuu acaaaauguu gucccccaaa auguagccau 120

ucguaucugc uccuaauaaa aagaaaguuu cuucacau 158

ULPI_041_01WO_SeqList_ST25.txt

<210> 6
<211> 114
<212> RNA
<213> Unknown

<220>
<223> Poly(A) 114 Tail

<400> 6
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 114

<210> 7
<211> 4599
<212> RNA
<213> Artificial Sequence

<220>
<223> Codon-Optimized AGL Coding Sequence

<400> 7
augggacacu ccaaacagau ccggauacug cugcugaacg agauggaaaa gcuggaaaag 60
acucuaauucc ggcucgagca gggauacgag cugcaguucc gccuggggccc cacccuacaa 120
gggaaggccg ugaccgucua caccaacuac ccuuucccg gcgaaacuuu caaccgggag 180
aaguuccggu cccuugacug ggaaaacccg acugagcgcg aggacgacuc ggacaaaauac 240
ugcaagcuga accuccagca gucgggcucu uuccaauauu acuucuugca aggggaacgag 300
aaguccggug gcggcuacau cgugguggac ccaauuuugc gcgugggggc ugacaaccac 360
gugcugccac uggauugugu gaccugcaa accuuccugg ccaagugccu cggcccccuc 420
gacgaauggg agucgcgccu gagaguggcg aaagagagcg gauacaacau gauucacuuc 480
acgccgcucc aaacgcuggg ucugagccgg ucaugcuacu cacuggcgaa ccagcucgaa 540
cugaacccccg auuucucccg gccaaacagg aaguacaccu ggaacgacgu gggacagcug 600
gucgagaagu ugaagaagga guggaacgug aucuguauc cgcagcugcu guacaaccac 660
accgcugcaa acucgaagug gauccaggaa cauccggaau gugccuacaa ccucgugaac 720

ULPI_041_01WO_SeqList_ST25.txt

agcccgacc ugaagccugc cuggguguug gauagggcac uguggagguu cuccugugac	780
guggcagagg gaaaguacaa ggagaagggu auucccgcac ucaucgaaaa cgaucaccac	840
augaauucga ucagaaagau uaucugggag gauaucuucc cuaagcugaa gcugugggag	900
uucuuucaag uugaugugaa caaggcaguc gaacaguuuc ggcggcuguu aacccaagaa	960
aaccgccgcg ugaccaaguc cgauccgaau cagcaucuga ccuaaucca ggacccggaa	1020
uaccgccggu uuggcugcac cguggacaug aacauugccc ugacuaccuu uaucccgcau	1080
gauaagggcc ccgccguau cgaagaugc ugcaacuggu uccacaagag gauggaggaa	1140
cugaacuccg aaaagcauag gcucauuuac uaccaccagg aacaggcagu gaacugccug	1200
cuggggaacg uguucuacga gcgacuggcu ggacacggac cgaaguagg acccgugaca	1260
aggaagcacc cacuggucac uagauacuuc accuucccau uugaggaaau cgacuucua	1320
auggaggagu cgaugaucca cuugccuaac aaggccugcu uucucauggc acauaacgga	1380
ugggucaugg gcgacgaucc ccuacggaau uuugcagaac caggcagcga ggucuaccuu	1440
cggcggaac ugauuugcug gggcgacucc gucaagcugc gcuacggcaa caagccugag	1500
gacugucccu accuuugggc acacaugaag aaguacacug aaauuacugc gacauacuuc	1560
caaggagucc gcuuagauaa uugucacucc accccgcugc auguggcgga guacaugcug	1620
gaugccgcaa gaaaccucca gccgaaucuc uacgugguag cagagcuguu caccgggagc	1680
gaggaccugg acaauguguu ugucacccgg cuggggaucc ccucccugau ccgggaggcc	1740
auguccgcu acaacucaca cgaggagggg agacuggugu accgcuacgg aggagaaccc	1800
gugggcagcu uugugcagcc uugccuccgg ccgcugaugc ccgcgauugc gcaugcucug	1860
uucauggaua ucacucacga uaacgagugc ccuauugugc acagaucgc cuacgacgcc	1920
cuuccuucca caaccaucgu guccauggca ugcugcgccu ccggcuccac ucgggguuac	1980
gaugagcugg ugccacacca gauuuccgug guguccgaag aacgcuucua caccaagugg	2040
aacccggaag cucugccguc aaacaccgga gaagugaacu uccaguccgg gaucaucgca	2100
gcgcgcugug cuauuagcaa gcugcaccag gagcugggag ccaagggguu cauccagguc	2160

ULPI_041_01WO_SeqList_ST25.txt

uauguggacc	aggucgauga	ggauaucguc	gcugucacga	gacacagccc	gucuauccau	2220
caaagcgucg	uggccguguc	ccggacugcg	uuccggaacc	cuaaaaccuc	auucuaaucc	2280
aaagaggugc	cccagaugug	caucccugga	aagaucgaag	aagucgugcu	ggaagcccgg	2340
accaucgagc	ggaacaccaa	gccguacagg	aaggacgaaa	acuccaucaa	ugguaccccu	2400
gacauuaccg	uggaaaucag	agaacauauc	cagcugaacg	aguccaagau	cgugaagcag	2460
gccggcgugg	cgaccaaggg	ucccaacgag	uacauucagg	aaaucgaguu	ugaaaaccug	2520
ucccccggaa	gcgugaucan	uuuccgggug	ucccuggacc	cgcaugcgca	agucgcuguc	2580
ggaauucugc	ggaauaccu	cacccaauuc	ucgccgcauu	ucaagagugg	uucccuggcg	2640
guggauaaug	ccgauccgau	ccugaagauu	ccuucgcgu	cccuggcauc	gagacucacc	2700
cuggcgaggau	ugaaccagau	ucuguaccgc	ugcgaauccg	aggaaaagga	ggacggaggc	2760
gguugcuacg	acauccccaa	cugguccgca	cuuaaguacg	cagggcugca	gggucuuaug	2820
agcgugcugg	cagaaaauucg	cccuaagaac	gauuugggac	accccuucug	caacaaccuc	2880
cgguccggag	acuggaugau	cgauuacgug	ucgaacagac	ugauuucgag	auccggcacc	2940
auugccgagg	ucggaaagug	gcuucaggcc	auguucuucu	accugaagca	gaucgagaga	3000
uaccugauuc	ccugcuacuu	cgacgcaauc	cugaucgggg	cguauaccac	ucuucuggac	3060
acugccugga	agcagauguc	cagcuucgug	caaaacggau	ccaccuucgu	caagcaucuu	3120
agccuggggcu	cagugcaguu	gugcggagug	ggaaaauucc	cuagccuccc	uauucuuuca	3180
ccggcgcuga	uggacgugcc	uuaccgccug	aacgaaauca	ccaaagagaa	ggagcagugu	3240
ugcgugucgc	uggccgcggg	ucugccgcau	uucuccuccg	gcaucuuccg	gugcuggggga	3300
agggacaccu	ucaucgcucu	gaggggaau	cugcugauua	ccgggcgcua	cguggaagcu	3360
cggaacauca	uccuggccuu	cgcggggaac	cugcgccacg	gccugauucc	aaacuugcuu	3420
ggcgaaggca	ucuacgcgcg	cuacaacugc	cgcgacgcgg	ucugguggug	gcuccagugc	3480
auucaagacu	acugcaagau	ggugccaaac	ggccuggaca	uccugaagug	cccggugucg	3540
agaauguacc	ccaccgacga	uucugcgccc	cugccggccg	guacucuuga	ccaaccucug	3600

ULPI_041_01WO_SeqList_ST25.txt

uucgaaguga uccaggaagc aaugcagaag cacaugcagg gcauucaguu ccgggagcgc	3660
aacgcagggc cgcaaaucga caggaacaug aaggacgaag gauucaacau caccgcggga	3720
guggacgaag agacuggcuu cgucuacggu ggaaaucggu ucaacugcgg gaccuggaug	3780
gacaagaugg gcgaaucaga ccgagcccg c aaccgcggaa ucccugccac cccccgggau	3840
gggagcgccg uggagauugu gggacugagc aagagcgcug ugcgugggcu gcuggaguug	3900
agcaagaaga acauuuuccc cuaucacgaa gugaccguga agcggcacgg aaaagcuau	3960
aaaguguccu acgacgagug gaacagaaag auccaggaca acuucgagaa gcuguuccac	4020
guguccgagg acccgucgga cuugaaugag aagcacccua accucgugca caagcgggga	4080
aucuacaagg acagcuacgg agcauccucg ccuuggugcg acuaucagcu gaggcccaac	4140
uucacuaucg caaugguggu ggccccagaa cuguucacua ccgaaaaggc cuggaaggca	4200
cuggagaucg cagagaaaaa gcugcugggc ccucugggca ugaaaacccu ggaccccgac	4260
gacauggugu acugcgggau cuacgaaac gcucuugaca augacaacua caaccuggc	4320
aagggaauca acuaucacca gggccccgag uggcuguggc cgauccgguu cuuccugcgc	4380
gccaagcugu acuuuuuccg gcugaugggc ccugagacua ccgcaaagac gaucgugcug	4440
gucaagaacg ugcugucacg gcacuacgug caucuggaac ggagcccgug gaagggguug	4500
cccgaacuga ccaacgaaaa cgcgcaguac uguccguucu cgugcgaaac ucaggccugg	4560
uccaucgcca cuauccucga aacucucuac gaccuguag	4599

<210> 8

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 8

augggacaua gcaaacaaau uaggauccuc cugcugaacg aaauggaaaa acuggaaaag	60
---	----

acgcuguuuc ggcuggaaca gggcuacgaa cuccaguuuc gccucgggcc aacgcugcaa	120
---	-----

ULPI_041_01WO_SeqList_ST25.txt

gggaaagcag ugaccgugua uacaaacuac ccauuccug gagagacauu caauaggag	180
aaguuuagaa gcuuagacug ggaaaaucca accgaacgag aagaugacag cgauaaguau	240
ugcaagcuga aucugcaaca guccggaagu uuucaguacu acuuuuugca agggaacgaa	300
aagagcggcg gagguuauau ugugguggau ccauuuuga gaguuggggc cgacaaccu	360
guacugccuc uugacugugu gacacugcaa acguuccugg ccaagugccu gggccguuu	420
gaugaguggg aaagccguuu gcgcguggcc aaggaaucg gguacaacau gauccauuuu	480
acccacugc aaaccuuagg ccugucaagg agcuguuaca gucuggccaa ccagcuggaa	540
cugaaucccg acuucuccg gccaaauagg aaguacacgu ggaacgacgu uggccagcua	600
guggagaagu ugaagaagga guggaacguu auuugcauca cugacgucgu guauaauac	660
acugccgcua auagcaaaug gauccaggaa cauccagagu gugccuauaa ccuggucaac	720
agcccgcauc ugaagccagc augggugcug gaccggcgcu uguggcgguu cucgugugac	780
guggccgaag gaaaguauaa ggagaagggu auuccggcuc ugauagagaa cgaccauau	840
augaauucca ucaggaagau cauugggag gacaucuuc caaaguugaa gcugugggag	900
uucuuucaag uggaugugaa caaggccgua gaacaguuuc gccgccuacu aaccaggaa	960
aauaggagag ucacaaaguc cgaucccaac caacaucuga ccuaauuca ggaccccgaa	1020
uaucgucgau uggcugcac cguggauaug aacaucgcu ugacaacuuu cauaccacac	1080
gacaaggguc cagccgccau ugaagagugu uguaacuggu uucacaaacg aauggaggag	1140
uugaacagug aaaagcaccg ccugaucau uaucaucagg agcaggccgu gaacuguuug	1200
cuggggaug uguucuauga gaggcucgcu ggucauggcc cuaagcuggg uccuguaaca	1260
agaaagcauc cauugugac gcgcuacuuu acauucccau uugaggagau cgacuuuucg	1320
auggaagagu ccaugauuca ucucccaac aaggccuguu uucugauggc ccacaacggc	1380
uggguuauug gggaugaucc guugagaaau uuugcggaac cagguucaga agugauucug	1440
cgucgggagc ugauuuuguu gggcgacucc gugaagcucc gcuaugggaa caagccugag	1500
gauuguccuu aucugugggc gcuaugaag aaguauacag agaucacugc cacauuuuu	1560

ULPI_041_01WO_SeqList_ST25.txt

cagggcguca ggcuggacaa uugucauagc accccgcuuc auguggcuga guauaugcug	1620
gacgcagcaa ggaaucugca gcccaaucug uauguggugg cggaacuguu uaccgggucc	1680
gaggaccugg acaauguauu uguaaccga uugggcaucu ccagccugau uagggaagca	1740
augagugcau acaacucaca cgaggagggg cggcugguuu aucgauaugg gggggaaccu	1800
gugggcagcu uuguacagcc augucuccgg ccguugaugc cugccauugc ucaugcgcuc	1860
uuuauaggaua uaacacauga caacgaugu ccaaucguuc auaggagugc uuacgacgcc	1920
cugccgagca caacgaucgu guccauggcc uguugugcaa guggcagcac acguggcuau	1980
gacgaauugg uacccaccca aauuagcgug guuuccgagg agagguucua uacaaagugg	2040
aauccagagg cuuugcccuc gaacaccgga gaggucaacu uucaaucugg cauaauugcc	2100
gcucgcugcg ccuauucuaa guugcaucag gaacuggggcg cgaagggguu uauacaaguc	2160
uacguugacc aaguggauga ggacaucguu gcugucacgc ggcacucacc uaguauucau	2220
caauccgugg uagcaguguc ucggacggcc uuuagaaacc caaagacauc auuuuacucg	2280
aaagaagugc cucaaauug uauaccuggc aaaaauagaag aggugguccu ggaagcccgc	2340
acgauagaga ggaauacuaa gccguauaga aaggaugaaa auucuaucuaa cggcacuccg	2400
gauauuacag uagaaaucag agagcacauu caacuuaug agucuaagau uguuaagcag	2460
gcugguguug caaccaaagg gccaaacgag uacauccagg agauugaauu ugaaaaucug	2520
ucaccgggcu ccgugaucau cuuuagagua ucuuuggauc cacaugcuca agucgcuguu	2580
ggaauccuga gaaaccaucu gacacaauuu uccccacauu uuaagagcgg cagccuggcc	2640
guggacaacg cagacccaau ccugaagauc ccuuugcau ccugggcuuc ccgccugacu	2700
cuggcugagc ugaaccagau cuuguaucgc ugugaaucug aagaaaaaga ggauggcgga	2760
ggcuguuaug acaucccaaa uuggagcgcc cugaaauaug ccggacucca aggccugaug	2820
uccguucugg ccgagauuag gccaaagaau gaucugggac acccauuuug uaacaacuug	2880
cggucuggag acuggaugau cgacuauguc agcaaccgcc ugauaagcag auccgguacu	2940
aucgcugaag ucggaaaaug gcugcaggcu auguuuuucu auuugaagca gauuccgaga	3000

ULPI_041_01WO_SeqList_ST25.txt

uauuugaucc ccuguuauuu ugaugccauu cugauuggug cuuauacuac ccugcuggau	3060
accgcuugga aacagaugag uagcuuuguc caaaauggcu ccaccuucgu aaagcaucug	3120
ucgcugggcu ccgugcaauu guguggcguc gggaaguucc cuagccugcc aauccugucc	3180
ccugcucuga uggacgugcc uuauccgucg aacgagauua cgaaggaaaa ggaacagugu	3240
ugcgugucac ucgcugcugg gcugccacac uuuucuucug gaauuuuucg guguuggggg	3300
cgggacaccu uuauugcccu gaggggcauu cugcugauua ccggccgcua ugucgaggcu	3360
aggaacauua uccuggcuuu cgcggguacc cuucggcacg gacucauacc caaccuccug	3420
ggagaaggaa ucuaugcacg guauaacugu cgggacgcag uuugguggug guugcagugu	3480
auucaagauu auugcaagau ggucccgaac ggacuggaca uacugaagug cccagugucc	3540
cggauguauc caacugacga cuccgcacca cugcccgcug ggacacuuga ccagccauug	3600
uuugaaguua uucaggaagc uaugcagaag cauaugcagg gaauucaguu uagggagaga	3660
aacgcuggac cccagauuga ccguaacaug aaggauagg gguuuuacau uaccgcggga	3720
guggacgaag agacuggcuu cgucuauggu ggcaaucggu ucaacugcgg caccuggaug	3780
gacaaaugg gcgaaagcga uagagcucga aaauaggggca ucccagcaac accacgggau	3840
ggcucugccg uggagauugu gggccugucu aagagcgag uucggugguu guuggaacuc	3900
agcaagaaga acauuuuucc auaccauga guuacaguga agcggcaugg gaaggccauc	3960
aagguuucuu acgacgaug gaauccggaag auccaggaua acuucgagaa acuguuucau	4020
guauccgagg acccuucuga ccugaacgaa aagcauccaa auuugguaca caaaagggga	4080
aucuauaagg auaguuaugg agccagcagc ccuuggugcg acuaucagcu acgaccaaac	4140
uucacuauug ccaugguagu agcaccagaa cucuuuacua ccgaaaaggc uuggaaggca	4200
cuggagaucg ccgagaagaa gcuguuaggu ccacugggca ugaaaacccu ggacccccgac	4260
gacaugguuu acuguggcau cuaugacaac gcucuggaca augacaacua uauucuggcc	4320
aagggcuuca acuaccacca gggcccagag uggcuauggc caauuggcua uuuucugcgg	4380
gcgaagcugu acuuuucacg auugaugggc ccagagacaa ccgcuaagac cauaguauug	4440

ULPI_041_01W0_SeqList_ST25.txt

gucaagaacg ugcugagucg gcacuauguc caccuggaaa ggagccccug gaaggggucg	4500
cccgagcuga ccaacgaaaa ugcacaguau uguccguuuu caugcgaaac ccaggcuugg	4560
uccaucgcga cgauccugga aaccucuaau gaucuguag	4599

<210> 9
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 9	
augggacacu caaaacagau aaggauccug cugcugaacg aaauggagaa acuggaaaag	60
accucguucc ggcucgaaca gggauaugaa cuccaguuaa ggcuggggccc aacucugcaa	120
ggaaaagcug ugaccgucua uacuaacuau ccuuuucccg gggaaacguu uaaccgagag	180
aaguuccgau ccuuggacug ggaaaauccc acagagcggg aggaugauuc ggacaaguau	240
ugcaagcuga aucugcaaca aucgggcagc uuucaguacu auuuucuaca aggaaacgaa	300
aagagcggcg ggggauacau ugugguggac ccaauccuga ggguuggggc agacaauac	360
guacugccuu uggacugugu gacauugcaa acauuccugg caaagugccu gggcccccuc	420
gaugaauagg agucacggcu gcgaguggcu aaggaaucug gcuauaaauau gauucacuuu	480
acaccauuac aaacgcuggg ccugagcaga aguugcuacu cgcuggcuua ucaguuggag	540
uugaacccgg acuuuagccg ucccaaccgg aaguauacau ggaaugaugu uggccaacuc	600
guggagaagc ugaagaagga guggaacgug aucuguauua ccgaugucgu guauaaucan	660
acagccgcga acuccaagug gauucaggag caccagaau gugcauauaa cuuggucaau	720
agccccacc ugaagccggc cuggguacug gaucggggccc uguggcgggu uucuugugac	780
guggcagaag gaaaguauaa ggaaaagggg auccagcac ugauagagaa cgaccaucau	840
augaacagca uuaggaagau cauuugggag gacauuuucc caaaguugaa gcucugggag	900
uuuuuccaag uugacgucaa caaggccgug gaacaauuuc gaaggcugcu gacacaggaa	960

ULPI_041_01WO_SeqList_ST25.txt

aacagacgug	uaacgaaguc	cgaccccaau	cagcauuuga	cgaucauaca	ggauccugag	1020
uauaggcgcu	uuggcuguac	ggucgacaug	aaauuugcuc	ugaccacuu	uaucuccau	1080
gacaaggggc	cagccgccau	cgaagagugu	uguaauuggu	uccacaagcg	cauggaagag	1140
uugaauuccg	aaaagcauag	guugaucaac	uaccaccagg	aacaggcagu	gaacugucug	1200
cugggcaacg	uguuuuuuga	gcggcuggcc	ggucauggcc	cuaagcuggg	accagugacc	1260
aggaagcauc	cacucguaac	cagauuuuu	accuucccg	uugaggagau	agauuuuagc	1320
auggaagaau	ccaugauuca	ccuaccgaac	aaggcuuguu	uucugauggc	ucacaacgg	1380
uggguuauagg	gcgaugaucc	ccugcgcaac	uuugcggaac	cgggcucuga	gguguacuug	1440
agaagggagc	ugauauguug	gggggacucc	gucaaacugc	gcuauaggga	uaagccugaa	1500
gauugcccuu	aucugugggc	ccauaugaag	aaguauaccg	agauuaccgc	gacauuuuu	1560
caaggagugc	ggcuggacaa	cugccacagc	acgccgcugc	auguggcaga	guauaugcug	1620
gacgcagccc	gcaaucugca	gccaaucug	uauguugugg	cugagcuguu	uacgggcucc	1680
gaggaccucg	acaacguuuu	cguaacccgg	cugggcauca	gcagccugau	ccgggaagcu	1740
augucagcgu	auaaauagcca	ugaggagggc	agacuggugu	acagauaugg	uggcgaacca	1800
guuggcagcu	uugugcagcc	cuguuugcgc	ccucugaugc	ccgccaucgc	acaugcuuug	1860
uucauggaca	ucacgcauga	uaacgaaugu	cccuuuguac	aucgcuccgc	cuaugacgcu	1920
uugccaacca	caacaucgu	guccauggcu	uguugugcaa	gcggcagcac	caggggauau	1980
gacgaauugg	uuccgcauca	aaucagcgug	guaucagagg	aaagguuuua	uacaaagugg	2040
aauccugaag	cccugccauc	caacaccggc	gaagugaacu	uucagucggg	uaucauugcc	2100
gcgcguugcg	cuauuagcaa	acugcaucag	gagcugggug	cuaaaggcuu	uauccaaguu	2160
uauugggaur	aaguagacga	agauuuuguu	gccguaacca	gacauagccc	cagcauucau	2220
caauccgugg	uugcuguguc	ccggacggcc	uucaggaacc	cuaaaacauc	cuuuuauucc	2280
aaggaagucc	cacagaugug	uauuccggga	aagauagagg	aagugguuuu	ggaggcucgc	2340
acgauugagc	ggaacaccaa	gccauauagg	aaggacgaga	acucgaucua	cggcacgccc	2400

ULPI_041_01W0_SeqList_ST25.txt

gacauuaccg uugaaaauucg cgagcacauu cagcugaacg aaucgaaaau ugucaagcag	2460
gccggcguag cgaccaaggg uccaaacgag uauauccagg aaucgaguu ugagaauucg	2520
ucgccugggu ccguaaucau uuuuagaguc agccuagacc cucacgcuca aguggccgua	2580
gggauccuga ggaaucaucu gacgcaguuu ucaccccauu uuaaguccgg cagccuggca	2640
guggauaacg ccgaccccau ccugaaaauu cccuuugcuu cccuggccuc gcgccugaca	2700
cuggcagaau ugaaucauac acuguaucgu ugcgagagcg aggagaagga ggacggaggg	2760
ggguguuauug auaucccgaa cugguccgcu cugaaauaug cgggauugca gggcuugaug	2820
aguguacugg ccgaaaauag acccaagaac gaucuggggcc auccguucug uaacaauucg	2880
cgauccggcg auuggaugau ugauuauguc agcaaccggu ugauaagccg gagugggaca	2940
auugcagaag ucggaaagug guugcaagcc auguucuuuu accugaagca gaucccucga	3000
uauugauac cauguuuuuu cgaugccauu cugaucgggg ccuaauacaac ucuguuagac	3060
acugcuugga aacagauguc uagcuucgug caaaacggcu caacauuugu uaagcacuug	3120
agccugggggu cugugcaguu guguggcgua ggaaaguuuu ccucacugcc cauccugucg	3180
cccgcccuga uggacgugcc cuaccgccuc aacgaaaaua ccaaggagaa agagcagugu	3240
uguguuagcu uggcugccgg uuugccgcau uuuucaagcg gaaucuuccg augcugggga	3300
cgcgauacgu ucauagcccu gagggguuuu cugcugauua cgggcagaua cguagaggcu	3360
cggaaauuuu uccuggccuu cgccggcacg cugcggcaug ggcugauucc gaacuugcug	3420
ggagagggca ucuacgcccg guacaacugc agggacgcug ugugguggug gcugcagugc	3480
auccaggacu auugcaaaaau ggugccaaac ggccuugaca uacugaagug uccugucucu	3540
cggauguauc caacggauga cuccgcaccc cugcccgcuu gaaccucgga ucaacccug	3600
uuugaaguga uacaggaagc aaugcagaag cacaugcagg gaauucaguu uagagagaga	3660
aaugcaggcc cccaaaucga cagaaacaug aaagaugaag gcuuuuacau cacggcugga	3720
guggacgaag aaacuggguu uguguauggc gggaauaggu uuaacugugg gacguggaug	3780
gacaagaugg gcgaauccga uagagcgcgg aacaggggaa ucccggcuac accccgggac	3840

ULPI_041_01WO_SeqList_ST25.txt

gguagugcug	uggagauugu	cggcuuaucc	aaauccgccg	ugcgugggcu	gcuggagcug	3900
uccaagaaga	acaucuuucc	uuaccaugaa	gugaccguga	agcggcaugg	aaaggccauc	3960
aaagucuccu	acgacgaug	gaauaggaag	auucaggaca	auuuugagaa	gcuguuuau	4020
guguccgagg	aucccagcga	ccugaacgag	aagcacccua	auuuggugca	uaagcggggc	4080
aucuauaagg	acuccuacgg	cgcuauguug	ccuuggugcg	acuaucagcu	gcggccaaac	4140
uuuacgauug	cgaugguggu	agccccgaa	uuguuuacga	cggaaaaggc	uuggaaggcc	4200
cuggagaucg	cagaaaagaa	gcugcugggg	cccuugggca	ugaaaacgcu	ggacccccgac	4260
gauaugguuu	auuguggcau	cuacgacaac	gccucgcaca	augacaauua	caaccuggca	4320
aaggguuuuu	acuaucauca	ggguccugaa	uggcuguggc	caauuggcua	cuuuuugcgg	4380
gccaagcugu	acuuuucacg	gcugauggga	ccugaaacca	ccgcuaagac	cauuguacug	4440
gucaagaacg	ugcugagcag	gcauuacgug	caucuggaaa	gaagcccug	gaagggucug	4500
ccagaguuga	cgaacgagaa	cgcgcaguau	uguccuuucu	cuugugaaac	gcaggcuugg	4560
aguauugcaa	ccauucugga	aacucuuuau	gaccuguag			4599

<210> 10
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 10	
augggccacu	caaagcagau uagaauuuu cuccugaacg agauggaaaa gcuggaaaag 60
accuguuucc	gccuggaaca gggauaugag cugcaguucc ggcuuggacc aaccuccaa 120
ggaaaggccg	ugaccgugua cacgaacuac ccuuuccgg gcgaaaccuu caacagagag 180
aaguuccgga	gccuggacug ggaaaacccg accgaaaggg aagaugauag cgacaaguac 240
ugcaagcuga	acuugcaaca guccggauca uuccaguacu acuuucugca agggaacgag 300
aagucgggcg	gcgguuacau cgugguggac ccuauucugc gcgugggagc cgacaaccac 360

ULPI_041_01W0_SeqList_ST25.txt

guguugccuc	uggacugugu	cacccugcaa	acuuuccugg	ccaagugccu	ggggccauuc	420
gaugaguggg	aaucgcggcu	gcggguggca	aaagaguccg	guuacaacau	gauccacuuc	480
accccacucc	aaaccugggg	ccuguccagg	ucgugcuacu	cgcuggccaa	ccagcuggaa	540
cugaacccgg	auuucucgcg	gccuaacaga	aaguauaccu	ggaaugacgu	cggacagcuc	600
gucgaaaagc	ugaagaagga	auggaacgua	aucuguauc	ccgaugucgu	guauaaccac	660
accgcugcca	acuccaagug	gauccaggaa	caccccgagu	gcgcuuacaa	cuuggugaac	720
ucgccccacc	ugaaaccggc	uugggugcug	gaccgggccc	uguggcgguu	cagcugcgau	780
guggcugagg	gaaaguauaa	ggagaagggu	aucccagcgc	ucauugaaaa	cgaccaccau	840
augaacucga	uccgcaagau	cauuugggag	gauauuuucc	caaaguugaa	acuguggggag	900
uuuuuccaag	uggacugaa	caaggcgguc	gaacaguuca	gacggcuguu	gacucaggaa	960
aaucggaggg	ucaccaaguc	ggaucccaac	cagcaccuga	ccauaaucca	ggauccagaa	1020
uaccggcggu	ucggcugcac	uguggacaug	aacauugccc	ugaccacuuu	uaucccacac	1080
gacaagggcc	cggccgccau	cgaagagugc	ugcaacuggu	uccacaagcg	gauggaggaa	1140
cugaacucgg	agaagcaccg	ccugaucaac	uaccaccagg	agcaggcggu	uaacugucug	1200
uugggaaacg	uguucuaacg	gcggguuggc	ggccacggcc	cuaagcuggg	cccggucacc	1260
cggaagcacc	cccuggucac	ccgcuacuuc	accuuuccau	ucgaggagau	cgacuucagc	1320
auggaagaau	cgaugaucca	ucugccgaac	aaggccugcu	uccugauggc	ucacaacgga	1380
uggguugaugg	gcgacgaucc	ucugagaaac	uuugccgagc	cgggcucgga	gguguaccug	1440
aggagggagc	ugaucuguug	gggggacagc	gugaagcuga	gauacggaaa	caagccagag	1500
gacugcccgu	accugugggc	ccauaugaag	aaguacaccg	agauuaccgc	aacauacuuc	1560
caaggggucc	ggcuggacaa	uugccacucg	acuccccugc	acguggcgga	guacaugcug	1620
gacgcggcca	ggaaccucca	gcccaccuc	uacgucgugg	ccgagcuguu	cacugguucc	1680
gaggaccugg	auaacguguu	cgugaccaga	cuggggauuc	ccucacugau	ucgcgaagcc	1740
auguccgcgu	auaacucgca	ugaagagggc	cgccuggugu	accgcuacgg	aggcgaaccu	1800

ULPI_041_01W0_SeqList_ST25.txt

gugggcagcu ucgugcaacc gugucugcgg ccucugaucg cugccauugc gcacgcccug	1860
uucaugggaca ucacccacga caacgaauugc cccaucgugc accgguccgc cuacgaugcc	1920
cucccuucga cuacuauugu guccauggcu ugcugcgcgu ccggcucgac ccgcggcuac	1980
gaugagcucg ugccgcauca gauuagcgug guguccgagg aaagguucua caccaagugg	2040
aauccagaag cgcugccgag caacaccggg gaggucaacu uccagucggg aaucaucgcc	2100
gcccgcugug ccaucuccaa gcugcaccag gaacuggggcg cgaaggguuu cauccaaguc	2160
uacgucgauc aggucgacga ggacaucgug gccgugacuc ggcauucccc gagcauucac	2220
caguccgugg uggccguguc gcgcaccgcc uuccgcaacc cuaagaccag cuucuauuca	2280
aaagaagugc cgcagaugug caucccugga aagaucgagg aggugguccu ggaagcgcgg	2340
acuaucgaaa ggaacaccaa gccauaccgc aaggacgaga acuccaucaa cgguaccccg	2400
gacaucaugc uggagaucgg cgagcauuu caacugaacg aguccaagau cgugaaacag	2460
gccggagugg caaccaaggg accgaacgag uauauccagg aaauugaguu cgagaaccuc	2520
uccccgggaa gcgugauuuu cuuccgcgug ucccuggacc cacacgcca aguggccguc	2580
ggcaucuugc ggaaccaccu gacucaguuc uccccgcauu ucaaguccgg aagccuggcg	2640
guggacaacg cggacccaau ccugaagaua cccuucgccu cacuggccuc gcgccugacu	2700
uuggccgagu ugaaucagau ccuguaccgc ugcgaauccg aagaaaagga ggacggugga	2760
ggcuguuacg acaucccuaa cugguccgca cugaaauacg caggacugca gggacugaug	2820
uccguccucg cugaaaucag accgaagaac gaucucggcc acccauucug caacaaccug	2880
agauccggcg auuggaugau ugauuacgug ucgaaccgac ugaucucccg cucggguacu	2940
auugcgggaag ucggaaaaug gcugcaggcc auguucuucu accugaagca gaucacacga	3000
uaccuuaucc cuugcuacuu ugacgcgauu cugauuggug ccuacacgac ccugcuggac	3060
acggccugga agcaaauguc cagcuuugug cagaacggca gcaccuucgu gaaacaccug	3120
ucgcugggau cggugcagcu cugcggcgug gggaaguuuu cgucacugcc gauccugucc	3180
ccggccuuga uggaugugcc cuaccggcug aacgaaaaua ccaaggagaa ggagcagugc	3240

ULPI_041_01WO_SeqList_ST25.txt

ugcgugucgc	uggccgccgg	gcugccucac	uucuccuccg	gaauuuuucg	augcugggggu	3300
agagacacuu	ucauugcgcu	gaggggggauu	cugcuuauua	ccggccgcua	cguggaagcc	3360
aggaacauca	uccuggcauu	cgccggaacc	cugcggcacg	ggcugauucc	caaccuccuc	3420
ggagaaggaa	ucuacgcucg	guacaauugc	cgggacgcag	ugugguggug	gcuccagugc	3480
auccaggacu	acugcaagau	ggucccuaac	ggacuggaca	uucugaagug	cccagugucc	3540
cggauguacc	ccacugacga	uucggcaccg	cugccggcug	guaccucgga	ccagccgcug	3600
uucgaaguga	uccaggaagc	caugcaaaag	cauauugcagg	ggauucaguu	ccgcgaaaga	3660
aaugccggac	cucagaucga	ccgcaacaug	aaggaugaag	gcuuuuacau	cacugccgga	3720
guggacgaag	agacuggauu	cgucuacggg	ggaaacagau	ucaauugcgg	uacauggaug	3780
gauaagaugg	gagaaagcga	cagagcucgg	aaauagaggaa	uuccggccac	accucgggac	3840
ggcucagccg	uggagaucgu	ggggcuaucc	aagucugccg	ugcgucggcu	guuggaacug	3900
ucgaagaaga	acauuuuccc	auaccacgag	gucaccguga	agcgccacgg	aaaggccauc	3960
aaagugucau	acgaugaaug	gaaccgcaag	auucaggaca	acuucgagaa	gcuguuccau	4020
gugucggagg	acccuuccga	ccugaacgaa	aagcauccaa	accuggugca	caagcggggg	4080
aucuacaagg	acuccuacgg	agcguccucc	ccuuggugcg	acuaccaacu	ccggccaaaau	4140
uucacgaucg	cgaugguggu	ggccccugaa	uuguuuacca	ccgaaaaggc	cuggaaggcc	4200
cucgagaucg	cagagaagaa	acugcuggga	ccccugggca	ugaaaacccu	ggacccggac	4260
gauauggugu	acugcggaau	cuacgacaac	gcccucgaua	augacaacua	uaacuuggcc	4320
aaggguuuca	auuaccacca	ggggccagag	uggcuguggc	caaucgguaa	cuuucugcgg	4380
gccaagcugu	acuucucgag	auugaugggg	cccgaaacca	ccgcuaagac	uauugugcuc	4440
gucaagaacg	ugcugucacg	gcauuacgug	caccuggaac	gcagcccaug	gaagggccug	4500
ccggaacuga	ccaacgagaa	cgcacaguac	uguccguucu	cgugugaaac	ucaggccugg	4560
agcauugcga	ccauccugga	aacucucuau	gaucuguag			4599

ULPI_041_01W0_SeqList_ST25.txt

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 11

augggccauu cuaagcaau ucggauacug cugcugaacg aaauggagaa acugaaaag	60
accuguuucc ggcuggagca gggcuauag cugcaguuaa gauugggccc gacacugcag	120
gggaaggcag ucaccguuaa uaccaauuau ccuuuuccug gggaacguu caaucgugaa	180
aaguuuucgga gccucgacug ggaaaacccu accgagcgag aggaugauag cgacaaguau	240
uguaaacuca aucugcagca gucaggcuca uucaauauu auuuucugca agggaacgag	300
aagucuggcg gaggcuacau ugucguugau ccgauacugc ggguuggcgc cgacaaccac	360
guguugccuc uggacugcgu gacacugcaa acauuucucg cgaaugcuu agggcccuuu	420
gaugaguggg aaagccgccu gagagucgcc aaggagagcg gcuauauau gauccauuuu	480
acgccgcugc aaacacuggg ucugucuagg ucauguuacu cauuggcuua ccagcuggag	540
cugaacccgg acuuuaguag gccuaacagg aaguauacau ggaaugaugu ugggcaacug	600
guggagaagc ugaagaagga auggaauguc auuuguauua ccgaugucgu guacaaccac	660
acggcagcaa auucaaagug gauucaggaa cauccggaau gcgcuuauaa ccucguaaac	720
ucaccgcacc ugaagccggc cuggguacug gacagggcac uguggagauu cucuugugau	780
guagccgagg ggaaguacaa ggagaaaggc auucccguc uuaugaaaa ugaccaccau	840
augaauagca uuaggaagau caucugggaa gauauuuuuc caaaacugaa gcugugggaa	900
uucuuucaag uugacgugaa caaggcuguc gagcaguua ggcggcuguu gaccaggag	960
aacagacggg uuacuaagag ugacccaac caacaucuga cgaucaucca ggaccagag	1020
uaucggcgc uuggcugcac aguggacaug acaucgccc ugaccacguu uaucccccac	1080
gauaagggcc cggccgcua ugaagagugu uguaacuggu uucauaagcg cauggaagaa	1140
cucaacagcg aaaagcauag acugaucaac uaucaucaag aacaggcugu gaauugucug	1200

ULPI_041_01WO_SeqList_ST25.txt

cugggcaacg	uguuuuuuga	gcgguuggcu	ggccauggcc	ccaaguuggg	ccccgugacg	1260
cgaaaacauc	cccucgugac	aagauauuuc	accuuuccau	uugaagagau	cgacuucuca	1320
auggaggaaa	gcaugauuca	ucugccaaac	aaggcuuguu	uucugauggc	ucauaauggc	1380
uggguuaugg	gcgacgaccc	ccugagaaaac	uuugcggaac	cagggucuga	aguguaucug	1440
cggagggaau	ugauuugcug	gggagacucu	guaaagcugc	gauaugggaa	caagccggag	1500
gauuguccgu	aucugugggc	ccauaugaag	aaguauacag	aaauuaccgc	aacguauuuu	1560
cagggcguac	gccuggauaa	uugucuuca	acgccacugc	auguggcaga	guacaugcua	1620
gacgcggcua	ggaaucugca	gccgaaucug	uangucgugg	ccgagcuguu	caccggcucc	1680
gaagauuugg	acaauuguuu	ugugacccgc	cucggauuaa	gcucacugau	uagagaagca	1740
augucggcuu	acaauucgca	ugaggaaggu	cgccuggugu	accguuacgg	cggcgaaccc	1800
gugggcaguu	uugugcaacc	cugucugagg	ccacuaaugc	ccgcuauugc	acaugcccug	1860
uuuauggaua	uaacacauga	uaacgaugu	ccaauuguac	accggagugc	uuauaugcu	1920
cugcccuaa	cgaccaucgu	aucaauaggca	ugcugugccu	ccggcucaac	acggggcuau	1980
gaugaacugg	uaccacacca	gaucagcgug	guuucugagg	aacgguuuuu	cacaaagugg	2040
aauccugagg	cguugccauc	caacacuggc	gaagugaacu	uucagagcgg	cauuauugcu	2100
gcccgaugug	cuauuucaaa	gcuacaccag	gaguuggggcg	caaagggguu	uauccagguu	2160
uauguggacc	agguugacga	agauaucgua	gcugugacgc	ggcacucucc	gucuauccau	2220
caauccgugg	uagccguguc	uaggacggcu	uuucgcaacc	ccaagacaag	uuucuauucc	2280
aaagaaguac	cccagaugug	uaucccgggg	aagauugagg	aagugguccu	ggaggccagg	2340
acgauagaac	ggaacacgaa	gccguaucgc	aaggacgaaa	acagcauaaa	cgggacgccu	2400
gacauacagg	uggagauucg	cgagcacauc	caguugaacg	aaucaaagau	agugaaacag	2460
gcuggggugg	cuaccaaggg	cccaaaugag	uacauucagg	agaucgaguu	cgaaaaccug	2520
uccccgggca	gugugaucau	cuuuagaguc	ucuuuggauc	cgcaugccca	ggucgcugug	2580
ggcaucuugc	gaaauaccu	gacgcaguuu	ucuccacauu	uuaagagugg	cucccuggcu	2640

ULPI_041_01WO_SeqList_ST25.txt

guggacaacg ccgacccaau cuugaagauu ccguucgcuu cccuggcuuc acgucugacu	2700
cucgcugaau ugaaucaaa ucuuguauaga ugcgaauagc aggagaagga agauggaggc	2760
gguuuuuacg acaucccga uuggucggca cugaaauacg ccggucugca aggcuuugaug	2820
agcguacugg ccgagauaag accaaagaac gaucugggcc acccauuuug caacaacuug	2880
cgguccggcg auuggaugau ugacuauguc ucuaaccgac ucaucaguag aagcgggaacc	2940
auagcugagg ucggcaaaug gcugcaagcc auguucuucu auuugaagca gaucucccgga	3000
uaucugaucc ccuguuuuuu ugacgcaauc cucauuggcg ccuauaccac acuccucgac	3060
accgcuugga agcagauguc cucauuugug caaaacggua gcaccuuugu gaagcaucug	3120
ucgcugggua gugugcagcu cugcggcgua gggaaguucc ccaguugcc gauccugagu	3180
ccagcucuaa uggacgugcc auauaggcug aacgagauua ccaaggaaaa ggagcagugu	3240
ugugucuccc uugcugccgg ucugcccau uuuuccucgg gcauuuuuag auguuggggg	3300
cgggacaccu ucaucgcucu gcggggcauc cugcugauca ccggcagaua uguggaggcu	3360
cggaaauua uucuggcuuu ugcugggacg cugcgccaug gacugauucc caaucuauug	3420
ggcgaaggca uauacgcuag guauaacugu cgggacgcug uuugguggug gcugcagugu	3480
auccaagauu auugcaagau ggucaccaac ggccuggaca uacugaaaug uccaguaucu	3540
aggauguacc cgacugacga cagugcuccc cugcccgcug gaacuuugga ucaacccug	3600
uuugagguua uucaagaggc uaugcagaag cacaugcagg gcauucagu ugcgaaagg	3660
aacgccggcc cccaaauuga ccguaacaug aaggacgaag gcuucaauau uacggcuggc	3720
guagaugaag aaacaggcuu uguguacggc ggcaaccgcu ucaacugugg gacauggaug	3780
gacaagaugg gcgaaagcga cagggcucgg aacaggggaa uuccagccac gccccgcgau	3840
ggcagugcug uugaaaauugu cggcuuguca aaauacagccg uacgcugguu auuagagcuc	3900
uccaagaaga acaucuuucc guaucaugaa gucacgguua agcggcaugg aaaagcaauc	3960
aaagugucuu acgacgaug gaauaggaaa auucaggaca acuuugagaa gcuguuccau	4020
gugucugagg acccguccga ucuuaacgag aagcauccca auuuggugca caagaggggc	4080

ULPI_041_01WO_SeqList_ST25.txt

aucuaauaagg acucauacgg ggcuagcuca ccuuggugug acuaucagcu gcgaccgaau	4140
uucacaauag cuauggucgu ggcuccggag cuuuuuacca ccgagaaggc uuggaaggcc	4200
cuggagauug ccgagaagaa gcuguuaggc ccguugggca ugaaaacguu ggauccagac	4260
gacauggucu auuguggcau cuaugacaac gcucucgaca augacaacua uaaccuagcu	4320
aagggcuuca auuaccauca gggcccggag uggcuguggc caauuggaua uuuccugagg	4380
gcaaagcugu auuucucccg gcugaugggc cccgagacaa cggcaaaaac cauagucuug	4440
guuaagaacg ugcugucccg ccacuaugug caccuugaaa ggaguccgug gaagggccug	4500
ccggagcuga caaacgaaaa ugcccaguac ugcccuuuuu cuugcgaaac ucaggcuugg	4560
agcaucgcca caaucuugga aaccucguau gacuuguag	4599

<210> 12
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 12	
augggacacu cgaagcagau ccgcauccuc cugcugaacg aaauggaaaa acuggaaaag	60
acacucuucc ggcuggaaca gggauaugag cugcaguuaa gacugggggc caccuuacaa	120
ggaaaggccg ugacugucua uaccaacuac ccguuccccg gagaaacuuu caaccgggaa	180
aaguuccggu cacuggacug ggaaaacccg accgagcggg aggacgauag cgacaaguac	240
ugcaagcuga accuucagca guccggaucg uuccaguacu acuuccugca agggaaacgag	300
aagucgggcg gaggauacau uguggucgau ccuauccugc gcgugggagc ugacaauac	360
gugcucccuc uggacugcgu gacgcugcaa accuuccucg caaagugccu cggaccuuuc	420
gacgaauagg aauccagacu ccgcguggcc aaggagucag gauacaacau gauccacuuc	480
acgccgcugc agaccuuggg acucucucgg agcuguuacu ccugggccaa ccagcuggaa	540
cugaacccag acuucucacg cccgaauaga aaguacaccu ggaacgacgu gggacagcuu	600

ULPI_041_01W0_SeqList_ST25.txt

guggagaagc ucaagaagga auggaacguc aucugcauaa cgcacgucgu guacaaccac	660
accgcagcca auucuaagug gauucaggaa caccgccgaau gcgccuacaa ucugguuaac	720
ucaccccacc uuaagcccg cuggguccua gacagagccc ucuggcgguu cucgugugac	780
guggccgagg gaaaguacaa ggagaaggga aucccccgcuc ugauugaaaa cgaccaccac	840
augaacucca uccgcaagau uaucugggag gacaucuucc cgaagcucaa gcugugggaa	900
uucuuccaag uggacgugaa caaggcgguc gagcaguuca gacgccugcu gaccaggaa	960
aaucggagag ugaccaagag cgauccuaac cagcaccuga ccaucaucca agaccagag	1020
uaccggcggu ucggaugcac uguggauaug aacaucgccc ugaccacuuu caucccucac	1080
gacaagggac cggccgcaau ugaggaaugc ugcaacuggu uccaauagcg gauggaggaa	1140
cugaacagcg agaagcaucg acugaucaau uaucaucaag agcaggcugu gaauugccug	1200
cugggaaacg uguucuaacga acggcuggcc ggacauggcc cuaagcuggg gccugugacc	1260
cggaagcacc cucuugugac ccgauacuuc accuucccg uugaagaaau ugauuucucc	1320
auggaagaau ccaugaucca ucugccaaac aaggccugcu uccuuauggc ccacauggc	1380
ugggucaugg gggacgaccc ucuucggaac uucgccgaac caggagcgga gguguaccuc	1440
agaagggagc ucaucuguug gggggauucc gugaagcuca gauacggaaa caagccagaa	1500
gauugccccc accuuugggc ccacaugaag aaguacaccg aaauaccgc cacauacuuc	1560
caaggagugc ggcuggacaa cugccauuca acucuccugc acgucgccga guacaugcug	1620
gacgcugcga gaaacuugca gcccaccuu uacguggugg ccgagcuguu caccgggagc	1680
gaggaccugg acaacguguu ugugaccagg cucggaauca cgucgcugau ucgcgaagcc	1740
augagcgccu acaacuccca cgaagagggu agacuggugu acagauacgg aggagagcca	1800
gugggauccu ucguccaacc gugccugcgg ccgcucaugc cugcgauccg acacgcgcug	1860
uucauggaca ucacccacga uaacgaangu ccuaucgugc auaggagcgc cuaugaugcc	1920
cuuccucca ccaccaucgu guccauggcg ugcugugccu cggguagcac caggggauac	1980
gacgagcugg ugccgcacca gaucucggug guguccgaag aacgguuuua cacuaagugg	2040

ULPI_041_01WO_SeqList_ST25.txt

aacccugagg cgcugccuuc caacaccgga gaagugaacu uccaguccgg uaucauugcc	2100
gcucgcugcg caaucagcaa acugcaccag gagcuuggug ccaaaggauu cauccaaguu	2160
uacgucgauc aaguggacga ggauauugug gccgucacua ggcacucucc aagcauucac	2220
caguccguag uggcaguguc gaggaccgcc uuccggaacc ccaagacuuc auuuuacucg	2280
aaagaggucc cacagaugug caucccugga aagaucgaag aaguggugcu ggaagcccgg	2340
accaucgaga ggaacacaaa gcccuaccgg aaggacgaga acuccaucaa cggaaccccc	2400
gacauuaccg uggaaaauag agagcacau caguugaacg agucgaagau cgugaagcag	2460
gccggagugg ccacuaaggg accaaacgag uacauccagg agaucgaguu ugaaaaccug	2520
uccccgggcu ccgugaucau cuuccgggug ucccuugacc cccaugccca aguggccguc	2580
ggaauccuua ggaaccaccu gaccaguuu ucgccccauu ucaaguccgg auccuuggcu	2640
gucgacaaug ccgaucccau ccugaaaauu ccguucgccu ccucgcuuc ccggcucacc	2700
cuugccgaac ucaaccagau ucuguaccgc ugcgagucag aagaaaagga agauggaggg	2760
ggaugcuacg acaucccgaa cugguccgcc cuuaauuacg ccgggcugca aggccugaug	2820
uccgugcugg cggagauuag accgaagaac gacuuggguc accuuuuug caacaacuug	2880
cggagcggag acuggaugau cgacuacgug ucgaaccggc ugauuagucg guccggaacc	2940
aucgccgaag ucggaaagug gcuccaagcc auguucuucu accugaagca aaauccccgc	3000
uaccugauac cgugcuacuu cgacgccauc cucauuggag ccuacaccac gcugcuggac	3060
accgccugga agcaaaugag cuccuucgug caaaacggau caaccuucgu gaagcaccug	3120
ucccugggua gcguccagcu cuguggcgug ggaaaguuu cgucccugcc aaauucgagc	3180
ccggcucuga uggaugugcc guaucgccug aacgagauca cgaaggagaa ggagcagugu	3240
ugcgugucgc uggcugcggg acugccucac uucucguccg gaaucuuccg cuguugggga	3300
cgcgacacgu uuauvcgguu gagggguuuu cuccucauuu ccggacgcua cguggaagcg	3360
cggaacauua uccuggcguu cgccggaacc cugcgccacg gucugauucc uaaucugcug	3420
ggagagggaa ucuacgcgcg guacaacugc cgggaugcug ugugguggug guugcagugc	3480

ULPI_041_01W0_SeqList_ST25.txt

auccaggacu auuguaaaau ggugccgaac ggccuggaca uccugaagug cccggugucc	3540
cggauguacc cgaccgauga uucagcacca cugccccgccg ggaccucgga ccagccccug	3600
uucgaaguca uucaggaagc gaugcagaag cauaugcagg gaauccagu cgcgaaaga	3660
aacgccggac cucagaucga ccgcaacaug aaggauagg gcuucaacau caccgcggga	3720
guggaugagg aaaccggcuu cgucuacgga gggaaccggu ucaacugcgg aaccuggaug	3780
gacaagaugg gagaguccga ccgcgcaaga aaccgcggua uuccugccac cccgcgggac	3840
ggaagcgcgg uggagaucgu cggacugagc aagagcgag ugcgcuggcg gcuggaguug	3900
uccaagaaga auauuuucc cuauacgag gucaccguga aacgccacgg gaaggccauc	3960
aaaguguccu acgaugagug gaaccgcaag auucaggaua acuucgaaaa gcuguuucau	4020
guguccgagg auccuccga ccucaacgaa aagcaccga accucgugca uaagaggggg	4080
aucuacaagg acagcuacgg ugccuccuca ccuuggugcg auuauagcu cgcgccgaac	4140
uucaccauug ccaugguggu cgccccugaa cuguuuacua ccgagaaggc cuggaaggcg	4200
cuggagauug ccgaaaagaa gcuguuggga cccucgggca ugaaaacccu cgacccccgac	4260
gacauggugu acugcggaau cuacgacaac gccucgaca acgacaacua caaccucgcg	4320
aagggguuca acuaccacca gggccccgaa uggcuguggc ccaucggaua cuuccuccgg	4380
gcgaaguugu acuucucccg ccugaugggc ccugaaacca cggccaaaac gaucgugcuc	4440
gugaagaacg ugcugucgag gcauuacgug caccucgagc ggagcccgug gaaggggcug	4500
ccggaacuga ccaacgagaa ugcacaguac ugccccuucu ccugcgaaac gcaggcuugg	4560
uccauugcca cuauuuugga aacgcuguac gaccuguag	4599

<210> 13

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 13

ULPI_041_01W0_SeqList_ST25.txt

augggacacu cgaacagau cagaauccug uugcucaacg aaauaggagaa gcuggaaaag	60
acccucuuuc ggcucgagca gggcuacgag cugcaguucc gccugggucc gacacuacaa	120
ggaaaggcag ugacugugua caccaacuac ccauuucccg gagaaaccuu caaccgggag	180
aaguuccggu cccuggacug ggaaaacca accgaacgag aggaugacuc cgacaaguac	240
ugcaagcuga accuccaaca guccgguuca uuccaguacu acuuucugca agggaacgag	300
aaguccggag gcggcuacau cguggucgac ccgauacuua gagugggagc cgacaaucau	360
guccugccuc uggacugcgu gaccugcaa accuucugg ccaaugucu gggcccguuc	420
gaugaguggg aaagccgccu cagagucgca aaggaguccg gauacaacau gauucacuuc	480
acuccgcugc aaaccucgg ucuguccgg ucgugcuauu cucuggcgaa ccagcuggag	540
cuaaaccg acuuucgcg ccaaaccgg aaguacaccu ggaacgacgu cgggcagcuu	600
guggaaaagc ugaagaagga guggaacgug aucugcauca cugacguggu guacaaccu	660
acggccgcca acucgaagug gauccaagag caucccgagu gcgcguacaa ccucgugaac	720
agcccgcuauc ugaagccugc uugggugcug gauagagccc ucuggagauu cagcugcgac	780
guggccgagg ggaaguacaa agaaaaggga auuccggccu ugauugagaa cgaccaucac	840
augaacucuaa uccgcaagau caucugggag gauaucuucc cuaagcuuaa guugugggag	900
uuuuuccaag uggacguaaa caaggcagug gagcaguuca gaagguugcu gacucaagaa	960
aacagacgcg ucacuaaguc cgauccuaac cagcaccuua ccaucauua agaccugag	1020
uaccgccggu uuggcugcac cgucgacaug aacaucgccc ugaccacuuu caucccgcau	1080
gacaagggcc cggcggcaau cgaggaaugc uguaacuggu uucauaagag gauggaggaa	1140
cugaauuccg agaagcaccg gcugauuaac uaccaccagg aacaggcagu gaauugccuc	1200
cugggcaacg uguuucacga acggcuggcu ggacacggac cgaagcuggg ucccugacu	1260
cgcaagcauc cgcucgugac ucgcuacuuc accuucccg uugaggagau ugacuucucc	1320
auggaagaau ccaugaucca ccuccgaac aaggcuugcu uccugauggc gcacaacgga	1380
ugggucaugg gcgacgaccc acugcgcaac uucgcugagc cuggcucgga ggucuaccug	1440

ULPI_041_01W0_SeqList_ST25.txt

agaagggaa	u	g	a	u	u	g	c	u	g	g	g	a	c	u	c	g	u	c	a	a	g	c	c	c	g	a		1500	
gauugcccc	a	c	c	u	g	u	g	g	g	c	u	c	a	u	g	a	a	g	a	a	g	a	a	c	u	g	c	1560	
cagggagucc	g	g	c	u	g	g	a	c	a	a	u	g	c	c	a	c	u	c	a	c	c	c	c	c	u	c	c	1620	
gaugcagcga	g	g	a	a	c	u	g	c	a	g	c	a	a	c	u	g	u	a	c	g	u	g	u	g	c	a	g	1680	
gaggaccucg	a	c	a	a	c	u	g	u	u	c	g	u	a	c	c	a	g	a	c	u	c	g	a	c	u	c	g	1740	
augucggccu	a	c	a	a	c	u	c	c	a	u	c	c	a	u	g	a	a	g	a	g	g	g	c	g	c	g	a	1800	
gugggaagcu	u	c	g	u	g	c	a	g	c	c	u	g	g	c	c	g	g	a	u	g	c	u	g	c	g	c	g	1860	
uucauggaua	u	c	a	c	u	c	a	c	a	g	a	a	c	u	g	c	a	a	c	u	g	c	a	a	c	u	g	1920	
cugccuucga	c	c	a	c	u	a	u	c	g	u	g	u	c	c	a	u	g	g	c	c	u	g	c	g	c	c	u	1980	
gaugaacucg	u	g	c	c	g	c	a	c	a	g	a	u	c	a	c	g	u	g	u	c	c	a	a	g	a	u	c	2040	
aaccccgaa	c	c	c	u	g	c	c	g	a	g	a	u	a	c	c	g	g	g	a	a	g	u	a	a	c	u	c	2100	
gcucgcugug	c	c	a	u	c	a	g	c	a	a	a	c	c	a	a	g	a	g	c	u	c	g	g	u	g	c	2160		
uacguggauc	a	g	g	u	c	g	a	c	a	g	a	u	a	u	g	u	g	g	c	c	g	u	g	a	c	c	a	2220	
caauccgucg	u	c	g	c	c	g	u	g	u	c	c	g	a	c	u	g	c	g	a	a	c	c	a	a	g	a	c	2280	
aaagaagugc	c	a	c	a	g	a	u	g	u	g	u	a	u	c	c	c	g	g	a	a	a	a	u	c	a	a	g	2340	
accuuugaga	g	g	a	a	c	a	c	a	a	a	g	c	c	u	a	c	c	g	a	a	g	a	c	a	g	a	a	2400	
gauauuacug	u	g	g	a	g	a	u	c	c	g	a	a	c	a	c	a	u	c	a	g	c	a	a	a	a	g	a	2460	
gcuggagugg	c	c	a	c	a	a	g	g	g	a	c	c	a	a	c	g	a	g	u	a	c	c	a	g	a	g	a	2520	
uccccuggcu	c	c	g	u	g	a	u	u	a	u	c	u	c	c	g	g	g	u	g	u	c	c	u	g	g	a	c	2580	
ggaauucuca	g	a	a	a	c	c	a	c	c	u	g	a	a	c	c	u	g	a	a	c	c	a	c	u	u	a	a	g	2640
guggacaacg	c	c	g	a	u	c	c	g	a	u	c	g	a	a	g	a	u	c	c	c	u	u	c	g	c	a	u	2700	
cucgcggaac	u	g	a	a	c	c	a	g	a	u	c	c	g	a	a	c	g	u	g	a	a	u	c	a	g	a	a	g	2760
ggcuguuacg	a	u	a	u	c	c	c	a	a	a	u	g	g	u	c	g	g	g	c	u	u	g	a	a	a	u	c	2820	
ucugugcugg	c	g	g	a	a	a	u	c	c	g	a	a	a	c	c	g	a	a	a	c	g	a	a	a	c	u	g	2880	

ULPI_041_01WO_SeqList_ST25.txt

cggagcggag acuggaugau cgauuacguc agcaacagau ugaucagccg gagcggcacu	2940
aucgccgagg ucggaaagug gcuccaggcc auguucuucu accugaagca gaucggccga	3000
uaccucaucc ccuguuacuu cgacgccauu cugaucgggg ccuacaccac ccugcuggac	3060
accgccugga agcagaugag caguuuugug caaacgggu ccaccuucgu gaagcaccuu	3120
ucacugggcu cagugcagcu cugcggcgug ggaaaguucc ccucgcugcc cauucugagc	3180
ccggcccuga uggacguccc uuaccggcug aacgagauca ccaaggagaa ggagcagugc	3240
uguguuuccc uggcugccgg gcugccacac uucucguccg gcaucuuccg gugcuggggc	3300
cgggauaccu ucauugcccu gcggggauc cugcuuauca ccggucgcu cguggaggcu	3360
cggaaacauua uucuggcguu cgccggcacc cuuagacacg gucugauucc gaauuuuug	3420
ggcgaaggaa ucuacgccag auacaacugu cgggacgccg ugugguggug gcuccagugc	3480
auucaggacu auugcaagau ggugccgaac ggccuggaca uccugaagug cccagugucg	3540
aggauguauc caaccgacga cagcgcaccu cugccggccg ggaccucga ccaaccucg	3600
uucgaaguca uucaagaggc uaugcagaag ccaugcagg guauucagu ccgggagcgg	3660
aacgcggggc cccagauuga uaggaacaug aaagacgagg gcuucaacau cacugccggc	3720
guggacgaag aaaccgguuu uguguacgga ggaaacagau ucaacugcgg uaccuggaug	3780
gacaagaugg gagaguucca ucgcgcgcgc aacagaggga ucccggcaac cccgcgggac	3840
ggauccgcgg uggaaaauugu gggacugagc aagagcggc ugccgguggcu ccuggaacug	3900
agcaaaaaga acaucuucc cuaccacga gugaccguga agcggcacgg aaaggccauc	3960
aaagucucau acgaugaug gaauaggaag auccaggaua acuucgagaa gcuguuucac	4020
guguccgagg auccuccga ucugaacga aagcauccga aucucgugca caagcgcggg	4080
aucuacaagg acucguacgg agcguccucc ccuuggugcg acuaucagcu gcggccuac	4140
uucaccauug ccauggucgu ggccccggag cuguucacaa cugagaaggc cuggaaggcc	4200
cuugaaaauug ccgagaagaa gcugcugggg ccuuugggga ugaaaacccu ggauccggac	4260
gacauggugu acugcggaau cuacgacaac gcccuggaca acgaauacua caaucuggcg	4320

ULPI_041_01WO_SeqList_ST25.txt

aagggcuuca auuaccacca gggcccggaa uggcucuggc cuauugggua cuuccugcgc	4380
gccaagcugu acuucucacg gcugauggga ccagagacua ccgccaagac uaucguccuc	4440
gugaagaacg ugcugucccg gcacuacgug caucuggaga ggagcccuug gaagggacuu	4500
ccugagcuga cgaacgaaaa cgcgcaguac ugcccuucu ccugcgaaac ccaggcuugg	4560
uccauugcca cuauacugga aaccuuauau gaccuguag	4599

<210> 14
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 14	
augggucacu ccaagcaau cagaauaua cuccuaacg aaauaggaaa gcuugaaaag	60
acacucuucc gcuuggagca gggcuacgag cugcaguucc ggcuugguucc gacacuucag	120
gggaaagccg ugaccgugua uaccaacuac cccuuuccgg gagaaacauu caaucgcgag	180
aaguuccggu cgcuggacug ggagaacccu accgagaggg aggacgacuc ugauaaguac	240
ugcaagcuga aucuccagca aucagguucu uuccaauauu acuuccuua aggaaacgaa	300
aaguccgggg gcggcuacau uguaguggac cccauccuuc gggucggcgc ggauaaccu	360
gugcuucccc ucgacugcgu gacucuccaa acuuuccucg ccaagugccu gggaccuuc	420
gaugaauagg aaucccgccu gcgcguggcc aaggagagcg gcuacaacau gauucacuuc	480
acuccgcuuc aaacccuugg gcugucccgc uccugcuauu cccuugcgaa ccagcuggaa	540
cuuaacccgg acuucucucg gccgaacaga aaguacacu ggaacgacgu gggccagcuu	600
gucgagaagc ugaagaagga auggaacgug aucugcauca ccgacguggu guacaaccac	660
accgaggcaa acuccaagug gauccaggaa cauccugagu gcgcgauacaa ccucgugaac	720
aguccgcauc ugaagccugc cugggugcug gauagagccc uguggcgcuu cuccugcgac	780
guggcugagg gaaaguacaa ggaaaagggc auccugcgc ucauugaaaa cgaucaccac	840

ULPI_041_01W0_SeqList_ST25.txt

augaacagca uccgcaagau uauuugggag gacauuucc cuaagcucaa gcucugggag	900
uuuuuccaag uugacgugaa caaggccgug gagcaguuca gacggcuccu gacucaagaa	960
aacagacgcg ucaccaaguc cgaccuaac caacaccuca ccaucaucca ggaucccga	1020
uaccgccggu uuggcugcac uguggacaug aacauugccc ugaccaccuu cauuccgcac	1080
gacaaaggcc cggccgcgau ugaagagugc uguaacuggu uccacaagag aauggaggag	1140
cugaacuccg aaaagcauag acuuaucaac uaccaccagg aacaggccgu gaacugccug	1200
cugggaaacg uguuuuacga gagacucgcc ggacacgguc caaguuggg uccugugacc	1260
cgaaacacc cgcucgucac ccgcuacuuu accuucccg ucgaggagau cgacuucagu	1320
auggaggaga gcaugaucca ucucccaac aaggccuguu uccucauggc ccacaacggc	1380
ugggucaugg gcgacgacc ccugcgcaac uucgcugaac ccggcucgga aguguaccuu	1440
cggaggaggc uuauuguug gggcgacagc gugaagcuua gauacggcaa caagccagaa	1500
gauuguccgu accugugggc gcacaugaag aaguacaccg agaucaccgc gacguacuuu	1560
caaggagugc ggucgauaa cugccacucc accccucugc acguggcaga guacaugcua	1620
gacgcggccc ggaaccucca gcccaccuu uacguggugg ccgaacucu cacugguucu	1680
gaggaucuug auaacguguu cgugacuagg cucggcauuu ccuccucau ccgggaagcc	1740
augucggccu auaacucca cgaggagggg cggcuggugu accgcuacgg aggcgaaccg	1800
gucggcagcu ucgugcagc guggcuccgc ccucugaugc ccgcuauugc ucacgcccuu	1860
uucauggaua ucacucacga uaaugagugc ccuaucgugc aucggagcgc cuacgacgu	1920
cucccuucca ccaccaucgu guccauggcg ugcugcgccu ccgguucaac caggggcuac	1980
gaugaacuug ugccgcacca gaucucaguc gucagcgagg aaagguucua cacuaagugg	2040
aaccugaag ccugcccuc uaacacgggc gaagugaacu uucagagcgg uaucauugcc	2100
gcuagaugcg caaucucca guugcaccag gaacugggag ccaagggguu cauccagguc	2160
uacguggacc aggucgacga ggacaucguc gccgugaccc ggcauuccc gagcauccau	2220
caguccgugg ucgccgugc acggaccgcc uuccgcaacc ccaagaccuc cuucuacucc	2280

ULPI_041_01WO_SeqList_ST25.txt

aaggaagugc cgcaaauug uaucccuggc aaaaucgagg aaguggugcu cgaagcgagg	2340
acgauugaga ggaauacuaa gccguacaga aaggacgaaa acuccaucaa cggcacccccg	2400
gacaucacug uggagaucgg ggagcacau cagcucaacg agagcaaaau ugugaagcag	2460
gccggcgucg cuacuaaggg cccaaacgag uacauccagg agauugaguu cgaaaaccuu	2520
agcccugggu cugugaucau cuuucgugug ucccucgacc cgcacgcaca ggucgcaguc	2580
gggauucucc ggaaccaucu gacucaguuc agccccacu ucaagagcgg cagccuugcc	2640
gucgacaacg ccgaucccau ccucaaaauc ccuuucgcau cccuugcguc gaggcuuacc	2700
cuggcggaau ugaaccagau ucuguaccgc ugcgagucgg aagaaaaaga ggauggcggc	2760
ggcugcuacg acauuccgaa cugguccgcc cugaaauacg cgggccuua gggccuuaug	2820
agcgucuugg ccgagaucgg cccaagaac gaccuggggc accccuuuug caacaaccuc	2880
agaagcggcg auuggaugau cgacuacgug ucgaacaggc ucaucagccg auccggcacu	2940
auagccgagg ucggaaagug gcugcaggcc auguucuuuu accucaaaca gaucccgagg	3000
uaccugaucc cgugcuacuu cgacgcuauu cucauuggcg ccuacacuac ccugcucgau	3060
accgcuugga agcagaugag cucauucgug caaacggaa gcaccuucgu gaagcaccuc	3120
ucccugggau cagugcagcu gugcggcgug ggaaaguucc cauccuuacc aaucucucg	3180
ccugcccuga uggacguccc uuauccgug aacgaaauc cgaaggagaa ggaacagugu	3240
ugugucucac uggcugccgg ccuccgcac uucucauccg gcaucuuccg gugcuggggu	3300
agagacacuu ucauugcgcu ccggggaauu cugcuuauc cuggccgcu cuggaagcc	3360
cgcaacauc uccuugccuu ugccgggacc cugcggcacg gccugauccc uaaccuucuc	3420
ggggagggca ucuaugcgcg auacaauugc cgggacgccg ucugguggug gcugcagugc	3480
auucaggacu auugcaagau ggugccaaau ggucuggaca uucugaagug uccagugucc	3540
cggauguacc cuaccgacga uagcgcccca cugcccgccg gaaccucga ucagccccug	3600
uucgaaguc uucaggaagc gaugcagaag cauauccagg guauccaau cgcgaaagg	3660
aaugccggcc caaaaucga cagaaauaug aaggaugagg gcuuaacau caccgccggc	3720

ULPI_041_01WO_SeqList_ST25.txt

guggacgagg agacugguuu cgucuacggc ggcaaucggu ucaauugcgg gaccuggaug	3780
gacaagaugg gcgaaagcga ccgggccaga aaccggggca uuccggcuac ccccgcgau	3840
ggcucggccg uggaaaucgu gggccucucc aagucugccg ugcgguggcu uuuggagcuc	3900
uccaagaaaa acaucuuccc guaccacgaa gugaccguga agagacacgg gaaggccauc	3960
aaaguguccu acgacgaug gaaccggaag auccaggaca acuucgaaaa gcuguuccac	4020
guguccgagg aucccuucca cuugaacgaa aagcacccca accucgugca caagcgcggc	4080
aucuacaagg auuccuacgg agcguccuca ccuuggugcg acuaucagcu gaggcccaac	4140
uucaccaucg caaugguggu ggccccugag uuguucacca cugagaaggc uuggaaggcc	4200
cuugagaucg ccgaaaagaa gcugcucggc ccgcugggga ugaaaacccu cgaccccgau	4260
gacauggugu acugcgggau auacgacaau gcacuagaca acgacaacua caaccuggcc	4320
aagggcuuca auuaccauca gggcccggag uggcuuuggc cuaucggcua cuuccugcgg	4380
gccaagcugu acuucucacg gcuuauggga ccggagacua cugcaaagac cauugugcuu	4440
gugaagaacg ugcuuucgcg ccacuacgug caucuggaac ggagccccug gaaggggcuc	4500
cccgagcuga ccaacgagaa cgcccaguac ugucccuucu ccugugaaac ccaggcuugg	4560
uccauugcca ccuucugga aaccuguac gaccucuag	4599

<210> 15

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 15

augggacacu caaagcaau ucggauccuc cugcucaacg agauggaaaa gcucgagaaa	60
accucuuuuc gccuggagca ggguaugag cuccaguucc gccugggucc gaccuccaa	120
gggaaggccg ucacugugua cacuaacuau ccuuccccg gagagacuuu caaccgggaa	180
aaguuccgca gccucgauug ggagaacccu acugaacggg aggacgauuc ggacaaguac	240

ULPI_041_01WO_SeqList_ST25.txt

uguaaacuga accuccagca gagcggcuca uuucaauacu acuuucugca agggaacgag	300
aaguccggag ggggguacau cgucguggau ccuaucuuuc gcgugggcg cgcacaaccau	360
gugcugccgu uggacugcgu gaccugcaa accuuccucg ccaaagccu cggaccuuuc	420
gaugaguggg aauccaggcu gagaguggcc aaggaaucgg gguacaacau gauucacuuc	480
accccucucc aaaccuggg ccugucucgg agcugcuacu ccugggcaa ccagcuggag	540
cugaaucucc acuuucuccg gcccaaccgg aaguacacuu ggaaugacgu gggccaacug	600
guggagaagc ucaagaagga guggaacgug aucugcauua ccgacguggu guacaaccau	660
accgccgcca acuccaagug gauccaggaa caccagaau gugccuacaa ccucgucaac	720
ucaccccacc ugaaaccagc augggugcug gaucgcgcc ucuggcgguu cucgugugac	780
guggccgaag gaaaguacaa agagaagggc auccugccc ugaucaaaa ugaccaccac	840
augaauucca uuagaaagau cauuugggag gacauuuucc cuaaguugaa gcucugggag	900
uucuuccaag uggacgucaa caaggccgug gaacaguuu gccggcuccu gaccaagaa	960
aaccgccg cg ugaccaaguc cgacccuac cagcaccuga ccuaaucca ggaccugag	1020
uacagacggu ucgguugcac ugucgacaug aacaucgccu ugacuacuuu cauccgcac	1080
gacaaggguc cugccgccau ugaggagugc ugcaacuggu uccacaagcg gauggaagaa	1140
cugaacuccg aaaagcaccg ccucaucaac uaccaccagg agcaggccgu gaacugccug	1200
cuggggaacg uauucucga gaggcuggcc ggacacggac ccaagcuggg acccgugacc	1260
agaaagcauc cucucgucac ccgcuacuuu acuuucccg uugaagagau cgacuucua	1320
auggaagagu cgaugaucca ucuccuac aaggccugcu uccugauggc ccacaacggc	1380
ugggugaugg gcgacgauc ucugagaaac uucgcugagc ccgguucgga aguguaccuu	1440
agacgggaac uuauuugcug gggcgauucc gugaaguac gcuacggaaa caagccugag	1500
gacugccuu accugugggc ccuaugaag aaguacaccg agauuaccgc caccuacuuc	1560
caaggggucc ggcuggacaa cugccacuca acuccucugc acguggcuga guacaugcug	1620
gaugcggccc ggaauuugca gccuaaccuu uacguggucg ccgaacuuu uaccgggucc	1680

ULPI_041_01WO_SeqList_ST25.txt

gaggaccugg	acaacguguu	cgugacucgg	cucggaauuc	ccucacugau	uagagaggcc	1740
auguccgc	acaacucgca	cgaagaaggc	cggcuggucu	aucgauacgg	cggcgaaccu	1800
gucggaagcu	ucguccagcc	cugccugcgg	ccgcugaugc	cagcgauccg	ccacgcccuc	1860
uucauggaca	ucacccauga	caacgaaugu	cccaucgugc	accgcuccgc	cuacgaugcc	1920
uugccaagca	ccaccaucgu	guccauggcg	ugcugcgcca	gcgguagcac	uaggggcuac	1980
gaugaacuug	ugccgcacca	gaucagcgug	guguccgagg	aaagguuuua	cacgaagugg	2040
aaccccgaag	cccugcccuc	caacacuggg	gaagugaacu	uccaguccgg	gaucauugcg	2100
gcccgcugcg	cgaucucgaa	gcuccaccag	gaacucggcg	cgaaggaauu	cauuaaguu	2160
uacguggacc	aggucgacga	ggacaucguc	gccgugacuc	gccacucucc	uucaauccau	2220
caauccgugg	uggcgguguc	gcggaccgcu	uuccggaacc	cuaagacuuc	guucuacucg	2280
aaagaagugc	cucagaugug	uaucuccgga	aagaucgagg	aaguggugcu	cgaggccagg	2340
acuauugaga	ggaauaccaa	gccuuaccgg	aaggauagga	acuccauuaa	cgguaucuccu	2400
gacaucaccg	ucgagaucgg	ggaacacauc	cagcucaacg	aaagcaaaau	cgugaagcag	2460
gccggcgugg	ccacuaaggg	uccgaacgag	uacauccagg	aaaucauuu	ugagaaccug	2520
ucgcccggaa	gcgugauuuu	uuuucgggug	ucccuggacc	cgcacgccc	agucgccgug	2580
gguaucugc	gcaaucaucu	gacucaguuc	agcccucacu	ucaaguccgg	cagccuugcc	2640
guggacaacg	ccgaucggu	ccugaagauc	ccauucgccu	cacucgccuc	acggcucaca	2700
cuggccgaac	ucaaucaugu	ccucuaucgc	ugugaauccg	aagagaagga	ggacggcgga	2760
ggauugcuacg	auauuccgaa	uugguccgca	cugaaauacg	ccggacugca	gggccucaug	2820
uccguguugg	ccgaaauccg	cccuaagaac	gaccucggcc	acccguucug	caacaaccuc	2880
agaucuggag	acuggaugau	cgauuacgug	ucaaacgcc	ugaucucgag	guccggcacu	2940
aucgccgaag	ucggaaagug	gcuccaagca	auguucuucu	accugaagca	gaucucucgc	3000
uaccugauac	cuuguuacuu	cgacgccauc	cucauuggcg	ccuacacuac	ucugcuggau	3060
acugccugga	agcaaaugag	cagcuucgug	cagaacggaa	gcacuuucgu	caagcaucug	3120

ULPI_041_01WO_SeqList_ST25.txt

ucgcucggga gcgugcagcu gugcggcguc ggaaaguuuu cuucccugcc cauucugucc	3180
ccugcccuca uggaugugcc guaccgccuu aacgagauca cuaaggagaa ggagcagugu	3240
ugcgugagcc uggcugccgg gcucccucac uucucguccg gaauuucag augcuggggc	3300
cgcgacaccu ucauugcccu gcgagggau cuguugauua cuggccgcu cugggaggcc	3360
aggaauauca uucucgccuu cgcgggaacc cugcggcacg gccugauccc uaaccuccug	3420
ggagaaggaa ucuaugcgag auauaacugc agggacgccg ugugguggug gcuccagugc	3480
auccaggacu acugcaagau ggugcccaac gggcuggaca uccucaagug ccccguguca	3540
cggauguacc caacugauga cagugcuccu cugccggccg guacucucga ccagccacuc	3600
uuugaaguca uccaggaggc caugcagaag cacaugcagg gcauucuuu cggggagagg	3660
aacgccggac cucagauuga ccggaacaug aaggacgagg guuucauuu cacugccggc	3720
guggacgaag aaaccggcuu cgucuacgga ggaaacagau ucaacugcgg aaccuggaug	3780
gauaagaugg gagaaagcga cagagcgccg aacaggggua ucccugccac ucccgggac	3840
ggaucggccg uggaauuugu gggacuucc aaguccgccg ugcgguggcu ccuggagcuu	3900
uccaagaaga acaucuucc uuaccauga gugaccguga agcggcacgg gaaggccauc	3960
aaagucuccu acgaugaug gaaccgcaag auucaggaca auuucgagaa gcuguuccau	4020
gucuccgagg acccuucuga ccugaaugag aagcaucca accuggugca caagagaggc	4080
aucuacaagg acagcuacgg agcuuccucc ccuuggugcg auuaccagcu ccggccaaac	4140
uucacuaucg ccaugguggu ggcgccugaa cuguucacca cugaaaaggc cuggaaggcg	4200
cucgaaaucg cggagaagaa gcugcucggg ccucucggga ugaaaacccu cgaccugau	4260
gauauggugu acugcggaau cuacgauaac gcccuagaca acgacaacua caaccuggcc	4320
aagggaauca acuaccaua ggggcccag agguuguggc cuauuggcua cuuccugaga	4380
gccaagcucu acuucucccg ccucaugggc cccgaaacca cugccaagac caucgugcuc	4440
gugaagaacg ugcucucccg gcacuaugug caccuugagc gcucgccaug gaagggacug	4500
cccgaacuga cuaacgagaa cgcccaguac ugccccuucu ccugugaaac ucaggccugg	4560

ULPI_041_01WO_SeqList_ST25.txt

uccaucgcca caauuuugga aaccucucac gaucuguag 4599

<210> 16
<211> 4599
<212> RNA
<213> Artificial Sequence

<220>
<223> Codon-Optimized AGL Coding Sequence

<400> 16
augggacauu caaagcaau cgggauucug uuacuuaug agauggagaa guuggagaaa 60
acaauuuuaa gacucgagca gggcuacgaa cugcaguuuc gccuggggccc uacacuacaa 120
gggaaagccg ucaccgucua uacaaacuau cccuuccccg gcgaaacguu uaaucgggag 180
aaguucagau cucuggauug ggagaacccc accgaacggg aggaugauag cgacaaguau 240
ugcaaacuga acuugcagca gucuggcagu uuucaguauu auuuucugca agggaaacgag 300
aaguccggcg gcggcuauau aguggucgac cccauccuuc gugugggagc ugacaaccuau 360
guguugcccc uggacugugu uacacugcaa accuuucugg cuaagugucu cgguccuuuu 420
gaugaauagg aaucgcggcu caggguggca aaagaauccg gcuauaaca gauucauuuu 480
acaccccugc aaaccucggg ccucucaagg aguuguuaca gccuugccaa ucagcuggag 540
cugaacccccg acuuuuuccg uccaaccgg aaguauacuu ggaacgacgu gggccagcuc 600
guagaaaagc ugaaaaagga guggaacgug aucuguauaa ccgauguggu cuauaaccac 660
accgccgcca auucgaagug gauccaggag caucccgagu gugcuuacaa ccucgugaac 720
agcccccauc uuaagccgc uuggguacuu gaccgcgcac uuuggagguu uuccugugac 780
guggcugagg ggaaguauaa ggaaaagggc aucccgccc ugaucgaaaa cgaccaucac 840
augaacagua uacggaagau cauuugggag gauauuuuuc ccaaacugaa gcugugggag 900
uuuuuucaag uggaugugaa caaggcuguc gaacaguuaa ggaggcugcu gacccaagaa 960
aauagacgcg ugacuaagag cgacccaaau cagcauuuga caauuauaca ggaccccgaa 1020
uaucgccgau ucggcuguac cguggauaug aaauaucgau ugacuacuuu caucacccac 1080

ULPI_041_01WO_SeqList_ST25.txt

gacaagggcc cgcgcgcuaugaggaugc uguaauuggu uucauaagcg aauggaggag	1140
cugaacucug agaagcacag gcugauuaac uaucaucagg aacaggccgu uaacuguuug	1200
cugggcaacg uguuuuauga acggcuugcc ggccauggcc cgaaguuggg cccgguuaca	1260
agaaagcauc cacuugugac ccgcuacuuc accuuuccgu ucgaagaaau ugacuuuagc	1320
auggaggagu caaugaucca ccuccccaac aaggcuugcu uucuuauggc ucacaacggc	1380
ugggugaugg gcgacgaccc ccugcggaac uuugcggaac cgggcucaga ggucuaucua	1440
agaagggaac ugauuugcug gggcgacucc gugaagcuga gauacgggaa caagccugag	1500
gauugccccc accucugggc ucacaugaaa aaguauacag aaauuacugc uacguauuuc	1560
caaggcguac ggcuugacaa cugucacucc accccacugc augucgccga guauauguua	1620
gaugcugcuc gaaacuugca gcccaauuug uauguggucg cugaacuguu uacggggagu	1680
gaagaucucg acaacgucuu uguaacccga cugggcauca gcucgcuaau ucgggaggcc	1740
auguccgcuu acaacuccca ugaggaaggc cgccuuguau aucguuauagg cggcgaaccc	1800
guggggagcu uuguucaacc gugucuacgc cccugaugc ccgccaucgc ucaugcccua	1860
uucauggaua ucacucauga caaugagugu ccuauugugc auaggagugc uuaugacgcc	1920
cucccaagca caacgaucgu guccauggcu uguugugcua guggcuccac aagaggcuau	1980
gacgaacugg ugccccauca aaucuccgua gugucugaag agagauuuua caccaagugg	2040
aaccccaggg cucucccuuc aaacacugga gagguuaacu uucaauccgg gauuauugcu	2100
gcuagaugug cuaucagcaa gcuccaucag gaacugggagc caaagggcuu cauucaaguu	2160
uauguagacc agguugacga ggacauaguu gcaguaacuc ggcauucccc uucgauacau	2220
cagucugucg uggccguguc caggacagca uuucgcaauc ccaagacuag cuuuuacucc	2280
aaagaaguac caaaaugug uauccccggg aagaucgagg aagugguacu ggaagcccgg	2340
acuauugaga gaaacaccaa gccguaucgg aaggacgaga auucuaucua cggcacaccu	2400
gauauaacag uggaaauacg cgaacacauu caguugaug aguccaagau cgugaagcag	2460
gccggcgucg ccaccaaggg cccaacgag uauauccagg agaucgaauu ugaaaaccug	2520

ULPI_041_01WO_SeqList_ST25.txt

agccccggcu	cuguuauuau	auuucggguu	ucauuagauc	cucaugcuca	aguggcugua	2580
ggcauccucc	ggaaucaucu	gacucaguuc	ucuccccacu	ucaagagcgg	cagccuggcu	2640
gucgauaaug	ccgacccccau	acuuuagauu	cccuuugcuu	cccuggcuuc	aaggcugacc	2700
cuggcugaac	uuaucaaaau	ccuauaccga	ugcgaaagcg	aagagaagga	agauggcggc	2760
ggcuguuacg	auaucccgaa	cuggagcgca	cugaaauaug	cuggcuuaca	aggccucaug	2820
aguguguugg	cugagauuag	acccaagaau	gacuugggcc	auccauuuug	uaacaaccug	2880
agaagcggcg	acuggaugau	cgauuacgug	ucuaaccgac	ucaucucccg	aagcggcacc	2940
auugcugagg	uuggcaaaug	gcugcaggcu	auguucuuuu	aucugaaaca	gauuccucgg	3000
uaccugauuc	ccuguuauuu	cgacgcuaau	cugauuggcg	cauauaccac	gcucuuggac	3060
acugcaugga	agcaaaugag	cucuuuuguc	caaaauggcu	ccacuuuugu	uaaacaucug	3120
aguuugggca	gcgugcaguu	auguggcggu	ggcaauuuuc	caagccugcc	cauacugucc	3180
cccgcucuga	uggacguccc	cuaccgacug	aacgagauca	ccaaggaaaa	ggaacagugc	3240
ugugugucuu	uagcugccgg	cuugccgcgu	uuuucagcg	gcauuuuccg	guguuggggg	3300
cgggacaccu	ucaucgcacu	gagaggcauu	cugcugauca	cuggccgcua	uguggaagcu	3360
aggaauauca	uuuuagccuu	ugcggguacc	cuucggcacg	ggcugauccc	caaucuuuug	3420
ggcgaaggca	ucuacgcacg	uuacaauugc	cgggaugccg	uaugguggug	gcuccaangu	3480
auccaggacu	auugcaagau	gguucccaac	ggccucgaua	uccugaagug	ccccguguca	3540
aggauguauc	cuaccgauga	cagugcuccc	cuuccugcug	gcaccugga	ucagccgcuc	3600
uuugaaguua	uucaggaagc	aaugcaaaag	cauauagcagg	gcauccaguu	ucgggaaaga	3660
aacgcuggcc	cgcaauuaga	caggaacaug	aaggauaag	gcuucaauau	uacugcuggc	3720
guagacgaag	agacaggguu	cgucuacggc	ggcaauaggu	uuuacugugg	cacuuggaug	3780
gacaagaugg	gcgaauauga	ccgcgcgagg	aacagaggca	uccagcaac	accgagggac	3840
ggcagcgug	uggagauugu	gggccugucu	aagucugccg	ugcgcuggcu	acucgaacug	3900
uccaagaaga	auaucuuucc	cuaucaugaa	gucaccguua	agcggcaugg	caaagcuauu	3960

ULPI_041_01WO_SeqList_ST25.txt

aagguuuccu augacgagug gaacaggaag auucaggaca auuuugagaa gcuguuccau	4020
gugucggagg acccuagcga ucucaacgag aagcacccca acuuaguaca uaagaggggc	4080
aucuauaagg auagcuauagg cgcuaagcagc ccuuggugug acuaucagcu ccgucccaac	4140
uuuaccuuug cuaugguagu ggcgcccagag uuguuuacca cugagaaggc uuggaaggcu	4200
cuugaaaauag ccgagaagaa auugcugggc cccugggca ugaaaacucu ggaucccgau	4260
gacaugguau acuguggcau cuaugacaau gcccuggaca augacaauua caaucuggcc	4320
aaggguuuuu acuaucacca gggcccugaa uggcuguggc cuauuggcua uuuucuucgc	4380
gccaagcuau auuuuaguag gcugaugggu ccagaaacaa cugcaaagac aaugugcuc	4440
gugaagaacg ugcuuucccg gcacuaugug caucuggaaa ggaguccaug gaaggguug	4500
ccggaauuga cuaacgagaa cgcccaguau uguccuuuu cuugugaaac ucaggccugg	4560
uccauugcua cuauucugga gacucuauau gacuuguag	4599

<210> 17

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 17

augggccauu ccaagcaau ucgaauucug cugcugaacg aaauggaaaa guuggaaaag	60
acauuguucc gucuggagca gggcuauag cugcaguuaa ggcuuggccc gacguugcag	120
ggcaaagccg ugaccguaua uacuaacuau cccuucccug gcgaaacguu uaaccgggag	180
aaguuucgcu cacuggacug ggagaauccc acggaaaggg aggacgauag cgauaaguac	240
uguaaacuga aucugcaaca gucuggcucu uucaauauu auuuucugca agggaacgag	300
aaguccgggg gcggcuauau uguaguggac cccaucuuga gagugggcgc ugacaaccu	360
guauuacccc uugauugcgu gacgcugcag accuuucugg cuaaguguuu gggcccuuc	420
gaugaguggg aaagccggcu gaggguggcc aaggagagcg gcuaauauau gauucauuu	480

ULPI_041_01WO_SeqList_ST25.txt

accccuuugc aaacucuggg gcucucuagg agcuguuauu ccuuggcuaa ucagcuggaa	540
cugaacccccg auuuuagccg gcccaaccga aaguauacau ggaugacgu ugggcaauug	600
guugagaagc ugaagaagga auggaauuu auauguauaa cagacuggu guauaaccau	660
accgcugcua auucgaaaug gauucaggaa cauccugagu gcgccuacaa ccucguaaa	720
aguccucauc ugaagccggc uugggugcug gacagggcuc ucuggagguu uagcugugau	780
guggccgaag ggaaguauaa ggagaagggu auccggcuc ugauugagaa cgaccaccau	840
augaacagua uaaggaagau aaucugggag gacauuuuuc ccaaacugaa gcugugggaa	900
uucuuucaag uggacgugaa caaggccguu gagcaguua gacguugcu gacacaagag	960
aaccgaaggg ucaccaagag ugaccuaac cagcaucuua cauuuaucca agaccccag	1020
uacgucggg uuggcuguac uguggacaug aacauugcuc ugaccacuu uauaccac	1080
gauaagggcc ccgcugcuau cgaagagugu uguaacuggu uccacaagag gauggaagag	1140
uugaacucug aaaagcaucg gcugauuaac uaucaucagg aacaggcugu gaacugccug	1200
cucggcaacg uauucuacga gagauuggcu ggccauggac cuaagcuggg cccgguuacg	1260
agaaagcauc cccugguuac gcgguauuuc accuucccg uugaggagau ugauuuuagc	1320
auggaagaau ccaugaucca ccuccccaac aaggcuugcu uccugauggc acauaacgga	1380
uggguuauug gcgacgacc ccugaggaau uuugcugaac ccggcuccga agucuaucug	1440
agaagggaa ugaucuguug gggcgacucc gucaaguugc gcuauaggaa uaagccugag	1500
gacuguccau aucugugggc ucacaugaag aaguacacag agaucaccgc uacauacuuu	1560
cagggcgua ggcucgaaa uugucacucc acaccucug acguggcuga guauaugcug	1620
gaugccgcua ggaauucca gccuaaccua uauguggucg cugaacuguu caccggcucu	1680
gaggaucugg auaauguguu ugugacacgc cugggcauca gcuccugau ccgcgaggcu	1740
augagugccu acaauaguc cgaggaaggc cggcuugugu accgcuaugg cggcgaaccc	1800
gugggcagcu uuguacaacc uugucugcg ccucugaugc ccgcuaugc ucacgcuug	1860
uuuauggaca ucaccauga caacgaugu ccuauuguac accggucugc uuacgacgu	1920

ULPI_041_01WO_SeqList_ST25.txt

cugccccucca caacuauugu caguauggcu uguugugcau cuggcucaac ucggggcuac	1980
gacgaauuag uuccgcauca aaucucugua guguccgaag aacgguuuua uacaaaaugg	2040
aaccccgaag cgcugcccuc caauaccggc gaagugaacu uucagucugg gaucaucgcc	2100
gcuagaugug cuauuuucuaa acugcaucag gaacugggcg cuaagggcuu uauucaagua	2160
uauugggacc aggucgacga ggacaucgua gcuguaaccc ggcacagccc cagcauacau	2220
cagagugucg uggcuguguc uagaacagcu uuuaggaacc caaagaccuc uuuuuacucc	2280
aaagagguuc cacagaugug uaucccgggc aaaaugagg aaguggucuu ggaggccagg	2340
acaauugagc guaauacgaa gcccuauaga aaggauagaga auagcaucuaa cggcacuccu	2400
gauaucacug uggaaauccg ggagcauua caguugaacg agucuaagau cguuaagcag	2460
gcugguguag cuaccaaggg ucccaaugag uauauucagg aaauugagu ugaaaaccug	2520
uccccgggca gcguauuuau cuuuagaguc agucucgacc cccaugcuca aguggcagug	2580
gggauccugc ggaaccuuu gacucaguuc ucccccauu uuaagucugg cagucuggcc	2640
guugacaacg cugaccccau auugaagauu cccuuugcuu ccugggcuuc acgguugacc	2700
cuggccgaac ugaaccagau uuuauaucgc ugcgaguccg aggaaaagga ggauggcggc	2760
ggcuguuaug acauucccaa cuggucugcu cuaaaauacg cugggcugca gggucugaug	2820
agugugcugg cugaaauucg cccaagaac gaccuggggc auccauuuug caacaucug	2880
cgcaguggcg acuggaugau ugauuaugug uccaaccggc ugauuagucg gagcggcacg	2940
aucgcagaag ucggcaaaug gcuccaggcu auguuuuuu accuaaagca gauaccaga	3000
uaccugaucc cuuguuuuu ugacgcuua cugauuggcg cuuacacaac cuugcuggac	3060
accgccugga agcagaugag cagcuuuguc caaaauggca gcacauucgu gaagcauuug	3120
ucucugggca gugugcagcu guguggcgug ggcaauuuc caaguugcc gauccugucu	3180
cccgcuuuga uggaugugcc auaucgacug aacgagauaa ccaaggaaaa ggaacagugu	3240
uguguuucac ucgcugcugg cuugcccau uuuuccucug gcauuuuccg auguuggggg	3300
agggacacau ucauugcccu gcgugggauc cugcugauua ccggccgcua cguugaggcu	3360

ULPI_041_01WO_SeqList_ST25.txt

agaaacauua uuuuggcuuu cgccggcaca uugagacaug gucugauacc caaucugcug	3420
ggcgaaggca uauaugcuag guauaacugu cgcgacgcag ugugguggug gcugcaaugc	3480
auccaggauu auugcaagau gguuccuaac ggccuagaca uccugaagug cccagugucc	3540
cggauguau caaccgauga uucagcucucc cugcccgccg guacacugga ccaaccccug	3600
uucgaaguua uucaagaagc caugcagaag cacaugcagg guauccagu ucgagagagg	3660
aaugcaggcc cccaaauuga ccggaacaug aaggauaga gcuucaacau uacugcuggc	3720
gucgacgaag aaacaggguu uguguacggc ggcaaccggu uuaacugcgg cacauggaug	3780
gacaagaugg gcgaaagcga ucgcgcacgg aaccggggca uccccgccac gccgcgugau	3840
ggcucugcug uggagaucgu cggccucagc aaauccgcug uccgauggcu guuagagcug	3900
ucgaagaaga acaucuucc cuaccaugag guuaccguua agagacacgg caaagcuau	3960
aaagugucuu augacgaug gaauagaaag auucaagaca acuucgagaa acuguuucan	4020
guguccgagg aucccagcga ccucaacgaa aagcauccca aucucguaca uaagaggggc	4080
aucuauaagg acucauauagg cgcaucaagu ccuuggugcg acuaucagcu ucggcccaau	4140
uucacgaucg caauggucgu ggcacccagc cuguuuacua cggagaaggc uuggaaggcu	4200
cuggaaaucg cugagaagaa gcugcugggc ccuugggca ugaaaacccu ggaucggac	4260
gacaugguau acuguggcau auaugacaac gcucuggaua acgauaacua uaaauuggcu	4320
aaaggcuua auuauacca gggcccagaa uggcuauagg ccuauaggcu cuuucugcga	4380
gcuaagcuau auuuuucucg ccugaugggc ccugaaacca ccgcuaagac uaucgugcuc	4440
gucaagaacg uacugaguag gcauuacgug cacuuggagc gcagcccaug gaaggguug	4500
ccggaacuga cgaacgaaaa cgcucaguau uguccguuuu cuugugaaac acaggcuugg	4560
aguauugcua cuauccugga aacucuguac gaucuguag	4599

<210> 18

<211> 4599

<212> RNA

<213> Artificial Sequence

ULPI_041_01W0_SeqList_ST25.txt

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 18

augggucauu ccaaacaaau uaggauccug cuguugaacg aaauggaaaa gcucgaaaag	60
accuuauuuuc gccuggagca gggcuauagag uugcaguuuuc ggcugggacc aacucugcaa	120
ggaaaggcug uaaccgugua cacaaauuau ccuuuucccg gcgaaacguu uaacagggag	180
aaguuu cgau cucuggauug ggagaacccc acugaaaggg aggacgauag cgacaaguau	240
ugcaaacuga aucugcaaca guccggcagc uuucaauuu auuuucugca agggaaacgaa	300
aagagcggcg ggggcuauau uguaguagac ccuauacuga gagucggagc ugacaaccuau	360
guucugcccu uggacugcgu gacacugcag accuuuuugg cuaagugucu gggcccccuu	420
gaugaauggg agagucggcu ccguguggcu aaggagagcg gguacaauau gauucacuuu	480
acccccugc aaacuuuggg acuuucccg agcuguuaua gccuggcaaa ccagcuggag	540
uugaauccug acuuuucccg cccaaauagg aaguauaccu ggaacgacgu gggccaacug	600
guggagaagc ugaagaagga guggaacgug aucugcauua ccgacgucgu guauaaucuu	660
accgccgcca acuccaagug gauacaggaa cauccugagu gugcuuacaa ccucguaaac	720
uccccgcauc ugaagccagc uuggguucug gauagagcuu uguggcgguu cagcugugac	780
gucgcugagg gaaaguacaa ggaaaaaggg auuccugcuc ucaucgagaa ugaccaccac	840
augaacagua uuaggaagau uaucugggag gacaucuucc cgaagcuaaa gcugugggaa	900
uucuuccaag uggacgugaa caaggccgug gagcaguucc gacggcuguu aacucaagag	960
aaccguaggg ucacaaauc ggauccuac cagcaccuga ccuuuaucca agaccggaa	1020
uauccggcgu ucggcuguac uguggacaug aacauugcuu ugacaacuuu cauuccccac	1080
gacaagggu cagcagcuau cgaggaaugu uguaaauuggu uucauaagcg gauggaggaa	1140
cucaacuccg agaagcauag gcugauuaau uaucaucagg aacaggcugu gaauugucug	1200
uugggcaaug uuuucucga acggcuggcu ggccauggcc ccaaacucgg accuguuaca	1260
aggaagcauc cacugguuac ccgauacuuu accuuuccau uugaagaaau ugauuuuagc	1320

ULPI_041_01WO_SeqList_ST25.txt

auggaagaau ccaugaucca ucuaccuaac aaggccuguu uucugauggc ccauaacggc	1380
uggguuaugg gagaugaucc cuugcggaac uuugcugaac ccggaucuga aguauaucug	1440
cggagagagu ugauuuuguug gggcgacucc guaaagcugc gcuauggcaa caagccagag	1500
gacugucccu auuuugugggc ucauauagaag aaguauaccg agaucaccgc uacguauuuu	1560
caaggcgua ggcucgauaa uugucauucc acuccgcugc auguugcuga guauauguug	1620
gacgcugcuc gaaaucugca gcccaauuug uauguugugg ccgagcuguu uaccggcucc	1680
gaggaccucg auaacguguu cgucacgcga cuaggcauca gcagcuugau ccgggaagcc	1740
auguccgcuu auaacuccca cgaggagggc cgacuggucu accgcuaugg cggcgaacca	1800
gugggcaguu uuguacagcc uugucugagg ccccucaugc ccgcuauugc ucaugcucug	1860
uucauggaca uuacucauga uaacgagugc ccuauagugc aucgguccgc cuacgacgcc	1920
cugccgagca cuacaauagu guccauggcu uguugugcaa gcgggagcac ccgcggcuau	1980
gacgagcugg ugccgcauca aaauaucgua gucagcgaag aaagguuuua uaccaagugg	2040
aacccggaag cuuuaccuuc caauacuggg gaagugaacu uucagagcgg cauuaucgcu	2100
gcgagaugug cuauauccaa guugcaucag gaacugggggg caaagggguu uauucaagua	2160
uauguagacc agguugacga agauauagug gcugugacac gccacagccc aagcauccac	2220
caguccgugg uggcugucag ccggacugcu uuucgcaacc caaagacaag cuuuuauagc	2280
aaagaagugc cgcagaugug cauucccggc aaaaugagg aggucgugcu ggaggcuagg	2340
acuaucgaaa ggaacaccaa gccguacagg aaggacgaga auuccaucaa cgggacucca	2400
gauauuacag uugagaucg ggaacauauu caguugaau agucgaagau uguuaagcag	2460
gcuggcguag cuacaaaggg gccaaacgag uauauucaag agauagaauu cgaaaaccug	2520
agccccgguu ccgugaucan uuuucgagug ucccuggauc cucaugcuca aguggccguu	2580
ggcauucuga gaaaccaucu cacacaauuu uccccucacu uaaaaagugg cagccuggc	2640
guggacaacg ccgauccgau ccugaagauu ccuucgcuu ccugggcuag ucgccugaca	2700
cuggcugaac uaaaccagau ucuuuaccgc ugugaaucug aagagaagga ggacggcggc	2760

ULPI_041_01WO_SeqList_ST25.txt

gguugcuaug acaucccaaa uuggagcgcu cugaaauacg cugggcugca gggccuacug	2820
agcguacugg cagaaaucag acccaagaac gaccuggggcc accccuucug uaacaaccug	2880
agguccggcg auuggaugau cgacuaugug ucgaacagac ucaucucaag aagcggcacu	2940
auagcugagg ucggcaaaug guugcaggcu auguucuuuu aucugaagca aaauccgcgg	3000
uaucugaucc cgugcuauuu ugacgcuaua ugaucggcg cuuauacgac uuugcuggac	3060
acagccugga agcagauguc cagcuucgug cagaacggca gcacuuuugu gaagcaccug	3120
agucuggggcu caguacagcu guguggcgug ggcaaaaucc ccagccuacc aaucugucc	3180
cccgucucuga uggauguucc cuaucggcug aacgagauua ccaaggaaaa ggagcaaugu	3240
ugcgugagcc uggcagccgg ccugccccac uuuucuuucc gcauuuuucg cuguugggggu	3300
cgggacaccu ucauugcacu gcgggggauc cuguuaauca ccggccgcuu uguggaagcu	3360
agaaaauuaa ucuuggcuuu ugcuggcacu cugcgacacg gccugauccc caaccuucug	3420
ggcgaaggca ucuaugcucg cuacaauugu cgcgacgccg ucugguggug gcugcagugu	3480
auacaggacu auugcaagau ggugccaaau ggccuggaca uccuuaaaug uccaguucc	3540
cggauguacc caacugacga cagcgucuccu uugccugccg gcacacugga ucagccgcug	3600
uuugaaguga uccaggaagc uaugcagaag cauauagcagg gcauucaguu ucgggaacga	3660
aaugcuggcc caaaauuga cagaaacaug aaggauagg guuuuaacau uacugcuggc	3720
gucgaugaag agacaggguu uguguacggc ggcaacaggu ucaacugcgg ccauggaug	3780
gacaagaugg gcgaauauga ccgugcuagg aacaggggaa uccccgcuac cccccgggac	3840
ggcagcgcuug uggaauucgu aggccugucc aaguccgccg uucgcuggcu ucuggagcug	3900
agcaagaaaa acaucuuucc cuaucacgag gucacaguga agagacaugg caaagcuau	3960
aaagucucuu acgaugagug gaauagaaag auucaagaua acuuugaaaa guuguuccac	4020
guuucggagg accccagcga ucugaauag aagcauccca acuuuggugca caagaggggc	4080
auauacaagg acucuuauug ugcuaagcag ccuuggugcg acuaucagcu gaggcccaac	4140
uuuacuaucg ccaugguggu ggcaccggaa cuguuuacaa cggagaaggc uuggaaggcu	4200

ULPI_041_01WO_SeqList_ST25.txt

cuggaaaauug cugagaagaa guugcuggggg ccccuggggga ugaaaacgcu ggauccagac	4260
gacaugggucu acugugggau auaugauaau gcucugggaua acgauaaaua caaucuggcu	4320
aaggguuuca acuaucauca ggggccagag uggcucuggc cuauuggcua uuuccuccgu	4380
gcuaagcugu acuuuucucg ccugaugggg ccagaaacua ccgcaaagac cauuguauug	4440
gucaagaacg ugcugagccg gcauuaugua cauuuggaac ggucuccuug gaaaggccug	4500
cccgaacuca cuaacgagaa cgcucaguau uguccguuuu ccugcgaaac ccaggcuugg	4560
uccaucgcaa cuauuuugga aaccucguac gaccuguag	4599

<210> 19
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 19	
auggggccaca gcaagcaau cagaauccuc cugcucaacg agauggaaaa gcucgaaaag	60
accuguuua gacuggagca gggauacgag cugcaguucc gccucggacc aaccuccaa	120
gggaaggccg ugaccgugua caccaacuac ccauucccg gagaaaccuu caaccgggaa	180
aaguucagga gccuggacug ggaaaacccc accgagagag aggacgacag cgacaaauac	240
ugcaagcuga accuccaaca gagcggcagc uuccaguacu acuuccucca aggaaacgaa	300
aagagcgggg gcggcuacau cgugguggac ccaauccugc gggucggggc agacaaccac	360
gugcugccac uggacugcgu gacacugcag accuuccugg ccaagugccu gggaccguuc	420
gacgaauggg agagccggcu gcggguggcc aaggagagcg gcuacaacau gauccacuuc	480
acaccgcucc aaaccucgg ccugagcaga agcugcuaca gccuagccaa ccagcuggag	540
cucaaccggg acuucagcag gcccaacaga aaguacaccu ggaacgacgu gggacagcuc	600
guggaaaagc ugaaaaagga auggaacgug aucugcauca ccgacuggu guacaaccac	660
accgccgcca acagcaaaug gauccaggaa caccggaaug gcgcguacaa ccucgugaac	720

ULPI_041_01W0_SeqList_ST25.txt

agccccacc ugaagccggc cuggguacug gacagggcac uguggcgcuu cagcugcgac	780
guggccgaag gaaaguacaa ggagaagggc aucccgccc ugaucgagaa cgaccaccac	840
augaacagca uccggaagau caucugggag gacaucuucc caaagcugaa gcugugggag	900
uucuuccaag ucgacgugaa caaggccgug gaacaguuca gacgccugcu gacgcaggag	960
aaccggaggg ucaccaagag cgacccaac cagcaccuga ccaucaucca agacccggaa	1020
uaccggagau ucggaugcac cguggacaug aacaucgccc ugaccacuu cauccacac	1080
gacaaggggc ccgcagccau cgaggagugc ugcaacuggu uccacaagag aauggaggaa	1140
cugaacagcg aaaagcaccg ccucaucaac uaccaccagg aacaggccgu gaacugccuc	1200
cucggaaacg uguucuacga gcggcuggcc ggccacgggc ccaagcuggg ccccgugacc	1260
cgcaagcacc cgcucgugac gcgguacuuc accuucccg ucgagaauu cgacuucagc	1320
auggaggaga gcaugaucca ccuccgaac aaggccugcu uccucauggc ccacaacggc	1380
ugggucaugg gcgacgacc gcugcggaac uucgcagagc ccggaagcga aguguaccuc	1440
agaagggagc ugaucugcug gggagacagc gugaagcugc gcuacggaaa caagcccag	1500
gacugcccgu accugugggc gcacaugaag aaguacaccg agaucaccgc caccuacuuc	1560
caaggagugc ggcuggacaa cugccacagc acccgccugc acguggccga guaucugcug	1620
gacgccgcca gaaaccucca gcccaaccuc uacguggugg cagagcuguu caccgggagc	1680
gaggaccucg acaacguguu cgugacccga cucggcauca gcagccugau ccgggaagca	1740
augagcgccu acaacagcca cgaggaaggg aggcuggugu acagauacgg aggagaaccc	1800
gugggcagcu ucgugcagcc gugccugagg cccugaugc cagccaucgc gcacgcgug	1860
uucauggaca ucaccacga caacgaugc ccgaucgugc accggagcgc auacgacgcc	1920
cugccgagca ccacgaucgu cagcauggcc ugucgcgcca gcggcagcac ccgaggauac	1980
gacgagcugg ugccccacca aaucagcguc gucagcgaag aacgcuucua caccaagugg	2040
aacccggaag cacugccgag caacaccgga gaagugaacu uccagagcgg aaucaucgcc	2100
gcgcgcugcg cgaucagcaa acugcaccag gaacugggag ccaagggguu cauccagguc	2160

ULPI_041_01WO_SeqList_ST25.txt

uacgucgacc aaguggacga ggacaucgug gcagugaccc gccacagccc cagcauccac	2220
caaagcgugg uggccgugag ccggacagcg uuccggaacc ccaagacgag cuucuacagc	2280
aaagaggugc cccagaugug caucccgga aagaucgagg aaguggugcu cgaagcgcg	2340
accaucgaac gcaacaccaa accguaccgc aaggacgaaa acagcauaaa cgggaccccc	2400
gacaucaccg uggagaucag ggaacacauc cagcugaacg agagcaagau cgugaagcag	2460
gccggggugc ccaccaaggg cccgaacgag uacauccagg agaucgagu cgagaaccuc	2520
agccccggga gcgugaucu cuucagaguc agccuggacc cacacgcca aguggccgug	2580
ggcauccucc ggaaccaccu gaccagauuc agcccgcacu ucaagagcgg gagccucgcc	2640
guggacaacg ccgacccgau ccugaagauc ccguucgca gccuggccag ccggcucacc	2700
cuggccgaac ugaaccagau ccuguaccgc ugcgagagcg aagaaaagga ggacggagga	2760
gggugcuacg acaucccaa cuggagcgcg cugaaauacg cgggccugca gggccugaug	2820
agcgugcugg ccgaaauccg cccaagaac gaccugggac acccauucug caacaaccug	2880
aggagcggag acuggaugau cgacuacguc agcaacagac ugaucagccg cagcggcacc	2940
aucgccgaag ucggaaaug gcuccaggcc auguucuuc accugaagca gaucgcg	3000
uaccugaucc cgugcuacuu cgacgccauc cugaucggcg ccuacaccac ccugcucgac	3060
accgccugga agcagaugag cagcuucguc caaacggga gcaccuucgu gaagcaccuc	3120
agccugggca gcgugcagcu gugcgggguc ggaaaaucc cgagccugcc cauccugagc	3180
ccggcgcuga uggacgugcc guacagacug aacgagauca ccaaggaaaa ggaacagugc	3240
ugcgugagcc uggccgccgg acugccgcac uucagcagcg gcaucuuccg gugcugggg	3300
cgcgacaccu ucaucgccc gcggggauc cugcugauca cgggccgcu cguggaggcc	3360
cggaacauca uccuggccuu cgccggcacc cugcggcacg gacugauccc gaaccuccug	3420
ggagagggga ucuacgccc guacaacugc cgggacgcc ugugguggug gcugcagugc	3480
auccaggacu acugcaagau ggugcccaac ggccuggaca uccugaagug ccccgucagc	3540
cggauguacc cgaccgacga cagcgcccc cuccccgcc gaaccucga ccagccacug	3600

ULPI_041_01WO_SeqList_ST25.txt

uucgaaguga uccaggaggc caugcaaaag cacaugcagg gaauccaguu cagagaacgg	3660
aacgccggac cccagaucga ccggaacaug aaggacgagg gauucaacau caccgccgga	3720
gucgacgagg aaaccggcuu cgucuacgga gggaaccggu ucaacugcgg aacauggaug	3780
gacaagaugg gagagagcga ccgcgccagg aaccgcggaa ucccagcaac cccgcgggac	3840
gggagcgagc uggaaaucgu ggggcugagc aaaagcgccg ugcgguggcu gcucgaacuc	3900
agcaagaaga acaucuuccc cuaccacgaa gucaccguga agagacacgg aaaggccauc	3960
aaagucagcu acgacgaug gaacaggaag auacaggaca acuucgagaa gcuguuccac	4020
gucagcgagg acccgagcga ccugaacgag aagcacccca accuggugca caagcgcggg	4080
aucuacaagg acagcuacgg cgcgagcagc ccguggugcg acuaccaacu gcgccccaac	4140
uucaccaucg ccauggucgu ggccccgaa cucuucacga ccgagaaagc guggaaggcg	4200
cuggagaucg cggaaaagaa gcuccucgga cccuggggga ugaaaacccu ggacccccgac	4260
gacauggugu acugcggcgu cuacgacaac gcgcuggaca acgacaacua caaccuggcc	4320
aagggcuuca acuaccacca gggccccgaa uggcuauaggc cgauccggcu cuuccugcgg	4380
gccaagcugu acuucagccg ccucaugggc ccagaaacca ccgcaaagac caucgugcug	4440
gucaagaacg uccugagccg ccacuacgug caccucgaga gaagcccaug gaagggacug	4500
cccgagcuga ccaacgaaaa cgcgcaguac ugcccguuca gcugcgaaac ccaggccugg	4560
agcaucgcca ccauccugga aacacuguac gaccucuag	4599

<210> 20

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 20

augggccaca gcaaacagau ccggauccuc cuacucaacg aaauggaaaa gcucgaaaag	60
---	----

acucuguucc ggcuggagca aggauacgag cuucaguucc gguugggccc gacgcugcag	120
---	-----

ULPI_041_01W0_SeqList_ST25.txt

gggaaggccg ugacagucua cacuaacuac ccauuccccg gcgaaaccuu caacagagag	180
aaguucaggu cccuggacug ggagaacccc acugaacggg aagaugauuc cgacaaguac	240
ugcaagcuga accuccagca auccgguuca uuccaguacu auuuccugca aggaaacgaa	300
aagagcggag gaggcuaau ugugguugau ccaauccuuc gcgugggagc cgauaaucau	360
gugcugccgc uggauugcgu caccugcaa accuuccugg cgaagugccu gggcccguuc	420
gacgaauggg aaucccgccu gcgggucgcc aaagaguccg gguauaaucau gauccacuuc	480
acccacucc aaacucuugg ccuguccaga uccugcuacu cguuggccaa ccaguuggag	540
cugaacccgg acuucucacg gccuaaccgg aaguacacuu ggaacgacgu cggacaacuc	600
guggagaagc ugaagaagga auggaauguc aucugcauua cugaugucgu guacaaccac	660
acggcggcga auucaaagug gauccaggag caccagaau gcgccuacaa ccuggucaau	720
ucccucacc ugaagccggc cugggugcuc gaccgggcc uugggcgguu uucgugcgau	780
guggcggagg gaaaguacaa ggagaagggg auuccggcuc ucaucgaaaa cgaccaccac	840
augaacucca uccggaaaau uaucugggag gacaucucc cgaaacuuaa gcugugggaa	900
uucuuccaag uggacgucaa caaggcugug gagcaguucc ggagauugcu gacacaagaa	960
aaccgccgcg ugacaaauc cgauccgaac cagcaccuga cuaucauua agaccggaa	1020
uaccggaggu uuggcugcac uguggacaug aaauacgcgc ugaccaccuu cauccgcac	1080
gacaagggac cggcagccau cgaagagugc uguaacuggu uccaauagag gauggaagaa	1140
uugaacucug aaaagcaccg gcugauuaac uaccaucagg aacaggccgu gaacugucuc	1200
cugggaaacg uguucuacga gcggcuggcg ggacacggac ccaagcucgg ccccgugacc	1260
cgcaagcauc cucucgugac uagauacuuu acuuuccau ucgaggagau cgacuucc	1320
auggaagaau caaugauca ccuccuaac aaggcuugcu uucugauggc acacaacggc	1380
ugggugaugg gcgaugaccc ccugaggaac uuugccgagc ccggcuccga aguguaccug	1440
aggagagagc ugauuugcug gggggacagc gucaagcugc gcuauggaaa caagccggaa	1500
gauugcccuu accucugggc ccacaugaag aaguacacug agauuacugc caccuacuuu	1560

ULPI_041_01WO_SeqList_ST25.txt

caaggagugc gccuggauaa cugccacuca accccgcugc augucgcuga guacauguug	1620
gacgcagccc ggaaucugca accgaaccuc uacguggugg cggagcuguu caccggcucg	1680
gaggauucug acaacguguu cgugacucgc cugggcaucu caucgcugau ccgggaagca	1740
auguccgccu acaacuccca ugaggagggu cggcuggugu accgcuacgg cggagaaccc	1800
gugggguccu ucgugcaacc gugccugcga cccugaugc ccgccaucgc ccaugcccuc	1860
uuuauaggaua uuaccacga uaacgaugc ccuauuguc accgcucagc cuaugaugcc	1920
cuccccucca ccaccaucgu guccauggcc ugcugcgccu cggggagcac ccgggguuau	1980
gacgagcugg ugccgcacca gauuucggug guguccgagg agagauucua caccaagugg	2040
aaccacagaag cucugccguc aaacacugga gagguuaacu uucaguccgg uauuauccg	2100
gcuagaugug cuauuagcaa acugcaccaa gagcuggggcg ccaaggguuu cauccaaguc	2160
uacgucgauc aggucgacga ggacaucgug gcugucacua ggcacucacc uagcauccac	2220
cagagcgugg uggccgugag ccgcacugcc uuccgcaacc caaagaccag cuuuuacucc	2280
aaggagggucc cucagaugug cauuccugga aagaucgaag aagugguccu ggaagcccgg	2340
accaucgaac gcaacaccaa gccuuaccgg aaggacgaaa acucgaucaa cgguacccccg	2400
gauauuacug uggagauucg cgaacacauc cagcucaaug aguccaagau cgugaagcag	2460
gccggagugg caaccaaggg gccgaacgag uacauccagg agaucgaauu cgagaauuc	2520
agcccuggca gcgugauc auuccgagug ucguuggacc cacaugcuca ggucgccgug	2580
ggcauccuga ggaaccaccu gaccaguuu uccccgcauu ucaaguccgg uucgcuggcc	2640
guggacaacg cagaucgguu ccugaagau cccuucgccu ccuuggcuuc gcgccucacg	2700
cuggccgagc ugaaccagau ccuguauaga ugcgaauccg aagaaaagga agauggaggc	2760
gguuuguuacg acaucccgaa cugguuccgcc uugaaauacg ccggacugca gggauugaug	2820
uccgugcugg ccgaaauuag accgaagaac gaccuggggc acccguucug caacaaccuc	2880
cggucggggg acuggaugau ugauuacgug ucgaaccgcc ucaucucccg guccgguacu	2940
auugcggagg ucggcaaaug gcuccaggcc auguucuucu acuugaagca aaucggccgc	3000

ULPI_041_01WO_SeqList_ST25.txt

uaccugaucc ccugcuacuu cgacgccauc cugaucggag ccuacaccac ccugcuggac	3060
accgcgugga agcagauguc uagcuucguc cagaacggcu ccaccuucgu gaagcaccug	3120
ucacugggcu cagugcagcu gugcggaguu ggaaaguucc ccagccugcc uauccugucc	3180
ccggccuuga uggacgugcc guacagacug aacgaaaucua cuaaggagaa ggagcagugu	3240
ugcgugucgu uggcggccgg gcugccccac uuuucaagcg ggauuuuccg gugcugggga	3300
agggacaccu ucaucgcgcu gcggggaauc cugcugauua ccggacgcuu cguggaggcu	3360
cggaacauca uucucgcguu cgccggcacc cugcgccaug gacugauccc uauucgcuu	3420
ggagagggca ucuacgcucg guacaacugc agagaugccg ugugguggug gcuccaauugc	3480
auccaggauu acugcaagau ggugccuaac ggucuggaca uucugaagug cccugugucc	3540
cgcauguacc ccaccgacga cuccgcccc cugccugccg gaaccucga ccagccucug	3600
uucgaaguca uccaagaggc caugcagaag cauauagcagg gcauucaguu ccgcgaacgg	3660
aacgcagggc cccagaucga ccggaacaug aaggauaag gguucaacau caccgccgga	3720
guggaugaag agacuggauu ugucuacgga ggaaaccgcu uuaacugcgg gaccuggaug	3780
gacaagaugg gagagucaga cagggccaga aaauagaggaa uccccgcgac cccgcgcgac	3840
ggauccgccg uggagauugu gggccuguca aagagcgccg uccgcuggcg gcuggaacuu	3900
agcaagaaga acaucuuccc uuaccacgaa gugaccguga aaagacaugg caaagccauc	3960
aaagucuccu acgacgagug gaaccgcaag auucaggaca acuucgaaaa gcuguuccac	4020
guguccgaag auccuccga ucugaacgag aagcauccga accuggucca uaagcgggga	4080
aucuacaagg auagcuacgg ugccagcucg ccguggugug acuaccagcu gcggccaaau	4140
uucacgauug cgauggucgu cgccccugag uuguucacca cugagaaggc cuggaaagcc	4200
cuggaaaauag ccgagaagaa auuacugggg ccacugggca ugaaaacucu cgaccucgac	4260
gacauggugu acuguggaau cuacgacaac gcacuggaca acgaaacua caauuuggcc	4320
aagggaauca acuaccacca ggggcccga uggcuguggc cuaucggaua cuuccuucgg	4380
gcaaagcuuu acuucucccg ccugauggga ccugaaacua cugccaagac cauugugcuu	4440

ULPI_041_01W0_SeqList_ST25.txt

gugaagaacg ugcucucacg gcacuacgug caccuggaga gaucggcgug gaaggggcuc	4500
ccggagcuga ccaacgagaa cgcccaguac ugcccuuucu ccugcgaaac ccaagccugg	4560
uccaucgcca ccuacuuga aacucuguac gaccuguag	4599

<210> 21
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 21	
augggacacu ccaagcagau ccgcauucuc cuccugaacg aaauggaaaa gcuggaaaag	60
accuuauuuc gccuugaaca aggauacgag cugcaguucc ggcuuggacc gaccuccaa	120
gggaaagcug ucaccgucua caccaacuac ccuucccg ggagagacuuu caacagagag	180
aaguucagau ccuuggacug ggaaaaucc acggaacggg aggacgacuc cgauaaguac	240
uguaaaacuga accuccagca gucgggaucg uuccaguacu acuuccuaca agggaaacgaa	300
aaguccggug gugguuacau cgugguggac ccgauccugc gcgugggggc cgacaaccac	360
gugcugccgc uggacugcgu gacucugcaa accuuccugg ccaagugccu cggcccguuc	420
gaugaauagg agucgagacu gcggguggcc aaggaaagcg gauacaauau gauucacuuu	480
acuccgcugc aaaccuggg gcugucgcgc uccugcuacu cgcucgcuaa ccagcuggaa	540
cugaauccgg auuuuagccg gccgaaccgg aaguauacgu ggaacgacgu cggucagcug	600
gucgagaagc ucaagaagga guggaacguc auuugcauca ccgacguggu cuacaaccac	660
accgcggcca auuccaagug gauccaggaa caccgggagu gcgccuauaa cuuggugaac	720
uccccuauu ugaaaccgc augggugcuc gaccggggcc uguggcgcuu uucaugcgac	780
guggcggaag ggaaguacaa ggaaaagggg auucccgccc ugauagagaa cgauaccac	840
augaacucaa uucgcaaaau caucugggag gauaucuucc cuaagcucaa gcugugggaa	900
uucuuccaag ucgacgucaa caaggccgug gaacaauuca gaaggcugcu gaccgaag	960

ULPI_041_01WO_SeqList_ST25.txt

aaccgcagag ugaccaaguc ggacccgaac cagcaccuga cuauaaucca ggacccggaa	1020
uaccgggaggu ucgguugcac cguggacaug aacauugcac ucaccaccuu caucccgcac	1080
gauaaggguc cggcggcgau ugaggagugc ugcaacuggu uccacaagcg gauggaggaa	1140
cugaacuccg agaagcaccg gcugauuaac uaccaccagg aacaggcugu gaauugccug	1200
cuggggaacg uguucuacga acggcucgcc ggccacggcc ccaagcuggg gcccgucacc	1260
cgcaagcacc cgcucgugac ucgauuuuc accuucccg ucgaggaaau ugauuucucc	1320
auggaggaa ucaugaucca ucugccgaac aaggccuguu ucuugauggc ccacaacggc	1380
ugggugaugg gagaugaccc gcugagaaau uucgccgagc cgggguccga gguguaccug	1440
aggagagaac ucaucugcug gggggauucc gugaaacugc gcuacgggaa caagcccag	1500
gacugccccc accugugggc acauugaag aaguacaccg aaauaccgc caccuacuuc	1560
caaggggucc ggcuaggaua cugccauuca acuccgcucc auguggccga auacaugcug	1620
gacgccgcac ggaacuugca gcccaccuc uacguggugg ccgaacucu cacugggagc	1680
gaagaucucg auaacuguu cgugacccgg cucggaauu cgagccugau ccgggaagcg	1740
augucugcuu acaacucga cgaagaggga aggcuggugu acagauacgg gggggagccc	1800
gugggauccu ucgugcaacc gugccugagg ccgcuaugc ccgcgaucgc ccacgcucug	1860
uucauggaca ucaccacga caacgaugu ccgauugucc accggagugc cuacgaugcc	1920
cuuccgagca cuaccaucgu gucgauggcg ugcugugcgu ccgggucuac ccgcggcuac	1980
gacgaacucg ucccgcacca aaucagcgug guguccgaag aacgguuua cacuaagugg	2040
aacccggagg ccuugcccuc gaacaccggg gaggucaacu uccagucggg aaucuuugcc	2100
gcccgaugug ccaucuaaa guugcaccaa gaacucgggg caaagggguu cauccaagug	2160
uacguggacc aaguggacga ggauuuugug gccgugacca ggcacucucc gucgauccac	2220
caguccgugg uggcagucuc aagaacugcc uuccggaacc ccaagaccuc cuuuuacucg	2280
aaggaagugc cccagaugug cauccgggc aaaucgaag aagucgugcu ggaagccaga	2340
acuaucgagc ggaacaccaa gcccuaccgg aaggauagaga acagcauaa cggcacgccc	2400

ULPI_041_01WO_SeqList_ST25.txt

gacaucaccg ucgagaucag agagcacauc cagcugaacg aauccaagau ugucaagcag	2460
gccggcguag cgacuaaggg acccaacgaa uacauccagg agaucgaauu cgaaaaccug	2520
ucgccgggau ccgugauuau cuuccgggug ucccuggacc cgcacgcca aguggccgug	2580
ggaauccugc ggaaccaccu gaccaguuc uccccgcauu ucaaguccgg cucccuggcg	2640
guggacaaug cagacccgau ucucaagauc ccguucgccu ccuuggccuc ccgccugacc	2700
cucgccgaac ucaaccagau ccuguaccgc ugcgagucug aggaaaagga ggacggggga	2760
ggaugcuacg auauuccgaa cuggucugcc cuuaauacg cgggacugca gggucugaug	2820
ucggugcucg ccgagaucag accgaagaac gaccuggggc acccguuuug caacaaccug	2880
agaucggggg auuggaugau cgacuacgug ucgaaccggc uuaucucucg cuccggcacc	2940
aucgccgagg ucgggaagug gcugcaggcc auguucuuc accugaagca aaauccgcgc	3000
uaccugaucc cgugcuacuu ugaugccauc cugaucggcg ccuacacgac cuugcuggac	3060
accgccugga agcagaugag cagcuucgug cagaacggcu ccacguucgu gaagcaucug	3120
ucccugggau ccgugcagcu uugcgguguc ggaaaguucc cuucgcugcc uauucugucc	3180
ccggcgcuga uggacgugcc guaccggcug aacgagauca cuaaggagaa ggaacaaugc	3240
ugcgugucac uggcggcggg acucacccac uucuccuccg ggauuuuccg guguugggga	3300
agagauaccu ucauugcccu gagggggaua cugcugauca ccggaagaua cguggaggcu	3360
agaaacauca uccucgccuu cgccggcacu cuccgccacg ggcugauccc caaccucuug	3420
ggcgaaggaa ucuacgcccg guauaacugc cgcgacgcgg ucugguggug gcuucagugc	3480
auccaggauu acugcaaaau ggugccgaau ggacuggaua uucugaagug ccccguaagc	3540
cgcauguacc ccaccgauga cuccgcaccc uugccggcgg gcaccuugga ccagccgcuc	3600
uuugaaguga uccaggaggc caugcagaag ccaugcagg guauccagu cagagagcgg	3660
aacgcugggc cgagauuga ccggaacaug aaggacgaag guuucaacau uacugccggg	3720
guggacgaag agacuggguu uguguacggg gggaaccgcu ucaacugugg gaccuggaug	3780
gauaagaugg gggagucgga ucgggcucgg aacaggggga uccccgcuac ccccgggac	3840

ULPI_041_01W0_SeqList_ST25.txt

gggucggcug	uggagauugu	cgggcugagc	aagagcgag	ugcgucggcu	gcuggagcuc	3900
agcaaaaaga	acaucuuccc	cuaccacgaa	gucacuguga	agagacacgg	aaaggccauc	3960
aagguguccu	acgaugagug	gaacagaaag	auccaagaca	acuucgagaa	gcuguuccau	4020
guguccgagg	auccgagcga	ucugaacgaa	aagcauccga	accucgugca	caagcgcggg	4080
aucuacaagg	acucguacgg	cgcauccucg	ccguggugcg	acuaccagcu	gcgccccaa	4140
uucaccauug	ccaugguggu	ggcgccccgag	cuguucacua	ccgagaaggc	cuggaaggcc	4200
cuggagauug	cugagaagaa	gcugcuggga	ccgcucggga	ugaaaacuuu	ggaccucgac	4260
gacauggugu	acuguggaau	cuacgacaau	gcgcucgaca	acgacaacua	caaucuugcg	4320
aagggaauca	acuaccauca	ggggccccgaa	uggcuguggc	cgauccggua	cuuccugcgc	4380
gccaagcugu	acuucucccg	gcugaugggg	cccagagaaa	ccgccaagac	cauugugcuc	4440
gucaagaacg	uucucucccg	gcacuacgug	caccucgaaa	gaucuccgug	gaaggggcug	4500
cccgagcuca	ccaacgagaa	cgcacaguac	ugcccguucu	ccugugaaac	ccaggccugg	4560
agcaucgcca	ccauauugga	aacucuguau	gaccucuag			4599

<210> 22
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 22	
augggacacu	cgaaacagau
cagaauccug	uugcugaacg
aaauggagaa	gcuggaaaag
60	
acccucuuc	ggcucgaaca
gggcuacgag	cugcaguucc
gccugggucc	gacucuacaa
120	
ggaaaggcag	ugacugugua
caccaacuac	ccauuucccg
gagaaaccuu	caaccgggag
180	
aaguuccggu	cccuggacug
ggaaaacca	accgaacgag
aggaugauuc	cgacaaguac
240	
ugcaagcuga	accuccaaca
guccggauca	uuccaguacu
acuuuuugca	aggaaacgag
300	
aaguccggag	gaggcuacau
cguggucgac	ccgauucuga
gagugggagc	ugacaaucan
360	

ULPI_041_01W0_SeqList_ST25.txt

guccugccuc	uggacugcgu	aaccugcaa	accuucuugg	ccaaugccu	gggcccguuc	420
gaugaguggg	aaagccgccu	ccgggucgca	aaggaaagcg	gcuacaauau	gauucacuuc	480
acuccucugc	aaaccucgg	ucugucccg	uccuguuauu	cccuggcgaa	ccagcuggag	540
cuuaaccccc	acuucucg	cccaaaccgg	aaguacaccu	ggaacgacgu	cgggcagcuu	600
guggaaaagc	ugaagaagga	guggaacgug	aucugcauca	cugacguggu	guacaaccac	660
accgcggcca	acuccaagug	gauccaagaa	caucccgauu	gcgcguacaa	ccucgugaac	720
agcccgcuauc	ugaagccugc	cugggugcug	gauagggccc	ucuggagauu	cagcugcgac	780
guggccgagg	ggaaguacaa	agaaaaggga	auucccgccc	ugauugagaa	cgaccaucac	840
augaacucuaa	uccgcaagau	caucuggggag	gauaucuucc	cuaagcuuaa	guuguggggag	900
uucuuccaag	uggacuguaa	caaggcagug	gagcaguuca	gaaggcugcu	gacucaagaa	960
aacagacggg	ucacuaaguc	cgaccuaac	cagcaccuua	ccaucauua	agaccggag	1020
uaccgccggu	uuggcugcac	ugucgacaug	aacaucgccc	ugaccacuuu	caucccgcau	1080
gacaagggcc	cggcggaau	cgaggaaugc	uguaacuggu	uucuaaagag	gauggaggaa	1140
cugaaauuccg	agaagcaccg	gcugauuaac	uaccaccagg	aacaggcagu	gaauugccuc	1200
cugggaaacg	uguucuacga	acggcuggc	ggacacggac	cgaagcuggg	ucccgugacu	1260
cgcaagcauc	cgcucgugac	ucgcuacuuc	accuucccg	uugaggagau	ugacuuuucc	1320
auggaagaau	ccaugaacca	ccucccgaa	aaggcuugcu	uccugauggc	gcacaacgga	1380
ugggucaugg	gcgacgaccc	acugcgcaac	uucgcugagc	cuggcagcga	ggucuaccug	1440
agaagggaac	ugauuugcug	gggagacucc	gucaagcugc	gcuauaggaa	caagcccga	1500
gauugcccu	accugugggc	ucacaugaag	aaguacacug	aaucacugc	cacguacuuc	1560
caggagucc	ggcuggacaa	uugccacucc	acccccucc	auguggccga	guacaugcuc	1620
gaugcagcga	ggaaucugca	gcccuaucug	uacgugguug	ccgaacuguu	caccggcucc	1680
gaggaccucg	acaacguguu	cgugaccaga	cuugggauuu	ccagccugau	ccgggaagcc	1740
augucggccu	acaacuccca	ugaagagggc	cgccuggugu	accgcuacgg	cggagaaccc	1800

ULPI_041_01WO_SeqList_ST25.txt

gugggaagcu ucgugcagcc uugccuccgg ccgcugaucg cugcgauccg ccacgcccug	1860
uucauggaua uuaccacga caacgaugc cccaucgugc aucgcucggc cuacgacga	1920
cugccuucga ccaccaucgu guccauggcc ugcugcgccu ccgguagcac ccgcggauac	1980
gaugaacucg ugccgcacca gaucagcgug guguccgaag aaagauucua caccaagugg	2040
aaccucgagg cacugccgag caacaccgga gaagugaacu uccagucggg uauuauccg	2100
gcucgcugug ccaucuccaa acuccaccaa gagcucggug ccaagggaau cauuaaguc	2160
uacguggauc aggucgacga agauauugug gccgugacca ggcacucacc uuccauccac	2220
caauccgucg ucgccguguc acggacugcc uuccggaacc ccaagacuuc guucuacucg	2280
aaagaggugc cacagaugug uauccccgga aagaucgaag aggucguccu ggaagcccgg	2340
accauugaga ggaacaccaa gccuuaccgg aaggacgaga acagcauca cgguaacccu	2400
gauauuacug uggagaucg cgaacacauc cagcucaacg aaucaaagau ugucaagcag	2460
gccggagugg ccaccaaggg acccaacgag uacauccagg agaucgaau ugaaccuc	2520
uccccuggcu ccgugauuau cuuccgggug ucccuggacc cucacgcca aguggccguc	2580
ggcauccuca gaaaccaccu gaccagauuc ucaccccau ucaaguccgg uucccuggcg	2640
guggacaacg ccgauccgau cuugaagau cccuuugcau ccugggccuc ccgccugacu	2700
cucgcggaac ugaaccagau ccuguaccgc ugcgaucag aggaaaagga ggacggcggc	2760
ggcuguuacg auauccgaa uugguccgcu cugaaauaug cgggacuua ggggcugaug	2820
agcgugcugg cggagaucg gccgaagaau gaccugggac acccauucug caacaacuug	2880
cggagcggag acuggaugau cgauuacguc agcaacagau ugaucucccg gagcggcacu	2940
aucgcggagg ucggaaagug gcuccaggcc auguucuuc accugaagca gaucuccga	3000
uaccucaucc cguguuacuu cgacgccauu cugaucgggg ccuacaccac ccugcuggac	3060
accgccugga agcagaugag cuccuuugug caaacgggu ccaccuucgu gaagcaccuu	3120
ucacugggau cagugcagcu cugcggcgug ggaaaauucc ccucgcugcc cauucugagc	3180
ccggcccua uggacgucc uuaccggcug aacgagaua ccaaggagaa ggagcagugc	3240

ULPI_041_01WO_SeqList_ST25.txt

uguguuuccc	uggcugccgg	acuuccacac	uucucguccg	gcaucuuccg	gugcuggggc	3300
cgggauaccu	ucauugcccu	gcggggcauc	uuguugauca	ccggucgcu	cguggaggcu	3360
cggaacauua	uucuggcauu	cgccggcacu	cugagacacg	gucugauucc	gaauuuuug	3420
ggcgaaggaa	ucuacgcccc	cuacaacugu	cgggacgccg	ugugguggug	gcuccagugc	3480
auucaggacu	auugcaagau	ggugccgaac	ggccuggaca	uccugaagug	cccagugucg	3540
aggauguauc	caaccgacga	cagcgcgccu	cugccggccg	ggaccucga	ccaacccug	3600
uucgaaguca	uucaagaggc	uaugcagaag	cacaugcagg	guauucaguu	ccgggagaga	3660
aacgcggggc	cccagauuga	uaggaacaug	aaagacgagg	gcuucaacau	cacugccggc	3720
guggacgaag	aaaccgguuu	uguguacgga	ggaaaccggu	ucaacugcgg	uaccuggaug	3780
gacaagaugg	gagaguuccga	ucgcgcgcgc	aacagaggga	ucccggaac	cccgcgggac	3840
ggaucagcgg	uggaaaauugu	gggacugagc	aagagcgccg	ugcgcuggcu	ccuggaacug	3900
agcaaaaaga	acaucuuccc	cuaccacgaa	gugaccguga	agcggcacgg	aaaggccauc	3960
aaagucucau	acgaugaaug	gaauaggaag	auccaggaua	acuucgagaa	gcuguuucac	4020
guguccgagg	aucccuuccga	ucugaacgaa	aagcauccga	aucucgugca	caagcgcggc	4080
aucuacaagg	acucguacgg	agccuccucc	ccuuggugcg	auuauagcu	gcggccuaac	4140
uucaccauug	ccauggucgu	ggcuccggag	cuguuacua	cugagaaggc	cuggaaggca	4200
cuugaaaauug	ccgagaagaa	gcugcuuggg	ccuuugggga	ugaaaacccu	ggauccggac	4260
gacauggugu	acugcggaau	cuacgacaac	gcccuggaca	acgacaacua	caaucuggcg	4320
aagggcuuca	auuaccacca	gggcccggaa	uggcuguggc	cuauugggua	cuuccugcgc	4380
gccaagcugu	acuucucacg	gcugauggga	ccagagacua	ccgccaagac	uaucguccuc	4440
gugaagaacg	ugcugucccg	gcacuacgug	caucuggaga	ggagcccuug	gaagggccuu	4500
ccugagcuga	ccaacgaaaa	cgcccaguac	ugcccuucu	ccugcgaaac	ccaggcuugg	4560
uccauugcca	cuauacugga	aaccuuauau	gaccuguag			4599

ULPI_041_01W0_SeqList_ST25.txt

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 23

augggacacu caaagcagau ccggauccuc cugcugaacg aaauggaaaa gcucgaaaag	60
acucuguucc ggcuggaaca gggcuauagag cugcaguucc ggcucggacc gacgcugcaa	120
ggaaaggccg ugacugugua caccaacuac cccuucccgaggagaaacuua uacagagaa	180
aaguuuaggu cccuggacug ggagaacccg accgagcggg aagaugacuc cgauaaguac	240
ugcaagcuga accuccagca gucgggaucg uuccaguacu acuuccuua aggaaaugag	300
aagucuggug gaggauacau cgugguggac cccauccuga gagugggagc cgacaauac	360
guccugccgc uugacugcgu gacucugcaa accuuccugg cuaagugucu cgggccguuc	420
gacgaauggg aguccaggcu gcgcguggcu aaggagagcg gcuacaauau gauccacuuc	480
accccgcuucc aaacucucgg acucucgcgc uccuguuacu cccucgcaa ccaacuggag	540
cugaacccgg acuucucccg cccgaaccgg aaguacaccu ggaacgacgu cggucagcuc	600
guggagaaaau ugaagaagga guggaacgug aucugcaua cggacgucgu guacaaccu	660
acugcggcca acuccaagug gauucaagag caccggaau gcgcuuacaa ccuugucaac	720
ucggcgacc ugaagccggc cugggugcug gacagagccc uguggcgguu cucgugcgac	780
guggcagagg gaaaguacaa ggagaagggc aucccgccc uuauugagaa ugaccaucac	840
augaacucca uccggaagau caucugggag gauaucuucc cgaagcucaa gcugugggaa	900
uucuucaag uggacgugaa caaggccgug gaacaguucc ggagacuccu gaccaagaa	960
aaccggagag ugaccaagag cgacccgaac cagcaucuga cuaucauua ggaucggag	1020
uauccggcgu uuggcugcac uguggacaug acaucgccc ucaccaccuu cauccgcac	1080
gacaagggcc ccgccgcgau cgaggaaugc ugcaacuggu uccacaagcg cauggaggag	1140
cugaaauccg aaaagcaccg ccugaucaac uaccaucagg agcaggcagu gaacugucuc	1200

ULPI_041_01W0_SeqList_ST25.txt

cugggaaacg	uguuuuacga	acggcuggcc	ggacacggcc	cgaagcuggg	ucccgugacc	1260
cgcaagcauc	cccucgugac	gcgguacuuc	accuucccg	ucgaggagau	cgacuucagc	1320
auggaagagu	ccaugaucca	ucugccgaac	aaggccugcu	uccucauggc	gcuaaauugu	1380
ugggucaugg	gagaugaucc	ccuccgaaac	uuugcggagc	cggguucgga	aguguauucg	1440
aggagggagc	ucaucugcug	gggagauagc	gugaaacuga	gauacgggaa	caagccggaa	1500
gauuguccgu	accugugggc	acacaugaag	aaguacaccg	aaauacugc	cacuuacuuc	1560
caaggaguuc	gccuggauaa	cugccauuca	accccucugc	augucgccga	guacaugcug	1620
gacgccgcuc	gcaaccuuca	gccgaaucuc	uacguggucg	ccgagcuguu	caccgguucc	1680
gaagaucugg	acaacguguu	cgugacuaga	cugggaaucg	gcagccugau	ccgggaagcg	1740
augagcgccu	acaacuccca	cgaagagggc	cggcucgugu	auagauacgg	cggagagccg	1800
gucgggagcu	ucgugcaacc	cugccugcgg	ccgcugaugc	ccgccauugc	ccacgccuug	1860
uucauggaua	ucacccacga	caacgagugu	ccgaucgugc	accggagcgc	guacgacgcg	1920
uuaccgucca	ccacgauugu	gucgauggcc	ugcugcgccu	ccggaucgac	ccgcggcuac	1980
gaugagcugg	ucccgcauca	aaucagcguc	gucagcgaag	aacggguucua	cacuaagugg	2040
aaccccgagg	cgcuccccuc	caacaccgga	gaagugaacu	uccaauccgg	cauuauucgu	2100
gcacgcugcg	cgauuagcaa	gcugcaucag	gagcuuggcg	cuaagggguu	cauacagguc	2160
uacguggauc	aggucgacga	ggacauugug	gccgugacuc	gccacucacc	guccauucac	2220
caaagcgugg	uggcuguguc	ccggaccgcu	uuccggaauc	ccaagaccuc	auucuacucc	2280
aaggagguuc	cgcagaugug	uaucgggga	aagauugagg	aagugguccu	agaggcucgc	2340
acuauugaac	gcaacaccaa	gccguacaga	aaggacgaga	auuccauuaa	cgggacccccg	2400
gauauuaccg	uagaaauuag	agaacauuu	cagcugaacg	agucgaagau	cgugaagcag	2460
gccggggugg	cuaccaaggg	cccgaacgaa	uacauccagg	agauugaguu	cgaaaaccuc	2520
ucccccggcu	cggugaucau	cuuccgggug	ucccucgauc	cccacgcccc	aguggccgug	2580
ggaaucuuga	gaaaccaccu	gacucaguuc	agcccgcuuu	ucaaguccgg	aucacucgcc	2640

ULPI_041_01WO_SeqList_ST25.txt

guggacaacg ccgacccgau ccuuaagauc cccuucgcau cacuggccag ccgccugacc	2700
cuugccgaac ugaaccaaau ccucuaucgg ugcgaauccg aggagaagga agaugggggu	2760
ggaugcuaug acauuccuaa cuggucugcc cugaaauacg caggccugca gggacugaug	2820
uccguccugg ccgaaauccg cccgaagaau gaccuggggc accccuuuug caacaauug	2880
agguccggag auuggaugau ugauuacgug uccaaccgcc ugauuucgcg gagcggcacc	2940
aucgcggaag ugggaaagug gcugcaagcc auguuuuucu accugaaaca gaucggcg	3000
uaccucaucc cgugcuauuu cgaugcgauu uugauuggcg cguacaccac ccuccucgac	3060
acugccugga agcaaaugag cuccuucgug caaaacgguu caaccuucgu caagcacuug	3120
ucgcugggggu cgguccaacu guguggaguc ggaaaguucc cgagccugcc cauccugucg	3180
ccugcccuga uggacgugcc cuaccggcug aacgagauca ccaaagaaaa ggagcagugc	3240
ugcguguccc uggcggccgg acucggccau uucucguccg ggaucuuuag auguugggga	3300
cgggauacuu ucauugcgcu gcgcggcauc uuguugauua ccggacgcua cguggaagcg	3360
cggaacauca uacucgccuu cgccggcacc cucagacacg gccugauccc gaaccuccug	3420
ggagaaggca ucuacgcacg auacaacugu cgggacgccg ucugguggug gcugcagugc	3480
auucaggacu acugcaagau ggugccgaac ggucuggaca uccugaagug ccccguguca	3540
aggauguacc cgaccgacga uagcgcaccg cugcccgcgg ggaccucgga ucagccgcug	3600
uucgaaguga uccaagaagc caugcaaaag cacaugcagg gaauucagu cagagaaagg	3660
aacgcaggcc cccagauuga ccggaacaug aaggacgagg guuucaacau caccgccggc	3720
gucgacgaag aaacggggguu uguguacggg ggcaaccgcu ucaacugcgg uacuuggaug	3780
gacaagaugg gggaaucaga ccgcgcccgc aaccgcggaa uuccggcgac cccgcgcgau	3840
ggauccgcag uggaaaucgu gggacugucc aagucggcug ugcgguggcu gcuggagcug	3900
uccaagaaga acaucuuccc guaccacgag gucaccguga agagacacgg gaaggccauc	3960
aagguguccu acgacgaug gaaccggaaa auccaggaua acuucgaaaa acuguuucac	4020
guguccgagg acccgucuga ccugaacgaa aaacacccga acuuggugca uaagcgggga	4080

ULPI_041_01WO_SeqList_ST25.txt

aucuacaagg acaguuacgg agccucaagc cgguggugcg acuaccagcu gcggcccaac	4140
uucacaaucg cgaugguggu ggcgcccgag cuuuucacca cggagaaagc cuggaaggca	4200
cuggagaucg cugaaaagaa gcugcucggu ccccucggca ugaaaacccu ggacccggac	4260
gauauggugu acugcggaau cuacgacaac gcccuggaca augauaacua caaccuggcc	4320
aagggaauca auuaccacca ggggccggaa uggcuguggc cgaucggcua cuuucugcgg	4380
gcuaagcucu acuucucgcg gcugauggga ccggaaacca cagccaagac cauuguccuc	4440
gucaagaacg ugcugucccg ccacuacgug caucuggagc ggucaccuug gaagggccug	4500
cccgagcuga ccaacgaaaa cgcccaguac ugcccccucu ccugugaaac ucaagccugg	4560
ucaaucgcca cuauucucga aacucuguac gaucuguag	4599

<210> 24
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 24	
augggucaca gcaaacaaau ccggauccug cugcucaacg aaauggaaaa gcuggaaaag	60
acucuguuuc gccuggagca gggcuacgaa cuccaauucc ggcuggggacc aacgcugcaa	120
gggaaggccg ucaccgugua caccaacuac cccuuucccg gagagacuuu uaaccgcgaa	180
aaguucagaa gccuggacug ggaaaacccc accgaacggg aagaugauuc ggacaaguac	240
ugcaagcuga accuccagca guccggcagc uuccaguacu acuuccucca agggaaacgaa	300
aagagcggag gaggcuaau cgugguggau ccgauccuuc gggugggagc ggacaaccuu	360
gugcugccuc uggacugcgu gacucugcaa accuuccuag ccaagugccu cggcccguuc	420
gaugaguggg aaucgggcu gcgaguggcg aaggagucag guuacaacau gauccacuuc	480
acaccgcugc agaccugggg acucuccggg uccugcuacu cauuggccaa ccagcucgaa	540
cugaacccgg acuucucaag gccgaaccgc aaguacacuu ggaugacgu gggacaguug	600

ULPI_041_01WO_SeqList_ST25.txt

guggagaagc ugaagaaaga guggaacgug aucugcauua cgcacguggu guacaaccac	660
acugcggcga auucgaagug gauucaggaa cauccggaau gcgccuacaa cuuggugaac	720
ucaccucacc ugaagccugc cugggugcug gacagagccc uguggcgcuu uagcugugac	780
gucgcggagg ggaaguacaa ggaaaaggga aucccgccc ugaucgagaa cgaccaccau	840
augaauagca uucgcaagau caucugggag gacauuuucc cgaagcugaa gcugugggag	900
uucuuccaag ucgacgucaa caaggccguc gaacaguucc ggcgguugcu gaccaagaa	960
aaccggagag ugacgaaaag cgaccuaac cagcaccuga caaucauua ggaccugaa	1020
uaccggagau ucggcuguac cgucgauaug aacaucgccc ugaccaccuu uauuccccac	1080
gauaagggcc cggcggccau cgaggagugu ugcaacuggu uccacaagag aauggaggaa	1140
cucaacuccg aaaaacaccg guugauuaau uaccaccagg agcaggcggu caacugccug	1200
cugggcaacg uguucucga aaggcucgu gggcacggcc cuaagcuggg accugucacu	1260
cggaaacacc cccuggugac ccgguacuuu acuuucccg ucaagagau cgacuucc	1320
auggaggaaa gcaugaucca ccuccuaac aaggccugcu uccugauggc ccacaacgga	1380
ugggucaugg gcgacgacc uuugcggaac uuugccgagc cgggcucaga gguguaccuu	1440
agacgcgaac ugaucugcug gggagacucc gugaagcugc gcuacgggaa caagcccga	1500
gauuguccuu accugugggc acacaugaag aaguacaccg aaauacugc caccuacuuc	1560
caaggagugc gcuuggauaa cugucauagc acuccucugc acgucgccga guacaugcug	1620
gacgcagcca ggaauccca gccuaaucug uacguggugg ccgaacuguu caccgggagc	1680
gaggaccugg acaacguguu cgugaccgg cugggcauca gcuccuuau ccgggaggcc	1740
auguccgcau acaacucgca cgaggagggu cgccuggugu accgcuacgg aggagagccg	1800
gucggaucau uuguccaacc augccuccgg cccugaugc cagcgauccg gcacgcucuc	1860
uucauggaca ucaccacga uaacgaugc ccuacuguc accggagcgc uuacgacgcc	1920
cucccuucga ccaccauugu guccauggcc ugcugcgccu ccgguuccac ucggggcuac	1980
gaugaacugg ugccacacca gauuuccgug gugucagagg agcgguucua caccaagugg	2040

ULPI_041_01WO_SeqList_ST25.txt

aacccagagg cgcugccgag caacacuggg gaagugaacu uccaguccgg aaucauagcg	2100
gcuagaugug caauuuccaa gcuucaucag gagcuggggcg ccaaaggauu cauccagguc	2160
uacguggacc aaguggauga ggacauugug gccgugacua gacauucacc gucaauccac	2220
caaucggucg uggcuguguc cagaaccgca uuccgcaacc ccaagacuuc cuucuacucc	2280
aaggaggucc cucagaugug caucccggga aagaucgagg aggguggugcu ggaggccagg	2340
accaucgaaa gaaacacuaa gccuuaccgg aaagacgaaa acuccaucaa ugggacucucc	2400
gauauuaccg uggagauuag ggagcacauu cagcugaacg aaucaaagau ugugaaacag	2460
gccggggucg caacuaaagg gccuaaugag uacauccaag agaucgaguu cgaaaaccug	2520
ucgccggguu ccgugaucau uuuccgugug uccuuggacc cucacgccc aaguggccgug	2580
ggcauccuga ggaaccaccu gaccaguuc agcccacacu ucaaguccgg auccuuggcu	2640
guggacaacg ccgauccaau ccucaagauc ccuuuugcgu cgcuggccuc acggcuuacc	2700
cuggccgagc ugaaccagau ccuguaccgc ugcgagagcg aagaaaagga ggacggugga	2760
gguugcuacg acauuccaaa cugguccgug cuuaaguacg ccggacucca gggucugaug	2820
uccgugcuug ccgagaucau accgaaaaac gaccuggggg accccuuuug caacaaccug	2880
agaagcggag acuggaugau cgacuaugug uccaaccggc ugauuucgag aagcgguacu	2940
auugccgagg ucggaaagug gcuccaagca auguuuuucu accugaagca aaauccccgc	3000
uaccugaucc cgugcuacuu ugacgccauu cugaucggug cauacacuac ccugcuggac	3060
accgcgugga agcagauguc cagcuucgug cagaauggcu ccaccuucgu caagcaucug	3120
ucccucggaa gcgugcagcu gugcggagug ggaaaguuu cguccgugcc aauccugucc	3180
cccgcgcuga uggaugucc guaccggcug aacgaaauc cgaaggaaaa agaacagugc	3240
ugcguguccc uugccgccgg acugccgcac uucuccuccg gcauuuuccg gugcugggga	3300
cgggauaccu ucaucgcgcu gagagguaau cugcugauua ccggcagaua uguggaagcc	3360
cgcaacauc auccuggccuu ugccggcacu cugcggcacg ggcucauccc uaaccucuug	3420
ggagaaggca ucuacgcgcu cuacaacugc cgggaugcug ugugguggug gcugcagugc	3480

ULPI_041_01WO_SeqList_ST25.txt

auccaggacu acuguaaaau ggugcccau ggccuugaca uccucaagug cccagugucc	3540
cggauguacc cgaccgauga cuccgcgccc cugcccgccg gcacccuuga ucaaccucug	3600
uucgaaguca uccaagaggc caugcagaag cacaugcagg gaauccaguu cagggaaaga	3660
aacgccgggc cucagaucga ccggaacaug aaggacgagg gauucaacau uaccgccgga	3720
guggacgagg aaacuggcuu cguguacggu ggaaaccgcu ucaacugcgg caccuggaug	3780
gauaagaugg gcgaauccga ucgcgccccg aaccggggaa ucccagcaac uccuagggac	3840
ggaagcgag ucgagaucgu ggggcugucc aaguccgccg ugcgguggc cuuagaacug	3900
uccaagaaga auauuuuccc cuaccacgag gucaccguga agcgccaugg aaaggccauc	3960
aaaguguccu augacgaug gaaccgcaag auccaggaca acuucgaaaa guuguuccac	4020
guguccgagg accccagcga ucugaacgaa aagcaccca accucgugca caagaggggc	4080
aucuacaagg acuccuacgg agcuagcucc ccuuggugcg auuaccaacu gcggccuau	4140
uucaccaucg ccauggucgu cgcacccgaa cucuucacca ccgagaaggc cuggaaggcu	4200
cuggaaaucg ccgaaaagaa gcuucugggc ccgcugggca ugaaaacucu cgauccugac	4260
gauauggucu acuguggcau cuacgacaac gcccuugaca acgacaacua caaccuggcc	4320
aagggcuuua acuaccacca gggcccagag uggcucuggc ccaucggaua uuuccugcgg	4380
gccaaacugu acuucucgcg guugaugggu ccggaaacca cagcuaagac caucgugcuc	4440
gugaagaug ugcugucgcg gcacuauguc caucuggagc gcucucccug gaagggacuc	4500
ccugagcuga cuaaugagaa cgcccaguac ugcccuuucu ccugcgaaac ccaggccugg	4560
agcaucgcua ccuuucuga aacgcucuac gaucuguag	4599

<210> 25

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 25

ULPI_041_01WO_SeqList_ST25.txt

augggacauu cuaaacagau ucggauccuc cugcucaacg aaauggagaa gcucgaaaag	60
accuguuucc ggcucgaaca gggauacgaa cugcaauucc ggcugggucc caccuuacaa	120
ggaaaggccg ugacugucua caccaacuac ccguuccccg gcgaaaccuu caauaggagag	180
aaguuccggu cccucgauug ggagaacccu acagagcgcg aagaugauuc agacaaguau	240
ugcaagcuga auuugcagca gagcggcuca uucaauacu acuuccugca aggaaacgaa	300
aagucgggcg gcggcuacau cgugguggac ccuauucuga gagugggcgc cgacaaccu	360
guccugccgc uggacugcgu caccugcaa acuuuccugg ccaagugccu ggguccuuuc	420
gaugaguggg aaucgagacu ucgcguggca aaggaucgg guuauauau gauucacuuc	480
accccgugc aaaccuggg ccugucgcg ucuugcuacu cacuggccaa ccagcuggaa	540
cugaaccag auuucucacg cccaaccgg aaguacaccu ggaacgaugu gggccagcuu	600
guggaaaagc ucaaaaagga guggaacgug aucugcauuu cugacguggu guacaaccu	660
acugcugcua auuccaagug gaucaggaa cauccugagu gugcuuacaa ccuggucaac	720
uccccacacc ugaagccgc cugggugcug gaucgggcuc ucuggcgguu cuccugcgau	780
gucgccgaag gaaaguacaa ggagaagggg auccugcgc ugaucgaaaa cgaccaucac	840
augaauucca ucaggaaaau caucugggag gacaucucc cuaagcucaa gcugugggag	900
uuuuucaaag ucgacugaa caaggccgug gagcaguua gacggcugcu gacacaagag	960
aaccgccgcg ucaccaagag cgacccgaau cagcaccuga cgauuaucca ggaccugag	1020
uacagaagau ucggcugcac cguggauaug aacauugcuc ugaccaccuu caucacccac	1080
gauaagggucc ccgcagccau ugaggagugc ugcaacuggu uccacaagag gauggaagaa	1140
cugaauuccg agaagcaccg gcugaucaau uaucaccaag aacaagccgu gaacugccuu	1200
cugggaaacg uuuuuacga gcgccuggcg ggacaugggc caaagcucgg gcccgugacc	1260
agaaagcacc cacuggucac ccgcuacuuc accuucccg ucgaggaaa ucgacuucagc	1320
auggaagagu cgaugaucca cuugccgaac aaggccugcu uccuauaggc ccauaacggu	1380
ugggucaugg gcgaugaccc ucuucgcaac uucgcggagc ccgguucuga agucuaucuu	1440

ULPI_041_01WO_SeqList_ST25.txt

cggcgggaac ugauuugcug gggagacucc gucaagcugc gcuacggaaa caagcccag	1500
gauugcccgu acuugugggc ccuaugaag aaguacacug agaucacugc cacuuacuuc	1560
caaggagugc gccucgacaa uugucacagc acuccgcugc acguggcgga guacaugcug	1620
gacgccgcuc gcaacuugca gccuaaucug uaugucgugg ccgaguuguu caccggcucg	1680
gaagaucugg acaacguguu cgucacucgc cugggaau cuccccugau ccgggaagcg	1740
augagcgccu acaacuccca cgaagagggg cgguuggugu accgcuacgg cggagagccu	1800
gucggaaguu ucgugcagcc cugucugagg cccugaugc ccgcuaucgc ccaugcgcug	1860
uuuauggaca ucacccacga caacgaugc ccuaauuguc accgcuccgc cuaugaugcc	1920
cugccucca ccacuauugu cagcauggcc ugcugcgccu cgggguccac ccggggauac	1980
gacgaacugg ugccccacca aauuuccgug guguccgagg aacgguucua caccaagugg	2040
aaccugaag cgugccauc gaacacugga gaagugaacu uucagucggg aauuaucgca	2100
gcccgaugcg ccaucagcaa gcugcaccag gaacucggcg caaaggguuu uauccaaguc	2160
uacguggacc aggucgacga ggacauuguc gccgugaccc ggcacucccc auccaaccac	2220
cagucugugg uggcuguguc aaggacggcu uuccggaacc caaagaccag cuucuacagc	2280
aaggaagugc cucagaugug caucccgggg aagaucgaag aaguggugcu ggaggccaga	2340
accaucgaaa gaaacaccaa gcccuauagg aaggacgaga acucgaucaa cgguaucuccg	2400
gacauuaccg ucgagauacg cgagcacauu cagcugaacg aguccaaaau cgugaagcag	2460
gccggggugg ccacgaaggg ucccaacgag uacauucagg agaucgaguu cgagaaccug	2520
agccccgggu ccgugaucau cuuccgcgug ucccuggacc cccacgcca aguggccgug	2580
ggcauccugc ggaaccaucu gaccaguuc uccccgcacu ucaagagcgg cucgcucgcc	2640
guggacaacg cggacccgau ccucaagauc ccuuucgcau cgucggccuc ccgccugacc	2700
cuggccgaac ugaaucau cuuguaccga ugcgaauacg aagagaagga ggacgggggg	2760
ggcugcuacg auaucccaa cuggagcgcg uugaaguacg caggauugca gggauugaug	2820
uccguccucg cugaaauccg cccgaagaac gaccugggac acccguuuug caacaaccug	2880

ULPI_041_01W0_SeqList_ST25.txt

agaucagggg auuggaugau cgauuacgug ucgaacagac ugaucucgcg cagcggcacu	2940
auugccgaag uggggaagug gcuccaggcc auguucuucu accugaagca gaucuccugg	3000
uacuugaucc cuuguuacuu cgacgccauc cugaucggag ccuacaccac ccugcuugac	3060
acugcaugga agcagauguc cagcuucgug caaaacggaa gcaccuucgu gaagcaccug	3120
ucccugggau ccgugcagcu cugcggcgug ggaaaguuuu cgucccucc cauccugagc	3180
ccugcccuua uggacgugcc guacaggcug aacgaaaaua ccaaagagaa ggagcaaugc	3240
uguguguccu uggcggccgg auugccgau uucuccuccg ggauuuuccg gugcugggga	3300
cgggacaccu uuaucgcacu gagggguauu cuccugauca ccggucgcu cugggaggcu	3360
cgcaacauua uucuggccuu cgcgggcacg cuuagacaug gauugauccc uaaccuucug	3420
ggagaaggga ucuacgcgcg guacaacugc cgcgaugccg ugugguggug gcugcagugc	3480
auccaggacu acugcaaaau ggugccgaau ggucuggaua uccugaagug uccgguuucg	3540
cggauguacc cuaccgacga cagcgccccu cuccggcccg gcacucucga ccagccccua	3600
uuugaagua uccaggaggc caugcaaaag cacaugcagg gcacacaguu cagagagagg	3660
aacgccggac cgagauuga ccggaacaug aaggacgagg gauucaacau uaccgcggga	3720
guggaugagg aaacugguuu cguguacggc ggaaaccggu uuaacugcgg cacuuggaug	3780
gacaagaugg gagaauccga ccgcgcccga aaccgcggaa uucccgccac ucccgcgcac	3840
ggcuccgccg uggaaaauugu gggacuguca aaguccgcag uccgcuggcu gcuggaacuc	3900
ucaagaaga acaucuucc guaccacgag gucaccguga agcggcacgg caaagcgauc	3960
aaagugucgu acgacgagug gaaccggaag auucaggaua acuucgagaa gcuguuucac	4020
guguccgaag auccaagcga ccugaacgag aagcauccca acuuggugca caagcgccgc	4080
aucuacaagg auuccuacgg agccagcagc ccguggugcg acuaccaacu ccgccccaac	4140
uucaccaucg ccaugguggu ggcgccggag cuguucacga cggagaaagc uuggaaggcu	4200
cucgaaaucg cggagaagaa gcugcugggu ccucugggga ugaaaacccu ggacccggac	4260
gauauggugu acugugggau cuacgacaac gcccuggaca acgacaacua caaccucgcc	4320

ULPI_041_01WO_SeqList_ST25.txt

aagggguuca acuaccacca gggacccgaa uggcucuggc caaucggaua cuuccugaga	4380
gcgaagcuuu acuucucgcg gcugaugggu ccugaaacca cggccaagac caucgugcuc	4440
gugaaaaaug ugcugucaag gcacuacgug caucuggaga ggucgccaug gaaggguug	4500
ccggaacuga ccaacgaaaa cgcacaguac ugccccuuuu cgugcgagac ucaggccugg	4560
uccaucgcca ccuucucga aacucucuac gaccuguag	4599

<210> 26

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 26

augggacacu cgaacagau cagaauccug uugcucaacg aaauggagaa gcuggaaaag	60
accucuuuc ggcucgaaca gggcuacgag cugcaguucc gccugggucc gacacuacaa	120
ggaaaggcag ucacugugua caccaacuac ccuuuuccg gagaaaccuu caaccgggag	180
aaguuccggu cccuggacug ggaaaacca accgaacgag aggaugacuc cgacaaguac	240
ugcaagcuga accuccaaca guccggauca uuccaguacu acuuucugca aggggaacgag	300
aaguccggag gcggcuacau cguggucgac ccauacuua gagugggagc cgacaaucau	360
guccugccuc uggacugcgu gaccugcaa accuucugg cgaaaugucu gggcccguuc	420
gaugaguggg aaagccgccu cagagucgca aaggaguccg gauacaacau gauucacuuc	480
acuccgcugc aaaccucgg ucuguccgg ucgugcuauu cucuggcgaa ccagcuggag	540
cuuaacccccg acuucucgcg ccaaaccgg aaguacaccu ggaacgacgu cgggcagcuu	600
guggaaaagc ugaagaagga guggaacgug aucugcauca cugacguggu guacaaccau	660
acggccgcca acucgaagug gauccaagag caucccgagu gcgcguacaa ccucgugaau	720
agcccgauc ugaagccugc ugggugcug gauagagccc ucuggagauu cagcugcgac	780
guggccgagg ggaaguacaa agaaaaggga auuccggccu ugauugagaa cgaccaucac	840

ULPI_041_01W0_SeqList_ST25.txt

augaacucuaa uccgcaagau caucuggggag gauaucuucc cuaagcuuaa guuguggggag	900
uucuuccaag uggacguaaa caaggcagug gagcaguuca gaagguugcu gacucaagaa	960
aacagacggg ucacuaaguc cgauccuaac cagcaccuua ccaucauua agaccugag	1020
uaccgccggu uuggcugcac cgucgacaug aacaucgccc ugaccacuuu caucccgcau	1080
gacaagggcc cggcggcaau cgaggaaugc uguaacuggu uucauaagag gauggaggaa	1140
cugaaauccg agaagcaccg gcugauuaac uaccaccagg aacaggcagu caacugccuc	1200
cugggcaacg uguuucacga acggcuggcu ggacacggac cgaagcuggg ucccugacu	1260
cgcaagcauc cgcucgugac ucgcuacuuc accuucccg uugaggagau ugacuucucc	1320
auggaagaau ccaugaucca ccuccgaac aaggcuugcu uccugauggc gcacaacgga	1380
ugggucaugg gcgacgacc acugcgcaau uucgcugagc cuggcucgga ggucuaccug	1440
agaagggaau ugauuugcug gggagacucc gugaagcugc gcuauggcaa caagccugaa	1500
gauugccccc accugugggc ucacaugaag aaguacacgg aaauacugc cacguacuuc	1560
caggagaucc ggcuggacaa uugccacucg accccgcucc auguggccga guacaugcug	1620
gaugcagcga ggaauugca gcccaaucug uacgugguug cagaacuguu cacuggcucc	1680
gaggaccucg acaacguguu cgugaccaga cuggggauuu ccucacugau ccgggaagcc	1740
augucggccu acaacuccca ugaagagggc cgccuggugu accgcuaugg gggagaaccc	1800
gugggaagcu ucgugcagcc uugccuccgg ccgcugaugc cugcgaucgc ccacgccug	1860
uucauggaua ucacucacga caacgaugc ccuauuguc aucgcucggc cuacgacgca	1920
cugccuucga ccacuaucgu guccauggcc ugcugcgccu ccgggagcac ccgcggauac	1980
gaugaacucg ugccgcacca gaucagcgug guguccgaag aacgguucua caccaagugg	2040
aacccgaag cccugccuag caauaccggg gaagugaacu uccaguccgg uauuauccg	2100
gcucgcugcg ccaucuccaa acuccaccaa gagcucggug ccaagggaau cauuaaguc	2160
uacguggauc aggucgacga agauauugug gccgugacca ggcacucacc uuccauccac	2220
caauccgucg ucgccguguc ccggacugcg uuucggaacc ccaagacuuc guucuacucg	2280

ULPI_041_01WO_SeqList_ST25.txt

aaagaagugc	cacagaugug	uaucgggga	aaaaucgagg	aggucgugcu	cgaagcccgg	2340
accuuugaga	ggaacaccaa	gccuuaccgg	aaagacgaga	acucuaucua	cgguaccccu	2400
gauauuacug	uggagaucg	cgaacacauc	cagcugaacg	aaucuaagau	cgucaagcag	2460
gcuggagugg	ccaccaaggg	acccaacgag	uacauccagg	agaucgaauu	ugaaaaccuc	2520
uccccuggc	ccgugauuau	cuuccgggug	ucccuggacc	cucacgcca	aguggccgug	2580
ggaauucua	gaaaccaccu	gaccagauuc	ucaccacacu	uuaguccgg	uucccuggcg	2640
guggacaacg	ccgauccgau	cuugaagauc	ccuucgc	cgucgccuc	ccgccugacu	2700
cucgcggaac	ugaaccagau	ccuguaccgc	ugugaauccg	aggaaaagga	ggacggcggc	2760
ggcuguuacg	auaucccaa	uuggucggcu	uugaaauacg	cgggacuua	ggggcugaug	2820
ucugugcugg	cggaaauccg	gccgaagaac	gaccugggac	accuauucug	caacaacuug	2880
cggagcggag	acuggaugau	cgauuacguc	agcaacagau	ugaucagccg	gagcggcacu	2940
aucgccgagg	ucggaaagug	gcuccaggcc	auguucuucu	accugaagca	gaucggccga	3000
uaccucaucc	ccuguuacuu	cgacgccauu	cugaucgggg	ccuacaccac	ccugcuggac	3060
accgccugga	agcagaugag	caguuuugug	caaacgggu	ccaccuucgu	gaagcaccuu	3120
ucacugggcu	cagugcagcu	cugcggcgug	ggaaaguucc	ccucgcugcc	cauucugagc	3180
cccggccuga	uggacgucc	uuaccggcug	aacgagauca	ccaaggagaa	ggagcagugc	3240
uguguuucc	uggcugccgg	gcugccacac	uucucguccg	gcaucuuccg	guguuggggc	3300
cgggauaccu	ucauugcccu	gcggggauc	cugcuuauca	ccggucgcu	cguggaggcu	3360
cggaaacaua	uucucgcguu	cgccggcacc	cugagacacg	gucugauucc	gaaucuguug	3420
ggcgaaggaa	ucuacgccag	auacaacugu	cgggacgccg	ugugguggug	gcuccagugc	3480
auucaggacu	auugcaagau	ggugccgaac	ggccuggaca	uccugaagug	cccagugucg	3540
aggauuauc	caacugacga	cagcgcaccu	cugccggccg	ggaccucga	ccaacccug	3600
uucgaaguca	uucaagaggc	uagcagaag	cacaugcagg	guauucaguu	ccgggagcgg	3660
aacgcggggc	cccagauuga	uaggaacaug	aaagacgagg	gcuucaacu	cacugccggc	3720

ULPI_041_01WO_SeqList_ST25.txt

guggacgaag aaaccgguuu uguguacgga ggaaaccgcu ucaacugcgg uaccuggaug	3780
gacaagaugg gagaauccga ucgcgcgcgc aacagaggga ucccggcaac cccucgggac	3840
ggauccgcgg uggaaaauugu gggacugagc aagagcgccg ugcgguggcu ccuggaacug	3900
uccaaaaaga acaucuucc cuaccacgaa gugaccguga agagacacgg aaaggccauc	3960
aaagucucau acgaugaaug gaacaggaag auccaggaua acuucgagaa gcuguuucac	4020
guguccgagg aucccuucca ucugaacgag aagcauccga aucuggugca caagcgcggg	4080
aucuacaagg acucguacgg agcguccucc ccuuggugcg acuaucagcu gcggccuaac	4140
uucaccauug ccauggucgu ggccccggag cuguucacaa cugagaaggc cuggaaggcc	4200
cuugaaaauug ccgagaagaa gcugcugggg ccuuugggga ugaaaacccu ggauccggac	4260
gacauggugu acugcggaau cuacgacaac gcccuggaca acgauaacua caaucucgcg	4320
aagggcuuca auuaccacca aggccccgaa uggcucuggc cuauugggua cuuccugcgc	4380
gccaagcugu acuucucacg gcugauggga ccagagacua ccgccaagac uaucguccuc	4440
gugaagaacg ugcugucccg gcacuacgug caucuggaga ggagcccuug gaagggacuu	4500
ccugagcuga cgaacgaaaa cgcgagacuac ugccccuucu ccugcgaaac ccaggcuugg	4560
uccauugcca cuauacugga aaccuuauau gaccuguag	4599

<210> 27
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 27	
augggacacu cgaaacagau ccggauccug uugcugaacg aaauggagaa gcuugaaaag	60
accuguuuuc ggcuaggagca gggcuacgag cugcaguucc gccugggucc gacacuacaa	120
ggaaaaggccg ugacugugua caccaacuac ccauuucccg gagaaacuuu caaccgggag	180
aaguuccggu cccuggacug ggaaaacca accgaacgag aggacgacuc ggacaaguac	240

ULPI_041_01WO_SeqList_ST25.txt

ugcaagcuga accuccaaca guccggauca uuccaguacu acuuucugca agggaaacgag	300
aaguccggag gcggcuacau cguggucgac ccgauacuua gagugggcgc cgacaaucau	360
guccugccuc uggauugcgu gaccugcaa accuucugg ccaaaugucu gggcccguuc	420
gaugaguggg aaagccgccu cagagucgca aaggaguccg gauacaacau gauucauuuc	480
accccgcugc aaaccuggg ucugucccg ucgugcuauu cucuggcgaa ccagcuggag	540
cuuaacccc acuucucgcg cccaaaccgg aaguacaccu ggaacgacgu cgggcagcuu	600
guggaaaagc ugaagaagga guggaacgug aucugcauca cugacguggu guacaaccau	660
acggccgcca acucgaagug gauccaggag caucccgagu gcgcuaacaa ccucgugaac	720
agcccgcacc uuaagccugc uugggugcug gacagagccc ucuggagauu caguugcgac	780
guggccgagg ggaaguacaa ggaaaaggga auuccggccu ugaucgagaa cgaccaucac	840
augaacucua ucaggaagau caucugggag gacaucuucc cuaagcuuaa guugugggag	900
uucuuccaag uggacguaaa caaggcagug gagcaguuca gaagguugcu gacccaagaa	960
aacagacgcg ucacuaaguc cgauccuaac caacaccuua ccaucauua agaccugaa	1020
uaccgccggu uuggcugcac cgucgauaug aacaucgccc ugaccacuuu caucccgcau	1080
gacaagggcc cggcggcaau cgaggaaugc uguaacuggu uucauaagag aauggaagaa	1140
cugaaauccg agaagcaccg gcugauuaac uaccaccagg aacaggcagu gaacugccug	1200
cugggcaaug uguucucaga acggcuggcu ggacacggac cgaagcuggg ucccugacu	1260
aggaagcauc cgcucgugac ucgcuacuuc accuucccg uugaggagau ugacuucug	1320
auggaagaau ccaugaucca ucugccuaac aaggcuugcu uccugauggc gcacaacgga	1380
ugggucaugg gcgacgacc ccugcgcaac uucgccgagc cuggcucgga ggucuaccug	1440
agaagggaac uuauucuguug gggagacucc gucaagcugc gcuacggaaa caagcccga	1500
gauugcccu accugugggc ucacaugaag aaguacacgg aaucacugc aacguacuuc	1560
caggagucc ggcuggacaa uugccacucc acccccuuc auguggccga guacaugcuc	1620
gaugcagcga ggaauucgca gccgaaucug uacgugguug ccgaacuguu cacuggcucc	1680

ULPI_041_01W0_SeqList_ST25.txt

gaggacuugg acaacguguu cgugaccaga cuggggau cuccucugau ccgggaagcc	1740
augucggccu acaacucca ugaagagggc cgccuggugu accgcuacgg gggagaaccc	1800
gugggaagcu ucgugcaacc uugccugcgg ccgcugaugc cugcgauccg ccacgcccug	1860
uucauggaca ucacucacga uaacgaugc ccgauugugc aucgcucggc cuacgacgca	1920
cugccgagca ccacuaucgu guccauggcc ugcugcgccu ccgggagcac ucgcggauac	1980
gaugaacucg ugccgcacca gaucagcgug guguccgaag aacgcuucua uaccaagugg	2040
aaccccgaag cgugccauc gaauacuggc gaagugaacu uccaguccgg uauuauccg	2100
gcucgcugug ccaucagcaa acugcaccaa gagcuuggug ccaagggauu cauuaaguc	2160
uacguggauc aggucgacga agauauugug gccgugacca ggcacucacc uuccauccac	2220
caucagucg uggccguguc ccggaccg cg uuccggaacc ccaagaccag cuucuacucg	2280
aaagaagugc cucagaugug uaucccggga aaaaucgaag aggucgugcu ggaagcccgg	2340
accuuugaga ggaacaccaa gccuuaccgg aaagacgaga acucuaucua cgguaccccu	2400
gauauuacug uggagaucgg cgaacacauc cagcugaacg aaucaaagau cgucaagcag	2460
gcuggagugg ccaccaaggg acccaacgag uacauccagg agaucgaauu ugaaaaccuc	2520
uccccuggcu ccgugauuau cuuucgggug ucccuggacc cucacgcccag aguggccgug	2580
ggaaauucua gaaaccaccu gaccaguu ucacccacu ucaaguccgg uucccuggcg	2640
guggauaacg ccgaccgcu uuugaagau cccuucgcu cgucggccuc ccgcccugcu	2700
cucgcggaac ugaaccagau ccuguaccgc ugugaauag aggaaaaaga ggacggcggc	2760
ggcuguuacg auauuccaa uuggucggcu uugaaauac cgggacuua ggggcugaug	2820
ucugugcugg cggaauccg gccgaagaac gaccugggac acccauucug caacaacuug	2880
cgguccggag acuggaugau cgauuacguc agcaacagau ugaucagccg gagcggcacu	2940
aucgcugagg ucggaaagug gcugcaggcc auguucuuc aucugaagca gaucggcgga	3000
uaccucaucc ccuguuacuu cgacgccauu cugaucgggg ccuacaccac ccugcuggac	3060
accgccugga agcagaugag caguuuugug caaacgggu ccaccuucgu gaagcaccuu	3120

ULPI_041_01WO_SeqList_ST25.txt

ucacugggcu cagugcagcu cugcggcgug ggaaaguucc ccucucugcc cauucugagc	3180
ccggcccuga uggacguccc uuaccggcug aacgagauca ccaaggagaa ggagcagugc	3240
ugcguuuccc uggcugccgg gcugccacac uucucguccg gcaucuuccg gugcuggggc	3300
cgggauaccu ucaucgccc ugcgggguau cugcuuauca ccggucgcu cguggaggcu	3360
cggaacauua uucuggcggu cgccggcacc cuuagacacg gacugauucc gaaccuuuug	3420
ggcgaaggaa ucuacgccag auacaacugu cgggacgccg ugugguggug gcuucagugc	3480
auucaggacu auugcaagau ggugccgaac ggccuggaca uccugaagug cccagugucg	3540
aggauguauc caaccgacga cagcgcaccu cugccggccg ggaccucga ccaaccucg	3600
uucgaaguca uucaggaggc uaugcagaag ccaugcagg guauucagu cgggagcgg	3660
aacgcggggc cgagauuga uaggaacaug aaagacgagg gcuucaacu cacugccggc	3720
guggacgaag aaaccgguu uguguacgga ggaaacagau ucaacugcgg uaccuggaug	3780
gacaagaugg gagaguccga ucgcgcgcgc aacagaggga ucccggcaac cccgcgggac	3840
ggauccgcgg uggaaauugu gggacugagc aagagcgccg ugcgguggcu gcuggaacug	3900
agcaaaaaga acaucuucc cuaccacgaa gugaccguga agcggcacgg aaaggccauc	3960
aaagucucu acgaugaug gaauaggaag auccaggaua acuucgagaa gcuguuucac	4020
guguccgagg auccuccga ucugaacgaa aagcaccga aucucgugca caagcgcggg	4080
aucuaaagg acucguacgg agcguccucc ccuuggugcg acuaucagcu gcggccuac	4140
uucaccauug ccauggucgu ggccccggag cuguucacaa cugagaaggc cuggaaggcc	4200
cuugaaaauug ccgagaagaa gcugcugggg ccuuugggga ugaaaacccu ggauccggac	4260
gacauggugu acugcggaau cuacgacaac gcccuggaca acgauaacua caaucuggcc	4320
aaggguuca auuaccacca gggccccgaa uggcuguggc ccuugggua cuuccugcg	4380
gccaaagcugu acuucucacg gcugauggga ccagagacua ccgccaagac uaucguccuc	4440
gugaagaacg ugcugucccg gcacuacgug caucuggaga ggagcccuug gaagggacu	4500
ccugaacuga cuaacgagaa cgcgcaguac ugccccuucu ccugcgaaac ccaggcuugg	4560

ULPI_041_01W0_SeqList_ST25.txt

uccauugcca ccuacugga aacccuuuau gaccuguag 4599

<210> 28
<211> 4599
<212> RNA
<213> Artificial Sequence

<220>
<223> Codon-Optimized AGL Coding Sequence

<400> 28
augggacacu cgaaacagau uagaauccug uugcucaacg aaauggagaa gcuggaaaag 60
accucuuuuc ggcuggagca gggcuacgag cuccaguucc gccugggucc gacacuacaa 120
gggaaggcag ugaccgugua caccaacuac ccuuuucccg gagaaaccuu caaccgggag 180
aaguuccggu cccuggacug ggaaaacca accgaacgag aggaugacag cgacaaguac 240
ugcaagcuga accuccaaca gucgggaucg uuccaguacu acuuucugca aggggaacgag 300
aaguccggag gaggcuaau cguggucgac ccgauacuua gagugggagc cgacaaccu 360
guccugccuc uggacugcgu gaccugcaa accuucugg ccaaagucu gggcccguuc 420
gaugaguggg aaagccgccu cagaguggcc aaggaguccg gguacaacau gauucacuuc 480
acuccgcugc aaaccucgg ucugucccg ucugcuauu cucuggcgaa ccagcuggag 540
cugaacccccg acuuucuccg ccaaaccgg aaguacaccu ggaacgacgu cgggcagcuu 600
guggaaaagc ugaagaagga guggaacgug aucugcauu cugacguggu guacaaccac 660
accgccgcca acucgaagug gauccaagaa caucccgagu gcgcguacaa ccucgugaac 720
agcccgcauc ugaagccggc uugggugcuc gauagagccc ucuggagauu cuccugugac 780
gucgccgagg ggaaguacaa agagaagggu auuccggccc ucauugagaa cgaccaucac 840
augaacucaa uccggaagau caucugggag gauaucuucc cuaagcuuaa guugugggag 900
uucuuccaag uggacguaaa caaggccgug gaacaguuca gaaggcugcu gacucaagaa 960
aaccgccgcg ucacuaaguc cgauccuac cagcaucuua ccuauacca agaccugag 1020
uauccgagg uugggugcac cgucgacug acaucgcac ugaccacuuu caucccgau 1080

ULPI_041_01WO_SeqList_ST25.txt

gacaaggggc	cggcggccau	cgaggaauc	uguaacuggu	uucuaaagag	gauggaggaa	1140
cugaauuccg	agaagcaccg	gcugauuaac	uaccaccagg	aacaggcagu	gaauugccuc	1200
cuggggaacg	uguucucga	aaggcuggcu	ggacacggac	cgaagcuggg	gcccugacu	1260
cgcaagcauc	cgucugac	ucgcuacuuc	accuucccg	uugaggagau	ugacuucua	1320
auggaagaau	ccaugaucca	ccuccgaac	aaggcuugcu	ucuugauggc	gcacaugga	1380
ugggucaugg	gcgacgacc	acugcgcaac	uucgcugagc	cuggcucgga	ggucuaccug	1440
agaagggaau	ugauuugcug	gggggacucc	gucaagcugc	gcuauggaaa	caagccggaa	1500
gauugcccu	accugugggc	ucacaugaag	aaguacacgg	aaauuacugc	cacguacuuc	1560
cagggcgucc	ggcuggacaa	cugccacucc	acuccccucc	auguggccga	guacaugcuc	1620
gacgcagcga	ggaaucugca	gcccuaucug	uacgugguug	cagaacuguu	cacugggucc	1680
gaggaccucg	acaauuguu	cgugaccaga	cuggggauuc	ccuccugau	ccgggaagcc	1740
augucggccu	acaacuccca	ugaagagggc	cgccuggugu	accgcuacgg	gggagaaccc	1800
guggggagcu	ucgugcagcc	uugccuccgc	ccgcugaugc	cugccaucgc	ccacgccug	1860
uucauggaua	ucacucauga	caacgaau	cccauugugc	aucgcucggc	cuacgacgca	1920
cugccuucga	ccacuau	guccauggcc	ugcugcgccu	ccgggagcac	caggggauac	1980
gacgaacucg	ugccgcacca	gaucagcgug	guguccgaag	agagauucua	uaccaagugg	2040
aaccccgaag	cgugcccag	caauaccggg	gaagugaacu	uccaguccgg	uauuauccg	2100
gcucgcugug	cgauagcaa	gcuccaccag	gagcucggug	ccaagggauu	cauuaaguc	2160
uacguggacc	aggucgacga	agauaucgug	gccgugacca	ggcacucacc	uuccauucac	2220
caauccgugg	ucgccguguc	ccggacugcu	uuucggaacc	ccaagacuuc	guucuacucg	2280
aaagaagugc	cacagaugug	uaucggggg	aaaaucaag	aggucguccu	cgaagcccgg	2340
accuuugaga	ggaacaccaa	gccuuaccgg	aaagacgaga	acucuaucua	cgggaccccu	2400
gauauuacug	uggagaucg	cgaacacauc	cagcugaacg	aaucuaagau	cgucaagcag	2460
gcuggagugg	cgaccaaggg	acccaacgag	uacauccagg	agaucgaauu	ugaaaaccuc	2520

ULPI_041_01W0_SeqList_ST25.txt

uccccuggcu	ccgugauuau	cuuccgggug	ucccuggacc	cucacgcca	aguggccgug	2580
ggaauucu	gaaaccaccu	gacucaguuc	ucacccacu	uuaagucugg	guccuggcg	2640
guggacaacg	ccgauccgau	cuugaaaauc	ccuuucgcaa	gccuggccuc	ccgccugacu	2700
uuggccgagc	ugaaccagau	ccuguaccgc	ugugaaucag	aggaaaagga	ggacgggggu	2760
ggcuguuacg	auaucccaa	cugguccgcu	uugaaauacg	cgggacugca	ggggcugaug	2820
uccgugcugg	cggaaucag	accgaagaac	gaucuggggc	acccuucug	caacaacuug	2880
cgguccggag	acuggaugau	cgauuacguc	agcaaccggu	ugaucagcag	aagcgguacu	2940
aucgccgagg	ucggaaagug	gcugcaggcc	auguucuucu	accugaagca	gaucccucga	3000
uaccucaucc	ccuguuacuu	cgacgccauu	uugaucgggg	ccuacaccac	ccugcuggac	3060
acugccugga	agcagaugag	caguuuugug	caaaugggu	cgaccuucgu	gaagcaccuu	3120
uccuugggcu	cagugcagcu	cugcggcgug	gggaaguucc	ccucgcugcc	cauucugucc	3180
ccggcccuga	uggacguccc	uuaccggcug	aacgagauua	ccaaggagaa	ggagcagugc	3240
uguguuucc	uggcugccgg	gcugccacac	uucucguccg	ggaucuuccg	gugcuggggc	3300
cgcgauaccu	ucauugcgcu	gcgggguauc	cugcuuauca	ccggucgcu	cguggaggcu	3360
cggaacauua	uccuugcauu	cgccgguauc	cugagacacg	gucugauccc	gaauucuc	3420
ggggaaggaa	ucuacgcaag	auacaacugc	cgggacgccg	ugugguggug	gcuccagugc	3480
auucaggacu	auugcaagau	ggugccgaac	ggacuugaca	uccugaagug	cccagugucg	3540
aggauguacc	cuaccgacga	cagcgcuccu	cugccggccg	ggaccugga	ccaaccucg	3600
uucgaaguca	uccaagaggc	uaugcagaag	cacaugcagg	guauucaguu	ccgggaacgg	3660
aacgcggggc	cccagauuga	uaggaacaug	aaggacgagg	guuucacau	cacugccggc	3720
guggacgaag	aaacuggguu	uguguacgga	ggaaacagau	ucaacugcgg	uaccuggaug	3780
gacaagaugg	gagaauccga	ucgcgcgcgc	aacagaggga	ucccggaac	cccgcgggac	3840
ggauccgcgg	uggaaaugu	gggacugagc	aagagcgccg	ugcgguggcu	ccuggaacug	3900
uccaaaaaga	acaucuucc	cuaccacgaa	gugaccguga	agcggcacgg	aaaggccauc	3960

ULPI_041_01WO_SeqList_ST25.txt

aaagucuc	acgaugaaug	gaaucggaag	auccaggaua	acuucgagaa	gcuguuucac	4020
guguccgagg	auccuccga	ucucaacgaa	aagcauccga	aucucgugca	caagcgcggg	4080
aucuacaagg	acucguacgg	ggcguccuca	ccuuggugcg	acuaucagcu	gcggccuaac	4140
uucacuaauug	cgauggucgu	ggccccggag	uuauucacaa	cggagaaggc	cuggaaggcc	4200
cuugaaaauug	cggagaagaa	gcugcugggg	ccucucggga	ugaaaacccu	ggauccggac	4260
gacauggugu	acugcggaau	cuacgacaac	gccccuugaca	acgauaacua	caaucuggcc	4320
aagggguuca	auuaccacca	ggggccggaa	uggcucuggc	ccauugggua	cuuccugcgc	4380
gccaagcugu	acuucucacg	gcugaugggg	ccagagacua	ccgccaagac	uaucguccuc	4440
gugaaaaacg	ugcugucccg	gcacuacgug	caucuggaga	ggagcccuug	gaagggacug	4500
ccagagcuga	cgaacgagaa	cgcgcaguac	ugcccccucu	ccugcgaaac	ccaggcuugg	4560
ucgauugcca	cuauacugga	aaccuuauau	gaccuguag			4599

<210> 29

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 29

augggacacu	caaagcaau	ccgcauucug	cugcugaacg	agauggaaaa	gcuggaaaag	60
acucuguucc	ggcuggaaca	ggguuacgag	cugcaguuuc	ggcucggccc	aaccuugcaa	120
gggaaggccg	ucacugucua	caccaacuac	ccuuuuccgg	gggagacuuu	caaccgcaaa	180
aaguuccgcu	cccuggauug	ggaaaacccc	acugaacggg	aggauagacag	cgacaaguac	240
ugcaaacuga	accuucagca	guccggcagc	uuccaauacu	acuuccugca	aggaaacgaa	300
aaguccggcg	gcgguuacau	cguggucgac	cccuuucuga	gaguggggagc	cgauaaucan	360
gugcugcccc	uggacugcgu	caccucgcaa	acuuuccugg	cuaagugccu	gggaccguuc	420
gaugaguggg	aauccaggcu	ccgcguggcg	aaggagucgg	gcuacaacau	gauucauuuc	480

ULPI_041_01WO_SeqList_ST25.txt

acucuccuuc aaacucuggg gcugagucgc agcuguuacu ccugggccaa ccagcuugaa	540
cugaacccag acuuuuuccg gccaaacaga aaguacaccu ggaacgacgu gggacaauug	600
guggagaagc uuaagaagga auggaacgug aucugcauca ccgauguggu guauaaccac	660
acugccgcga auuccaagug gauccaggaa cccccgagu gugcuuacaa ccuugugaac	720
agcccucacc uuaagccggc cugggugcug gaucgggcgc uguggagauu cuccugcgac	780
guggccgaag ggaaguacaa agaaaaggga aucccgccc ugauugaaaa ugaccaccu	840
augaacucca uucggaaaau cauugggaa gauaucuuuc ccaagcugaa gcucugggag	900
uucuuucaag uggaugugaa caaggcugug gagcaguucc ggcggcugcu gaccaggag	960
aaccgccgcg ugaccaaguc cgaccuaac cagcacuuga ccaucauaca ggaccggaa	1020
uaccggagau ucggcugcac ugucgacaug aacauugccc ucacuaccuu cauccgcgcau	1080
gacaagggac ccgcugccau cgaagagugc ugcaacuggu uccacaagcg gauggaagaa	1140
cugaacucug aaaagcacag gcugaucaac uaccaccagg aacaggcugu gaacugccug	1200
cugggcaacg uguucuacga gagacucgca ggacacggac cgaagcuggg ccugugacc	1260
cggaagcauc cucuggucac ccgcuacuuc accuucccg ucgagagau cgacuuuucg	1320
auggaagaau cgaugaucca ccuccuaac aaggccugcu uccucauggc gcacaacggc	1380
ugggucaugg gcgacgaccc gcugagaaac uucgccgagc ccgggagcga aguguaccuc	1440
cggcgggaaac uuauuugcug gggagauagc gugaagcuua gauauggcaa caagccugag	1500
gacugcccau accugugggc gcacaugaag aaguacacug aaauuaccgc gaccuacuuc	1560
caaggagucc gacucgacaa cugccacagc accccacuuc acgucgcgga guacauguug	1620
gaugccgcac ggaauccu	1680
gaggaccuug acaauguguu cgugacuaga cuggggau	1740
auguccgcgu acaacucca cgaagaggga cggcuggugu accgcuacgg cggagagccc	1800
gugggaagcu ucgugcagcc cugccugcgg ccucugaugc cggccaucgc ucacgcccug	1860
uucauggaua ucacucacga caaugagugu ccuauugc acaggagcgc guacgacgcc	1920

ULPI_041_01W0_SeqList_ST25.txt

cugccucca cuacuaucgu gucgauggcc ugcugcgcaa gcgguucuac ccgcgguuac	1980
gacgagcuug ucccgcacca aaauuccgug gugucagagg aacgguucua caccaagugg	2040
aacccggagg cccugccuuc aaacaccggc gaagugaacu uccaguccgg aaucauugcc	2100
gcccgcugug ccauuucaaa guugcaccag gagcugggcg ccaagggauu cauucagguc	2160
uacguugacc aggucgacga agauaucgug gccguuacua gacauucacc gagcauccau	2220
cagagcgugg ucgcagucag caggacugcc uuccgcaacc cgaaaaccuc guucuacucc	2280
aaggaagugc cccagaugug uaucccgga aaaaauaggg aaguggugcu ggaggcccgg	2340
accaucgagc ggaacacuaa gcccuaccgg aaggacgaga auucaaucaa cggaaccccu	2400
gacaucaccg uggagaucgg cgagcauau caacugaacg agucgaagau cgucaagcag	2460
gcuggggugg caacuaaggg cccuaacgag uacauucagg agauugaauu cgagaaccug	2520
ucccccgggu ccgugaucau uuuccgcgug ucccuggacc cacaugcuca aguagcagug	2580
gggauccuga gaaaccaccu gaccaguuu agcccgcacu ucaaguccgg aucccuggcc	2640
guggauaacg ccgacccgau ccugaagau cccuuugcau cccuggcuuc ccggcugacc	2700
uuggccgaac ugaaccagau ucuguaccgc ugcgaauacg aggaaaagga ggacggaggc	2760
ggaugcuacg auaucccaa uuggucggcg cugaaguacg ccggccuua aggacugaug	2820
uccgugcugg ccgagaucag gccgaagaau gaccugggac acccguucug caacaacuug	2880
agaagcgggg acuggaugau ugauuacgug ucgaaccggc ugaucucccg gagcggcacc	2940
aucgcggaag ucggaaagug gcuccaggcc auguucuucu accugaagca gauccccgc	3000
uaccugaucc ccugcuacuu cgacgcgauc uugauugggg cauacaccac ucugcuugac	3060
accgcuugga agcagauguc cuccuucgug caaaacggau ccaccuucgu gaagcaccug	3120
ucccuuggau cagugcagcu gugcggcgug ggaaaguucc cuucccuucc cauccugucc	3180
ccugcgcuga uggacgugcc guaccggcuc aacgagauc ccaaggaaaa ggaacagugc	3240
ugcgugucac uggcagccgg ccuuccgcau uuucgagcg gaauuuucag auguuggggc	3300
agagacaccu ucauugcgcu gcgcgguauc cugcugauua ccggcagaua cguggaagcc	3360

ULPI_041_01WO_SeqList_ST25.txt

agaaacauua uccuggcguu ugcugguacu cugcggcacg gacugauucc uaaccuguug	3420
ggagagggga ucuacgcccg guacaauugc agagaugccg ugugguggug gcugcagugc	3480
auccaggacu acugcaagau ggugcccaac ggacuugaca uucugaagug cccggugucg	3540
cgcauguacc ccaccgacga cucugcgccc cugccggccg guaccugga ucagccgcug	3600
uucgaaguga uccaggaagc caugcaaaag cacaugcagg gaauucaguu cagggaacgg	3660
aacgcaggcc cgcaaaucga ccggaacaug aaggacgaag gcuucaacau uaccgccggc	3720
guggacgagg aaaccggcuu cguguacggc ggcaaccggu ucaauugcgg uacuuggaug	3780
gacaagaugg gagaaagcga ccgcgccagg aaucggggca uuccugccac cccgcgggac	3840
gguagcgcgg uggagaucgu gggccuuucg aaguccgcgg uccgcuggcu ucuggagcug	3900
ucuaagaaaa acaucuuccc uuaccacgag gucaccguga aacgccacgg aaaggccauc	3960
aagguguccu acgacgaug gaacaggaag auccaggaca acuuugagaa gcuguuucac	4020
gugucggaag auccguccga ccugaacgaa aagcacccca accuugucca uaagcgcggg	4080
aucuacaaag auucguauug ugcauccucc ccuuggugcg acuaccaacu ccggccgaac	4140
uucaccaucg caaugguggu ggccccggag cuguucacca cugaaaaggc cuggaaggcc	4200
cuggaaaucg ccgaaaagaa gcugcuggga ccgcugggga ugaaaacccu ggaccccgau	4260
gauauuggugu acugcgggau cuacgacaac gcccuggaua acgacaacua caaccuggcc	4320
aagggcuuca acuaccaua ggguccggag uggcuguggc caaucggaua cuuccugagg	4380
gccaagcugu acuucucccg cuugaugggc cccgaaacua ccgcaaagac uaucgugcuc	4440
gugaagaacg uccugucccg gcacuacgug caucuggaac ggucgccgug gaaaggccug	4500
ccagaguuga ccaacgagaa ugcccagauu ugcccguucu caugcgaaac ccaagccugg	4560
agcauugcca cuauucugga aaccucucac gaccuguag	4599

<210> 30

<211> 4599

<212> RNA

<213> Artificial Sequence

ULPI_041_01W0_SeqList_ST25.txt

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 30

augggacacu cgaacagau cagaauccug uugcucaacg aaauggagaa gcucgaaaag	60
accuguuucc ggcuggagca ggguuacgag cugcaguucc gccugggucc gacucuacaa	120
gggaaagcag ugacggucua caccaacuac ccguuucccg gagaaaccuu caaccgggag	180
aaguuccggu cccuggacug ggagaacccg accgaacgag aggaugacuc agauaaguac	240
ugcaagcuga acuugcaaca guccggguca uuccaguacu acuuucugca aggggaacgag	300
aagucaggag ggggcuacau cguggucgac ccgauacugc gcgugggagc cgacaaucac	360
guccugccgc uggacugcgu gaccugcaa accuuccucg ccaaauugucu ggggccguuc	420
gaugaguggg aaagccgccu cagagucgca aaggaguccg gauacaacau gauucacuuc	480
acuccgcugc aaacacucgg ucugucccg ucgugcuauu cucuggcgaa ccagcuggag	540
cuuaaucccg acuuuucccg cccgaaccgg aaguacaccu ggaacgacgu cgggcagcuu	600
guggaaaagc ugaagaagga guggaacgug aucugcauca cugauguggu guacaaccac	660
acggcugcca acucgaagug gauccaagag caucccgagu gcgcguacaa cuuggugaac	720
agcccgauc ugaagccugc uugggugcug gacagagccc ucuggagauu cagcugcgac	780
guggccgagg ggaaguacaa agaaaagggg auuccggccu ugauugagaa cgaccaucac	840
augaacucaa uccgcaagau cauuugggag gauaucuucc cuaagcuuaa guugugggag	900
uucuuccaag uggacguaaa caaggcagug gagcaguuca gaaggcugcu gacucaagaa	960
aacagacgcg ucacuaaguc cgacccgaac cagcaccuua ccaucaucca agaccggag	1020
uaccgccggu ucggcugcac cgucgauaug aacauagccc ugaccacuuu caucucccau	1080
gacaaggggc cggcggaau cgaggaauugc uguaacuggu uucauaagcg gauggaggaa	1140
cugaauuccg agaagcaccg gcugauuaac uaccaccagg aacaggcagu gaacugccug	1200
cugggcaacg uguucuacga acggcuggcu ggacacggac cgaagcuggg ucccgugacu	1260
agaaagcauc cgcucgucac ucgcuacuuc accuucccg uugaggagau ugacuucucc	1320

ULPI_041_01WO_SeqList_ST25.txt

auggaagaau ccaugaucca cuugccgaac aaggcuugcu uccugauggc gcacaacgga	1380
ugggucaugg gcgacgaccc gcugaggaau uucgcggagc cggguucgga aguguaccug	1440
agaagggaac ucauuugcug gggagacucc gucaagcugc gcuaugggaa caagcccag	1500
gauugccccc accugugggc ucacaugaag aaguacacgg aaauccaccgc cacguacuuc	1560
caggagaucc ggcuggacaa uugccacucc acccccucc auguggccga guacaugcug	1620
gaugcagcgc gcaaucugca gccgaaucug uacgugguug cagagcuguu cacugggucc	1680
gaggaccucg acaacguguu cgugacuaga uuggggauuc ccuccucau ccgggaagcc	1740
augucggccu acaacuccca ugaggagggg aggcuggugu acagauacgg cggcgaaccc	1800
gugggaagcu ucgugcagcc gugccuccgg ccgcugaugc cggccaucgc ccacgcccug	1860
uucauggaua ucacucacga caacgaugc ccgaucgugc aucgcucggc cuacgacgca	1920
cugccgucca ccacuaucgu guccauggca ugcugcgccu ccgggagcac ccgcggauac	1980
gaugagcucg ugccgcacca gauuagcug guguccgaag aacgcuucua uaccaagugg	2040
aaccccgaag cccugccguc caauaccggg gaagugaacu uccaguccgg uauuauccg	2100
gcucgcugug cgauucgaa acuccaccaa gagcucggug ccaagggguu cauucagguc	2160
uacguggauc aggucgacga ggauuuugug gcagugacca ggcacucacc uagcauccac	2220
caauccgugg ucgccguguc acgcacugcg uuucggaacc ccaagaccuc guucuacucg	2280
aaagaagugc cgcagaugug uaucccggga aagaucgaag aggucgugcu ggaagcacgg	2340
accauugaga ggaacaccaa gccuuaccgg aaagacgaga acucuaucua cgggaccccg	2400
gauaucacug uggagauucg cgaacacauc cagcugaacg aaucaaagau cgucaagcag	2460
gcuggagugg ccaccaaggg acccaacgag uacauccagg agaucgaauu ugaaaaccuc	2520
uccccggggg ccgugaucua cuuccgggug ucccuggacc cccacgcca aguggccgug	2580
ggaauucuca gaaaccaccu gaccagauuc ucacccacu uuaagucggg uucccuggcc	2640
guggacaacg ccgauccgau ccucaagauc ccguucgcu cgcuggccuc ccgccucacu	2700
cucgcggaac ugaaccagau ccuguaccgc ugugaucag aggaaaagga ggacggcggc	2760

ULPI_041_01WO_SeqList_ST25.txt

ggcuguuacg auauuccgaa uuggucggcu uugaaauacg cgggacuuca ggggcugaug	2820
ucuguccugg cggaaauccg gccgaagaac gaccuggggc acccguucug caacaacuug	2880
cggagcggag auuggaugau cgacuacguc agcaacagac ugaucagccg gagcggcacu	2940
aucgccgagg ucggaaagug guugcaggcc auguucuucu accugaagca gaucccccga	3000
uaccucaucc cguguuacuu cgacgccauc cugaucgggg ccuacaccac ccuccuggac	3060
accgccugga agcagaugag caguuuugug caaaacgggu ccaccuucgu gaagcaccuu	3120
ucccuggggcu cagugcagcu gugcggcgug ggaaaguucc ccucgcugcc cauucugagc	3180
ccggcccuga uggacguccc cuaccggcug aacgagauca ccaaggagaa ggagcagugc	3240
uguguguccc uggcugccgg gcugccgcac uuuucguccg gcaucuuccg gugcuggggg	3300
cgggauaccu ucauugcccu gcggggaaucc cuucuauca ccggucgcu uguggaggcu	3360
cggaacauua uucuggcguu cgccggaacc cugagacacg ggcu gauucc gaaccucuug	3420
ggggaaggga ucuacgccc cuacaacugu cgggacgccg ugugguggug gcuccagugc	3480
auucaggacu auugcaagau ggugcccaac gggcuggaca uccugaagug cccggugucg	3540
aggauguauc cgaccgacga cagcgcaccc cugccggccg ggaccucga ccaaccucg	3600
uucgaaguca uucaagaggc uaugcagaag ccaugcagg guauucaguu ccgggaacgg	3660
aacgcggggc cccagauuga uaggaacaug aaagacgagg guuucaacau cacugccggc	3720
guggacgaag aaaccgguu uguguacgga ggaaaccggu ucaacugcgg uaccuggaug	3780
gauaagaugg gagaauca gaugcgcgcg aacagaggga ucccggcaac cccgcgggac	3840
ggaucggcug uggaaaauugu gggacugagc aagagcgccg ugcgguggcu gcuggaacug	3900
agcaaaaaga acaucuucc cuaccacgaa gugaccguga agcggcacgg aaaggccauc	3960
aaagucuccu acgaugaug gaauaggaag auccaggaua acuucgagaa gcuuuuucac	4020
guguccgagg auccuccga ucugaacgag aagcauccga aucucgugca uaagaggggg	4080
aucuacaagg acuccuacgg agcguccucc ccuuggugcg acuaucagcu gcggccuaac	4140
uucaccauug ccauggucgu ggccccggag cucuuuacaa ccgagaaggc cuggaaggcc	4200

ULPI_041_01W0_SeqList_ST25.txt

cucgaaaauug ccgaaaagaa gcugcuggggg cccuugggga ugaaaacccu ggauccggac	4260
gacauggugu acugcggaau cuacgacaac gcccuggaca acgacaacua caaucuggcg	4320
aagggcuuca auuaccacca ggggccggaa uggcucuggc cgauugggua cuuccugcg	4380
gccaagcugu acuucucacg gcugauggga ccggagacua ccgccaagac caucguccuc	4440
gugaagaacg ugcugucg cgacucgug caccuggaga ggagccccug gaaggggcuu	4500
cccgagcuga cgaacgagaa cgcgagacu ugucccuuc ccugcgaaac ccaagccugg	4560
uccauugcca cuauacugga aaccuuauau gaccuguag	4599

<210> 31
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 31	
augggccaca gcaagcagau ccggaucuug cugcugaacg agauggagaa gcuggagaag	60
accuguuua ggcuggagca gggcuacgag cugcaguucc gguuggggccc caccuugcag	120
ggcaaggccg ugaccgugua caccaacuac cccuuccccc gcgaaaccuu caacaggagg	180
aaguuccggu cccuggacug ggagaacccc accgagaggg aggacgacuc cgacaaguac	240
ugcaagcuga accugcagca guccggcuucc uuccaguacu acuuccugca gggcaacgag	300
aaaagcggcg gcggcuacau cgugguggac cccaucuugc ggguggggcg cgacaaccac	360
gugcugcccu uggacugcgu gaccugcag accuucuugg ccaagugcuu gggcccuuc	420
gacgaguggg agagcaggcu gaggguggcc aaggaguccg gcuacaacau gauccacuuc	480
accccuugc agaccuggg ccuguccagg uccugcuacu cccuggccaa ccaguuggag	540
uugaacccc acuucuccag gcccaacagg aaguacaccu ggaacgacgu gggccagcug	600
guggagaagu ugaagaagga guggaacgug aucugcauca ccgacuggu guacaaccac	660
accgccgcca acagcaagug gauccaggag caccgccagu gcgccuacaa ccuggugaac	720

ULPI_041_01W0_SeqList_ST25.txt

uccccccacu ugaagcccg	cuggguguug gacagggccc	uguggcgguu cuccugcgac	780
guggccgagg gcaaguacaa	ggagaagggc aucccgccu	ugaucgagaa cgaccaccac	840
augaacucca uccggaagau	caucugggag gacaucuucc	ccaagcugaa gcugugggag	900
uucuuccagg uggacgugaa	caaggccgug gagcaguuca	ggaggcugcu gacccaggag	960
aacaggcggg ugaccaaguc	cgaccccaac cagcaccuga	ccaucaucca ggaccccgag	1020
uacaggcggu ucggcugcac	cguggacaug aacaucgccc	ugaccacuu caucggccac	1080
gacaagggcc ccgccccaau	cgaggagugc ugcaacuggu	uccacaagag gauggaggag	1140
uugaacuccg agaagcaccg	gcugaucaac uaccaccagg	agcaggccgu gaacugccug	1200
uugggcaacg uguucuacga	gcggcuggcc ggccacggcc	ccaagcuggg ccccgugacc	1260
aggaagcacc ccuuggugac	cagguacuuc accuucccu	ucgaggagau cgacuucucc	1320
auggaggagu ccaugaucca	ccugcccaac aaggccugcu	uccugauggc ccacaacggc	1380
ugggugaugg gcgacgaccc	ccugcggaac uucgccgagc	ccggcuuccga gguguaccug	1440
aggagggagc ugauugcug	gggcgacagc gugaaguugc	gguacggcaa caagcccgag	1500
gacugcccu accugugggc	ccacaugaag aaguacaccg	agaucaccgc caccuacuuc	1560
cagggcgugc ggcuggacaa	cugccacucc acccccugc	acguggccga guacauguug	1620
gacgccgcca ggaacuugca	gcccaacuug uacguggugg	ccgagcuguu caccggcagc	1680
gaggaccugg acaacuguu	cgugaccagg cugggcauca	gcuccuugau cagggaggcc	1740
augagcgccu acaacagcca	cgaggagggc agguuggugu	accgguacgg cggcgagccc	1800
gugggcuccu ucgugcagcc	cugcuugagg cccuugaugc	ccgccaucgc ccacggccug	1860
uucauggaca ucacccacga	caacgagugc cccaucgugc	acagguccgc cuacgacgcc	1920
cugcccagca ccaccaucgu	guccauggcc ugcugcgcca	gcggcagcac caggggcuac	1980
gacgaguugg ugccccacca	gaucuccgug guguccgagg	agcgguucua caccaagugg	2040
aaccccgagg ccuugcccuc	caacaccggc gaggugaacu	uccagagcgg caucaucgcc	2100
gccaggugcg ccaucagcaa	gcugcaccag gagcugggcg	ccaagggcuu cauccaggug	2160

ULPI_041_01W0_SeqList_ST25.txt

uacguggacc	agguggacga	ggacaucgug	gccgugacca	ggcacucucc	cagcauccac	2220
caguccgugg	uggccguguc	caggaccgcc	uucaggaacc	ccaagaccuc	cuucuacagc	2280
aaggaggugc	cccagaugug	caucuccggc	aagaucgagg	agguggugcu	ggaggccagg	2340
accaucgaga	ggaacaccaa	gcccuaacagg	aaggacgaga	acuccaucaa	cggcaccccc	2400
gacaucaccg	uggagaucag	ggagcacauc	cagcugaacg	agagcaagau	cgugaagcag	2460
gccggcgugg	ccaccaaggg	ccccaacgag	uacauccagg	agaucgaguu	cgagaacuug	2520
ucucccgcca	gcgugaucu	cuucagggug	agccuggacc	cccacgcccc	gguggccgug	2580
ggcauccugc	ggaaccaccu	gaccagauuc	agccccacu	ucaaguccgg	cagccuggcc	2640
guggacaacg	ccgacccccau	cuugaagauc	ccuucgccu	cccuggccuc	cagguugacc	2700
uuggccgagc	ugaaccagau	ccuguaccgg	ugcgaguccg	aggagaagga	ggacggcggc	2760
ggcugcuacg	acaucuccaa	cugguccgcc	cugaaguacg	ccggccugca	gggcuugaug	2820
uccguguugg	ccgagaucag	gcccagaac	gacuugggcc	acccuucug	caacaacuug	2880
agguccggcg	acuggaugau	cgacuacgug	agcaaccggc	ugaucucccg	guccggcacc	2940
aucgccgagg	ugggcaagug	guugcaggcc	auguucuucu	accugaagca	gaucucccg	3000
uaccugaucc	ccugcuacuu	cgacgccauc	uugaucggcg	ccuacaccac	ccugcuggac	3060
accgccugga	agcagauguc	cagcuucgug	cagaacggcu	ccaccuucgu	gaagcaccug	3120
uccuugggcu	ccgugcagcu	gugcggcgug	ggcaaguucc	ccuccugcc	cauccugucc	3180
cccgcccuga	uggacgugcc	cuacagguug	aacgagauca	ccaaggagaa	ggagcagugc	3240
ugcguguccc	uggccgccgg	cuugccccac	uucuccuccg	gcaucuuccg	gugcuggggc	3300
agggacaccu	ucaucgcccc	gaggggcauc	cugcugauca	ccggccggua	cguggaggcc	3360
aggaacauca	ucuuggccuu	cgccggcacc	cugaggcacg	gccugauccc	caaccugcug	3420
ggcgagggca	ucuacgccag	guacaacugc	cgggacgccg	ugugguggug	gcugcagugc	3480
auccaggacu	acugcaagau	ggugcccaac	ggccuggaca	uccugaagug	ccccgugucc	3540
aggauguacc	ccaccgacga	cuccgcccc	uugcccgccg	gcaccucgga	ccagcccuug	3600

ULPI_041_01W0_SeqList_ST25.txt

uucgagguga uccaggaggc caugcagaag cacaugcagg gcauccaguu ccgggagagg	3660
aacgccggcc cccagaucga ccggaacaug aaggacgagg gcuucaacau caccgccggc	3720
guggacgagg aaaccggcuu cguguacggc ggcaaccggu ucaacugcgg caccuggaug	3780
gacaagaugg gcgagagcga cagggccagg aacaggggca uccccgccac ccccagggac	3840
ggcuccgccg uggagaucgu gggccugagc aaguccgccg ugcggugguu gcuggaguug	3900
uccaagaaga acaucuucc cuaccacgag gugaccguga agaggcacgg caaggccauc	3960
aagguguccu acgacgagug gaacaggaag auccaggaca acuucgagaa gcuguuccac	4020
guguccgagg accccuccga cuugaacgag aagcacccca accuggugca caagcggggc	4080
aucuacaagg acagcuacgg cgccuccagc cccuggugcg acuaccagcu gaggcccaac	4140
uucaccaucg ccaugguggu ggcccccgag cuguucacca ccgagaaggc cuggaaggcc	4200
uuggagaucg ccgagaagaa guugcugggc cccugggca ugaagaccuu ggaccccgac	4260
gacauggugu acugcggcau cuacgacaac gccuuggaca acgacaacua caaccuggcc	4320
aagggcuuca acuaccacca gggccccgag uggcuguggc ccaucggcua cuuccugcgg	4380
gccaauguu acuucuccag guugaugggc cccgaaacca ccgccaagac caucguguug	4440
gugaagaacg ugcugucccg gcacuacgug caccuggaga ggucggcug gaagggccug	4500
cccgagcuga ccaacgagaa cgcccaguac ugccccuua gcugcgaaac ccaggccugg	4560
uccaucgcca ccauccugga aaccuguac gacuuguag	4599

<210> 32

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 32

augggacaua gcaagcagau ccggaucuug cugcugaaug aaauggaaaa gcuggaaaag	60
---	----

accuguuua gacuggaaca gggauacgaa cugcaguucc gguugggacc uaccuugcag	120
--	-----

ULPI_041_01WO_SeqList_ST25.txt

ggaaaggcug ugaccgugua caccaauuac ccuuucccug gagaaaccuu caauagagaa	180
aaguucagau cucuagauug ggaaaauccu accgaaagag aagaugauuc ugauaaguac	240
ugcaagcuga aucugcagca gucuggaucu uuccaguacu acuuccugca gggaaaugaa	300
aagagcggag gaggauacau cgugguggau ccuaucuugc ggguggggagc ugauaaucan	360
gugcugccuu uggauugcgu gaccugcag accuucuugg cuaagugcuu gggaccuuuc	420
gaugaauggg aaagcagacu gagaguggcu aaggaaucug gauacaauau gauccauuuc	480
accccuuugc agaccuggg acugucuaga ucuugcuacu cucuggcuua ucaguuggaa	540
uugaauccug auuucucuag accuaauaga aaguacaccu ggaaugaugu gggacagcug	600
guggaaaagu ugaagaagga auggaaugug aucugcauca ccgauguggu guacaaucau	660
accgcugcua auagcaagug gauccaggaa cauccugaau gcgcuuacaa ucuggugaau	720
ucuccucauu ugaagccugc uuggguguug gauagagcuc uguggcgguu cucuugcgau	780
guggcugaag gaaaguacaa ggaaaaggga auccugcuu ugaucgaaaa ugaucaucau	840
augaauucua uccggaagau caucugggaa gauaucuucc cuaagcugaa gcugugggaa	900
uucuuccagg uggauuguaa uaaggcugug gaacaguuca gaagacugcu gaccaggaa	960
aaauagacggg ugaccaaguc ugauccuaau cagcaucuga ccaucaucca ggauccugaa	1020
uacagacggg ucggaugcac cguggauaug aaauaucgcuc ugaccaccuu cauccucau	1080
gauaaggggac cugcugcuau cgaagaugc ugcaauuggu uccaauagag aauggaagaa	1140
uugaauucug aaaagcaucg gcugaucaau uaccaucagg aacaggcugu gaauugccug	1200
uugggaaaug uguucuacga acggcuggcu ggacauggac cuaagcuggg accugugacc	1260
agaaagcauc cuuuggugac cagauacuuc accuucccuu ucgaagaaau cgauuucucu	1320
auggaagaau cuaugaucca ucugccuaau aaggcuugcu uccugauggc ucauaaugga	1380
ugggugaugg gagaugauc ucugcggaau uucgcugaac cuggaucuga aguguaccug	1440
agaagagaac ugaucugcug gggagauagc gugaaguugc gguacggaaa uaagccugaa	1500
gauugcccuu accugugggc ucauaugaag aaguacaccg aaauaccgc uaccuacuuc	1560

ULPI_041_01WO_SeqList_ST25.txt

cagggagugc ggcuggauaa uugccauucu accccucugc auguggcuga auacauguug	1620
gaugcugcua gaaauuugca gccuaauuug uacguggugg cugaacuguu caccggaagc	1680
gaagaucugg auaauguguu cgugaccaga cugggaauca gcucuugau cagagaagcu	1740
augagcgcuu acaauagcca ugaagaagga agauuggugu accgguacgg aggagaaccu	1800
gugggaucuu ucgugcagcc uugcuugagg ccuuugaugc cugcuaucgc ucaugcucug	1860
uucauggaua ucacccauga uaaugaaugc ccuaucgugc auagaucugc uuacgaugcu	1920
cugccuagca ccaccaucgu gucuauaggc ugcugcgcua gcggaagcac cagaggauac	1980
gaugaauugg ugccucauca gaucucugug gugucugaag aacgguucua caccaagugg	2040
aauccugaag cuuugccuuc uaauaccgga gaagugaauu uccagagcgg aaucaucgcu	2100
gcuagaugcg cuaucagcaa gcugcaucag gaacugggag cuaagggauu cauccaggug	2160
uacguggauc agguggauga agauaucgug gcugugacca gacauucc uagcauccau	2220
cagucugugg uggcuguguc uagaaccgcu uucagaaauc cuaagaccuc uuucuacagc	2280
aaggaagugc cucagaugug caucccugga aagaucgaag aaguggugcu ggaagcuaga	2340
accaucgaaa gaaauaccaa gccuuacaga aaggaugaaa auucuaucua uggaaccccu	2400
gauaucaccg uggaauacag agaacauauc cagcugaaug aaagcaagau cgugaagcag	2460
gcuggagugg cuaccaaggg accuaaugaa uacauccagg aaucgaauu cgaaaauuug	2520
ucuccuggaa gcgugaucuu cuucagagug agccuggauc cucaugcuca gguggcugug	2580
ggaauccugc ggaaucaucu gaccaguuc agcccucuu ucaagucugg aagccuggcu	2640
guggauaaug cugauccuau cuugaagauc ccuuucgcuu cucuggcuuc uagauugacc	2700
uuggcugaac ugaauacgau ccuguaccgg ugcgaucug aagaaaagga agauggagga	2760
ggaugcuacg auaucccuua uuggucugcu cugaaguacg cuggacugca gggauugaug	2820
ucuguguugg cugaaaucag accuaagaau gauuugggac auccuuucug caauauuug	2880
agaucuggag auuggaugau cgauuacgug agcaaucggc ugaucucucg gucuggaacc	2940
aucgcugaag ugaggaaagug guugcaggcu auguucuucu accugaagca gaucuccugg	3000

ULPI_041_01WO_SeqList_ST25.txt

uaccugaucc cuugcuacuu cgaugcuau c ugaucggag cuuacaccac ccugcuggau	3060
accgcuugga agcagauguc uagcuucgug cagaau ggau cuaccuucgu gaagcaucug	3120
ucuuugggau cugugcagcu gugcggagug ggaaaguucc cuucucugcc uauccugucu	3180
ccugcucuga uggauugucc uuacagauug aaugaaa uca ccaaggaaaa ggaacagugc	3240
ugcgugucuc uggcugcugg auugccucau uucucuucug gaau cuuccg gugcugggga	3300
agagauaccu ucaucgcucu gagaggaau cugcugauca ccggacggua cguggaagcu	3360
agaaauauca ucuuggcuuu cgcuggaacc cugagacaug gacugauccc uauucugcug	3420
ggagaaggaa ucuacgcuag auacaauugc cgggaugcug ugugguggug gcugcagugc	3480
auccaggauu acugcaagau ggugccuaau ggacuggaua uccugaagug cccugugucu	3540
agaau guacc cuaccgauga uucugcuccu uugccugcug gaaccugga ucagccuuug	3600
uucgaaguga uccaggaagc uaugcagaag cauau gcagg gaauccagu cccgggaaaga	3660
aaugcuggac cucagaucga ucggaauaug aaggau gaag gauucaauau caccgcugga	3720
guggaugaag aaaccggauu cguguacgga ggaaau cgg ucaauugcgg aaccuggaug	3780
gauaagaugg gagaaagcga uagagcuaga aaauagaggaa ucccugcuac cccuagagau	3840
ggauucugcug uggaaau cgu gggacugagc aagucugcug ugcggugguu gcuggaauug	3900
ucuaagaaga auau cuucc uuaccauga gugaccguga agagacaugg aaaggcuau c	3960
aaggugucuu acgaugaau gaauagaa auccaggau auuucgaaaa gcuguuccau	4020
gugucugaag auccuucuga uuugaau ga aagcaucca aucuggugca uaagcgggga	4080
aucuacaagg auagcuacgg agcuucuagc ccuuggugcg auuaccagcu gaggccuaau	4140
uucaccaucg cuau gguggu ggcuccugaa cuguucacca ccgaaaaggc uuggaaggcu	4200
uuggaaaucg cugaaaagaa guugcuggga ccucugggaa ugaagaccuu ggauccugau	4260
gauau ggugu acugcggaau cuacgaaua gcuuuggaua augauaa uua caaucuggcu	4320
aagggaauca auuaccau cgggaccugaa uggcuguggc cuaucggaua cuuccugcgg	4380
gcuaaguugu acuucucuag auugauggga ccugaaacca ccgcuaagac caucguguug	4440

ULPI_041_01W0_SeqList_ST25.txt

gugaagaaug ugcugucucg gcauuacgug caucuggaaa gaucuccuug gaagggacug 4500
 ccugaacuga ccaaugaaaa ugcucaguac ugcccuuua gcugcgaaac ccaggcuugg 4560
 ucuaucgcu ccauccugga aaccucguac gauuuguag 4599

<210> 33
 <211> 158
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Modified Xenopus beta-globin 3'UTR

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (157)..(157)
 <223> n is unlocked nucleomonomer UNA-A

<400> 33
 cnagugacug acuaggaucu gguuaccacu aaaccagccu caagaacacc cgauggagu 60
 cucuaagcua cauaauacca acuuacacuu acaaaauguu gucccccaaa auguagccau 120
 ucguaucugc uccuaauaaa aagaaaguuu cuucacnu 158

<210> 34
 <211> 158
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Modified Xenopus beta-globin 3'UTR

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> n is unlocked nucleomonomer UNA-U

ULPI_041_01W0_SeqList_ST25.txt

```

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is unlocked nucleomonomer UNA-A

<220>
<221> misc_feature
<222> (4)..(4)
<223> n is unlocked nucleomonomer UNA-G

<220>
<221> misc_feature
<222> (155)..(155)
<223> n is unlocked nucleomonomer UNA-A

<220>
<221> misc_feature
<222> (156)..(156)
<223> n is unlocked nucleomonomer UNA-C

<220>
<221> misc_feature
<222> (157)..(157)
<223> n is unlocked nucleomonomer UNA-A

<400> 34
cnnnugacug acuaggaucu gguuaccacu aaaccagccu caagaacacc cgaauggagu      60

cucuaagcua cauaauacca acuuacacuu acaaaauguu gucccccaaa auguagccau      120

ucguaucugc uccuaauaaa aagaaaguuu cuucnnnu      158


<210> 35
<211> 158
<212> RNA
<213> Artificial Sequence

<220>
<223> Modified Xenopus beta-globin 3'UTR

<220>
<221> misc_feature
<222> (2)..(2)
<223> n is unlocked nucleomonomer UNA-U

<220>

```

<221> misc_feature
 <222> (8)..(8)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (20)..(20)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (26)..(26)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (32)..(32)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (38)..(38)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (44)..(44)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (50)..(50)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (56)..(56)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (62)..(62)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (68)..(68)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (74)..(74)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (80)..(80)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (86)..(86)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (92)..(92)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (98)..(98)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (104)..(104)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (110)..(110)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (116)..(116)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature

<222> (122)..(122)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (128)..(128)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (134)..(134)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (140)..(140)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (146)..(146)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (152)..(152)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (157)..(157)
 <223> n is unlocked nucleomonomer UNA-A

<400> 35
 cnaguganug acunggauch gguuancacu anaccagncu caanaacacn cgaaungagu 60
 cncuaagnua caunauacn acuuanaacu anaaaaunuu guncaccaan auguanccau 120
 unguaucngc uccnaauaan aagaanguuu cnucacnu 158

<210> 36
 <211> 158
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Modified Xenopus beta-globin 3'UTR

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (8)..(8)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (20)..(20)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (21)..(21)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (26)..(27)
 <223> n is unlocked nucleomonomer UNA-C

<220>

<221> misc_feature
 <222> (32)..(33)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (38)..(39)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (44)..(44)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (45)..(45)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (50)..(51)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (56)..(57)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (62)..(62)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (63)..(63)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (68)..(68)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (69)..(69)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (74)..(75)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (80)..(81)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (86)..(86)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (87)..(87)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (92)..(92)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (93)..(93)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (98)..(98)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (99)..(99)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (104)..(105)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature

<222> (110)..(111)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (116)..(116)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (117)..(117)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (122)..(122)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (123)..(123)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (128)..(128)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (129)..(129)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (134)..(134)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (135)..(135)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (140)..(141)
 <223> n is unlocked nucleomonomer UNA-A

ULPI_041_01WO_SeqList_ST25.txt

```

<220>
<221> misc_feature
<222> (146)..(146)
<223> n is unlocked nucleomonomer UNA-A

<220>
<221> misc_feature
<222> (147)..(147)
<223> n is unlocked nucleomonomer UNA-G

<220>
<221> misc_feature
<222> (152)..(153)
<223> n is unlocked nucleomonomer UNA-U

<220>
<221> misc_feature
<222> (157)..(157)
<223> n is unlocked nucleomonomer UNA-A

<400> 36
cnnguganng acunngaucn nguuanacu annccaggnu caannacacn ngaaunnagu      60

cnnuaagnna caunnuacn ncuuanncuu annaaaunnu gucnnccaan nuguanncau      120

unnuaucnnc uccnnauaan nagaannuuu cnncacnu      158


<210> 37
<211> 158
<212> RNA
<213> Artificial Sequence

<220>
<223> Modified Xenopus beta-globin 3'UTR

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is unlocked nucleomonomer UNA-C

<220>
<221> misc_feature
<222> (2)..(2)
<223> n is unlocked nucleomonomer UNA-U

<220>

```

<221> misc_feature
 <222> (3)..(3)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (4)..(4)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (5)..(5)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (7)..(7)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (8)..(8)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (10)..(10)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (11)..(11)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (13)..(13)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (15)..(16)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (17)..(17)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (18)..(18)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (19)..(19)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (20)..(20)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (23)..(24)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature

<222> (25)..(25)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (26)..(27)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (28)..(28)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (29)..(29)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (30)..(30)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (31)..(33)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (34)..(35)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (36)..(36)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (37)..(37)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (38)..(39)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (40)..(40)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (41)..(41)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (42)..(43)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (44)..(44)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (45)..(46)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (47)..(47)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (48)..(48)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (49)..(51)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (52)..(52)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (53)..(54)

<223> n is unlocked nucleomonomer UNA-A

<220>

<221> misc_feature

<222> (55)..(55)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature

<222> (56)..(57)

<223> n is unlocked nucleomonomer UNA-G

<220>

<221> misc_feature

<222> (58)..(58)

<223> n is unlocked nucleomonomer UNA-A

<220>

<221> misc_feature

<222> (59)..(59)

<223> n is unlocked nucleomonomer UNA-G

<220>

<221> misc_feature

<222> (60)..(60)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature

<222> (61)..(61)

<223> n is unlocked nucleomonomer UNA-C

<220>

<221> misc_feature

<222> (62)..(62)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature

<222> (63)..(63)

<223> n is unlocked nucleomonomer UNA-C

<220>

<221> misc_feature

<222> (64)..(64)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature
 <222> (65)..(66)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (67)..(67)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (68)..(68)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (69)..(69)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (70)..(70)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (71)..(71)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (72)..(72)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (73)..(73)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (74)..(75)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (76)..(76)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (77)..(77)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (78)..(79)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (80)..(81)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (82)..(82)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (83)..(84)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (85)..(85)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (86)..(86)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (87)..(87)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (88)..(88)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature

<222> (89)..(90)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (91)..(91)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (92)..(92)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (93)..(96)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (97)..(97)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (98)..(98)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (99)..(100)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (101)..(101)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (102)..(102)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (103)..(107)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (108)..(111)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (112)..(112)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (113)..(113)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (114)..(114)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (115)..(115)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (116)..(116)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (117)..(118)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (119)..(119)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (120)..(121)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (122)..(122)

<223> n is unlocked nucleomonomer UNA-C

<220>

<221> misc_feature

<222> (123)..(123)

<223> n is unlocked nucleomonomer UNA-G

<220>

<221> misc_feature

<222> (124)..(124)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature

<222> (125)..(125)

<223> n is unlocked nucleomonomer UNA-A

<220>

<221> misc_feature

<222> (126)..(126)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature

<222> (127)..(127)

<223> n is unlocked nucleomonomer UNA-C

<220>

<221> misc_feature

<222> (128)..(128)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature

<222> (129)..(129)

<223> n is unlocked nucleomonomer UNA-G

<220>

<221> misc_feature

<222> (130)..(130)

<223> n is unlocked nucleomonomer UNA-C

<220>

<221> misc_feature

<222> (131)..(131)

<223> n is unlocked nucleomonomer UNA-U

<220>

<221> misc_feature
 <222> (132)..(133)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (134)..(134)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (135)..(136)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (137)..(137)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (138)..(142)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (143)..(143)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (144)..(146)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (147)..(147)
 <223> n is unlocked nucleomonomer UNA-G

<220>
 <221> misc_feature
 <222> (148)..(149)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (150)..(150)
 <223> n is unlocked nucleomonomer UNA-U

ULPI_041_01W0_SeqList_ST25.txt

<220>
 <221> misc_feature
 <222> (151)..(151)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (152)..(153)
 <223> n is unlocked nucleomonomer UNA-U

<220>
 <221> misc_feature
 <222> (154)..(154)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (155)..(155)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (156)..(156)
 <223> n is unlocked nucleomonomer UNA-C

<220>
 <221> misc_feature
 <222> (157)..(157)
 <223> n is unlocked nucleomonomer UNA-A

<220>
 <221> misc_feature
 <222> (158)..(158)
 <223> n is unlocked nucleomonomer UNA-U

<400> 37
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 60
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 120
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 158

<210> 38
 <211> 100
 <212> RNA
 <213> Unknown

ULPI_041_01W0_SeqList_ST25.txt

<220>

<223> Poly(A) 100 Tail

<400> 38

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 100

<210> 39

<211> 110

<212> RNA

<213> Unknown

<220>

<223> Poly(A) 110 Tail

<400> 39

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 110

<210> 40

<211> 9

<212> RNA

<213> Unknown

<220>

<223> Triple Stop Codon

<400> 40

auaagugaa 9

<210> 41

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 41

augggccaca gcaagcagau ccggauccug cugcugaacg agauggagaa gcuggagaag 60

accuguuucc ggcuggagca gggcuacgag cugcaguucc ggcuggggccc caccucgag 120

ULPI_041_01W0_SeqList_ST25.txt

ggcaaggccg ugaccgugua caccaacuac cccuuccccg gcgagaccuu caaccgggag	180
aaguuccgga gccuggacug ggagaacccc accgagcggg aggacgacag cgacaaguac	240
ugcaagcuga accugcagca gagcggcagc uuccaguacu acuuccugca gggcaacgag	300
aagagcggcg gcggcuacau cgugguggac cccauccugc ggguggggcg cgacaaccac	360
gugcugcccc uggacugcgu gaccugcag accuuccugg ccaagugccu gggcccuuc	420
gacgaguggg agagccggcu gcggguggcc aaggagagcg gcuacaacau gauccacuuc	480
acccccugc agaccuggg ccugagccgg agcugcuaca gccuggccaa ccagcuggag	540
cugaacccccg acuucagccg gcccacccg aaguacaccu ggaacgacgu gggccagcug	600
guggagaagc ugaagaagga guggaacgug aucugcauca ccgacguggu guacaaccac	660
accgccgcca acagcaagug gauccaggag caccgccagu gcgccuacaa ccuggugaac	720
agccccacc ugaagccgc cugggugcug gaccgggccc uguggcgguu cagcugcgac	780
guggccgagg gcaaguacaa ggagaagggc aucccgccc ugaucgagaa cgaccaccac	840
augaacagca uccggaagau caucugggag gacaucuucc ccaagcugaa gcugugggag	900
uucuuccagg uggacgugaa caaggccgug gagcaguucc ggcggcugcu gaccagggag	960
aaccggcggg ugaccaagag cgaccccaac cagcaccuga ccaucaucca ggaccccgag	1020
uaccggcggg ucggcugcac cguggacaug aacaucgccc ugaccaccuu caucaccac	1080
gacaagggcc ccgccccaau cgaggagugc ugcaacuggu uccacaagcg gauggaggag	1140
cugaacagcg agaagcaccg gcugaucaac uaccaccagg agcaggccgu gaacugccug	1200
cugggcaacg uguucuacga gcggcuggcc ggccacggcc ccaagcuggg ccccgugacc	1260
cggaagcacc cccuggugac ccgguacuuc accuucccu ucgaggagau cgacuucagc	1320
auggaggaga gcaugaucca ccugcccaac aaggccugcu uccugauggc ccacaacggc	1380
ugggugaugg gcgacgaccc ccugcggaa uucgccgagc ccggcagcga gguguaccug	1440
cggcgggagc ugauugcug gggcgacagc gugaagcugc gguacggcaa caagcccag	1500
gacugcccu accugugggc ccacaugaag aaguacaccg agaucaccgc caccuacuuc	1560

ULPI_041_01W0_SeqList_ST25.txt

cagggcgugc	ggcuggacaa	cugccacagc	acccccugc	acguggccga	guacaucug	1620
gacgccgccc	ggaaccugca	gcccaccug	uacguggugg	ccgagcuguu	caccggcagc	1680
gaggaccugg	acaacguguu	cgugacccgg	cugggcauca	gcagccugau	ccgggaggcc	1740
augagcgccu	acaacagcca	cgaggagggc	cggcuggugu	accgguacgg	cggcgagccc	1800
gugggcagcu	ucgugcagcc	cugccugcgg	ccccugaugc	ccgccaucgc	ccacgcccug	1860
uucauggaca	ucacccacga	caacgagugc	cccaucgugc	accggagcgc	cuacgacgcc	1920
cugcccagca	ccaccaucgu	gagcauggcc	ugcugcgcca	gcggcagcac	ccggggcuac	1980
gacgagcugg	ugccccacca	gaucagcgug	gugagcgagg	agcgguuucua	caccaagugg	2040
aaccccgagg	cccugcccag	caacaccggc	gaggugaacu	uccagagcgg	caucaucgcc	2100
gcccggugcg	ccaucagcaa	gcugcaccag	gagcuggggcg	ccaagggcuu	cauccaggug	2160
uacguggacc	agguggacga	ggacaucgug	gccgugaccc	ggcacagccc	cagcauccac	2220
cagagcgugg	uggccgugag	ccggaccgcc	uuccggaacc	ccaagaccag	cuucuacagc	2280
aaggaggugc	cccagaugug	cauccccggc	aagaucgagg	agguggugcu	ggaggcccgg	2340
accaucgagc	ggaacaccaa	gcccuaccgg	aaggacgaga	acagcaucua	cggcaccccc	2400
gacaucaccg	uggagaucgg	ggagcacauc	cagcugaacg	agagcaagau	cgugaagcag	2460
gccggcgugg	ccaccaaggg	ccccaacgag	uacauccagg	agaucgagu	cgagaaccug	2520
agccccggca	gcgugauc	cuuccgggug	agccuggacc	cccacgccc	gguggccgug	2580
ggcauccugc	ggaaccaccu	gaccaguu	agccccacu	ucaagagcgg	cagccuggcc	2640
guggacaacg	ccgaccccau	ccugaagauc	cccuucgcca	gccuggccag	ccggcugacc	2700
cuggccgagc	ugaaccagau	ccuguaccgg	ugcgagagcg	aggagaagga	ggacggcggc	2760
ggcugcuacg	acaucaccaa	cuggagcgcc	cugaaguacg	ccggccugca	gggcccugaug	2820
agcgugcugg	ccgagaucgg	gcccagaag	gaccuggggc	accccuucug	caacaaccug	2880
cggagcggcg	acuggaugau	cgacuacgug	agcaaccggc	ugaucagccg	gagcggcacc	2940
aucgccgagg	ugggcaagug	gcugcaggcc	auguucuucu	accugaagca	gaucacccgg	3000

ULPI_041_01W0_SeqList_ST25.txt

uaccugaucc ccugcuacuu cgacgccauc cugaucggcg ccuacaccac ccugcuggac	3060
accgccugga agcagaugag cagcuucgug cagaacggca gcaccuucgu gaagcaccug	3120
agccugggca gcgugcagcu gugcggcgug ggcaaguucc ccagccugcc cauccugagc	3180
cccgcccuga uggacgugcc cuaccggcug aacgagauca ccaaggagaa ggagcagugc	3240
ugcgugagcc uggccgccgg ccugccccac uucagcagcg gcaucuuccg gugcuggggc	3300
cgggacaccu ucaucgccc ugcggggcauc cugcugauca ccggccggua cguggaggcc	3360
cggaacauca uccuggccuu cgccggcacc cugcggcacg gccugauccc caaccugcug	3420
ggcgagggca ucuacgccc guacaacugc cgggacgccg ugugguggug gcugcagugc	3480
auccaggacu acugcaagau ggugcccaac ggccuggaca uccugaagug ccccgugagc	3540
cggauguacc ccaccgacga cagcgcccc cugcccgccg gcaccugga ccagccccug	3600
uucgagguga uccaggaggc caugcagaag cacaugcagg gcauccaguu ccgggagcgg	3660
aacgccggcc cccagaucga ccggaacaug aaggacgagg gcuucaacau caccgccggc	3720
guggacgagg agaccggcuu cguguacggc ggcaaccggu ucaacugcgg caccuggaug	3780
gacaagaugg gcgagagcga ccggggcccgg aaccggggca uccccgccac cccccgggac	3840
ggcagcgccg uggagaucgu gggccugagc aagagcgccg ugcgguggcu gcuggagcug	3900
agcaagaaga acaucuucc cuaccacgag gugaccguga agcggcacgg caaggccauc	3960
aaggugagcu acgacgagug gaaccggaag auccaggaca acuucgagaa gcuguuccac	4020
gugagcgagg accccagcga ccugaacgag aagcaccaca accuggugca caagcggggc	4080
aucuacaagg acagcuacgg cgccagcagc cccuggugcg acuaccagcu gcggcccaac	4140
uucaccaucg ccaugguggu ggcccccgag cuguucacca ccgagaaggc cuggaaggcc	4200
cuggagaucg ccgagaagaa gcugcugggc cccugggca ugaagaccu ggaccccgac	4260
gacauggugu acugcggcau cuacgacaac gcccuggaca acgacaacua caaccuggcc	4320
aagggcuuca acuaccacca gggccccgag uggcuguggc ccaucggcua cuuccugcgg	4380
gccaagcugu acuucagccg gcugaugggc cccgagacca ccgccaagac caucgugcug	4440

ULPI_041_01W0_SeqList_ST25.txt

gugaagaacg ugcugagccg gcacuacgug caccuggagc ggagccccug gaagggccug	4500
cccgagcuga ccaacgagaa cgcccaguac ugccccuua gcugcgagac ccaggccugg	4560
agcaucgcca ccauccugga gaccuguac gaccuguag	4599

<210> 42
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 42	
augggccaaua gcaagcagau ccggauccuc cugcucaaug agauggaaaa gcucgagaaa	60
accucuuuaa ggcuggagca aggcuaagaa cuccaguuaa ggcucgggcc cacucuccag	120
ggcaaggccg ugaccgucua caccaacuau cccuucccug gcgagaccuu caacaggagg	180
aaguuuucggu cccucgacug ggagaacccc accgagaggg aggacgacuc cgacaaguau	240
ugcaagcuga accuccaaca auccgggucc uuccaguacu acuuucugca aggcaacgaa	300
aagagcgggg gcggguauau uguggucgau cccauccucc gggugggggc ugacaaccac	360
guccuuuuuu uggauugcgu gacucugcag accuuccucg cuaaaugccu gggcccuuuc	420
gacgaguggg agagcaggcu gcgggugggc aaggaguccg gguauauau gauccacuuc	480
acccccugc agacucuggg ccucucccgg uccuguuaua gccuggccaa ccagcucgag	540
cucaaccucg auuucuccag gccuaacagg aaauacaccu ggaacgacgu gggccagcuc	600
gucgagaagc ucaagaaaga guggaacgug aucugcauca cugacuggu cuauaaccac	660
acugcugcua acagcaagug gauucaggag cccccgagu gcgccuacaa ccuggucaac	720
uuuuuuuau ucaagccugc cuggguccuc gauagggccc uguggcgguu uagcugcgac	780
guggccgagg gcaaguauaa ggagaaaggc auuccugcuc ugaucgagaa cgaccaucac	840
augaacagca uucggaagau uaucugggaa gacaucuucc ccaaacugaa gcucugggag	900
uuuuuuuau ugcagcucaa uaaggcugug gaacaauuca ggaggcugcu gaccaagag	960

ULPI_041_01W0_SeqList_ST25.txt

aaccggcgggg ucaccaaau	cgacccaau caaucuga	cuaucacca agaccugag	1020
uauaggcggu ucgggugcac	ugucgacaug aaauaagccc	ucacuacuuu uauuccccac	1080
gauaaaaggcc ccgccgccau	cgaggagugu ugcaacuggu	uccacaagag gauggaagag	1140
cucaacuccg aaaaacaccg	gcucaucaau uaccaccagg	agcaggccgu gaacugucug	1200
cugggcaacg ucuucuacga	gcggcucgcu gggcacgggc	ccaagcucgg cccugucacu	1260
aggaaacacc cucucgugac	ccgguaacuuc acuuuucccu	ucgaggaaau ugauuuuagc	1320
auggaggagu ccaugaucca	ccuccccaac aaggcuugcu	uccucauggc ccauaacggc	1380
ugggucaugg gcgacgaccc	ucugcggaau uucgcugagc	ccggguccga gguguaucuc	1440
cggaggggagc ugauuguuug	gggcgauagc gugaagcucc	gguaacggcaa caagcccga	1500
gauugcccuu accucugggc	ccaugaag aaguauacug	agauuacugc cacuuacuuu	1560
cagggcgucc ggcuggacaa	uugucauucc accccucugc	augucgcuga auauaugcug	1620
gacgcugcuc ggaaccugca	acccaaccuc uacgucgucg	cugagcucuu uaccggcagc	1680
gaggaccucg auaacgucuu	cgugacuagg cucgggauca	gcagccucau uagggaagcc	1740
augagcgccu acaacagcca	cgaggaaggg aggcucgugu	aucgguacgg cggcgagccu	1800
gugggcagcu ucgugcagcc	cugccugcgg ccucucaugc	ccgcuaucgc ccacggccuc	1860
uucauggaca ucacucacga	caacgaugc ccuaucgucc	acagguccgc uuacgacgu	1920
cugcccagca cuaccaucgu	guccauggcc ugcugcgua	gcggcagcac caggggguac	1980
gacgagcucg ucccucacca	gaucuccguc guguccgagg	agcgguuua uacuaaaugg	2040
aaccugagg cccugccuuc	cauaccggg gaggugaacu	uucaaagcgg gaucaucgcc	2100
gcccggugcg cuauuagcaa	gcugcaccag gaacugggcg	ccaaaggguu cauccagguc	2160
uacgucgacc aaguggacga	ggauauuguc gccgucacca	ggcauucccc uagcauucac	2220
caguccgugg ucgcugucuc	caggacugcu uuucggaacc	ccaaaacuuc cuucuauagc	2280
aaagaggucc cuaaaugug	uauuccuggg aagaucgagg	aggucgugcu ggaggcuagg	2340
acuaucgaaa ggaauacuaa	gccuuaccgg aaagacgaaa	acuccauuaa cggcaccccc	2400

ULPI_041_01W0_SeqList_ST25.txt

gacauuaccg ucgagaucag ggagcacauc cagcugaacg agagcaagau cgugaagcaa	2460
gcuggcgugg ccaccaaggg ccccaacgag uacauccaag agauugaguu cgagaaucug	2520
ucccccggca gcgugaucau uuuuagggug agccuggacc cccacgcca agucgcugug	2580
ggcauccugc ggaaccaccu cacccauuu agcccucuu ucaaguccgg gagccucgu	2640
guggauaacg ccgaccuau ucuaagauc ccuuucguu cccuggccag caggcucacu	2700
cuggcugaac ucaaccagau ucuguaucgg ugcgaguccg aggagaaaga ggacgggggc	2760
ggcuguuacg auauucccaa uugguccgcc cugaaguacg ccgggcugca agggcucaug	2820
uccguccugg ccgagauuag gcccaaaaac gaucuggggc acccuuucug caacaaccug	2880
agguccggcg acuggaugau cgauuacguc agcaaucggc ugaucucccg guccggcacu	2940
aucgcugagg uggggaagug gcugcaggcu auguuuuuu aucucaaaca gauuccccgg	3000
uaucugauuc ccugcuauuu cgacgcuaau cucaucgggg ccuacaccac ucugcucgac	3060
accgccugga aacagauguc cagcuucgug cagaacgggu ccaccuucgu caagcaucug	3120
ucccuggggg ccgugcaacu cugcggcguc ggcaaguuc cuagccucc cauccugucc	3180
ccugcccuca uggacguccc uuaccggcuc aacgaaaua ccaaggagaa agaacaaugc	3240
ugcgugucc ucgcugccgg gcucccucac uucuccuccg ggaucuuucg guguuggggc	3300
cgggacacuu ucaucgccc gagggggauu cuccucauca cuggccggua cgucgaggcc	3360
cggaacauca uccucgccuu cgucgggacc cuccggcacg ggcucaucc uaaccuccuc	3420
ggggagggca ucuacgccag guauaacugc cgggacgcug ucugguggug gcuccagugc	3480
auucaggacu auugcaagau ggugcccaac gggcucgaua uccucaaauug ucccugagc	3540
aggauuacc cuaccgacga cuccgcuccu cugccugcug ggaccucga ccagccccug	3600
uucgagguca uucaggaggc caugcaaaag cauauagcagg ggauucaguu ucgggagcgg	3660
aacgcugggc cccagauuga ccggaauaug aaagaugagg gguucaacau uacugccggc	3720
guggacgagg agaccggcuu cguguacggc ggcaaccggg uuaacugcgg gaccuggaug	3780
gacaagaugg gcgagagcga uaggguuagg aaauaggggca uucccgccac cccagggac	3840

ULPI_041_01W0_SeqList_ST25.txt

ggcuccgcug ucgagaucgu cgggcucagc aaguccgcug ugcgguggcu gcucgagcuc	3900
agcaagaaga acaucuuucc uuaccacgag gucaccguca agaggcacgg caaagcuauu	3960
aaagucuccu acgacgaug gaauaggaag auucaagaua auuucgagaa acucuuccac	4020
gugagcgagg auccuuccga ccucaacgag aaacacccca accucgugca uaagcggggg	4080
auuuauaagg acagcuacgg cgcuuccagc ccuuggugcg auuaccagcu ccggccuaac	4140
uucaccauug ccaugguggu cgccccugaa cucuucacua ccgagaaggc cuggaaggcu	4200
cuggaaaucg ccgagaagaa gcugcugggg cccugggga ugaagacucu cgaccccgac	4260
gacauggugu auugcggcau cuacgauaac gccucgaua acgauaaaua uaaccuggcu	4320
aagggguuuu acuaccauca agggccugaa uggcucuggc cuaucggcua cuuccugcgg	4380
gccaagcucu acuucagcag gcugaugggg cccgaaacua cugccaaaac uauugugcug	4440
gugaagaacg ugcugagccg gcacuacgug caccuggaaa ggagcccuug gaagggccug	4500
cccgagcuca ccaacgaaaa cgcccaguau ugcccuuuu gcugcgagac ccaagccugg	4560
uccaucgcua cuauccugga aaccucguac gaccucuag	4599

<210> 43

<211> 4599

<212> RNA

<213> Artificial Sequence

<220>

<223> Codon-Optimized AGL Coding Sequence

<400> 43

auggggcaca gcaaacaaau caggauccuc cuccucaacg aaauggaaaa acucgaaaaa	60
accucuuca ggcucgaaca aggguacgaa cuccaauuca ggcucgggcc gacgcuccaa	120
gggaaagcgg ugacggugua cacgaacuac ccguucccg gggaacguu caacagggaa	180
aaauucagga gccucgacug ggaaaacccg acggaaaggg aagacgacag cgacaaauac	240
ugcaaacuca accuccaaca aagcgggagc uuccaauacu acuuccucca agggaacgaa	300
aaaagcgggg ggggguacau cgugguggac ccgauccuca gggugggggc ggacaaccac	360

ULPI_041_01W0_SeqList_ST25.txt

gugcucccg	ucgacugcg	gacgcuccaa	acguuccug	cgaaaugccu	cgggccguuc	420
gacgaauggg	aaagcaggcu	caggguggcg	aaagaaagcg	gguacaacau	gauccacuuc	480
acgccgcucc	aaacgcucgg	gcucagcagg	agcugcuaca	gccucgcgaa	ccaacucgaa	540
cucaacccgg	acuucagcag	gccgaacagg	aaauacacgu	ggaacgacgu	ggggcaacuc	600
guggaaaaac	ucaaaaaaga	auggaacgug	aucugcauca	cggacguggu	guacaaccac	660
acggcggcg	acagcaaau	gauccaagaa	cacccggaau	gcgcguacaa	ccucgugaac	720
agcccgacc	ucaaacggc	gugggugcuc	gacagggcg	ucuggagguu	cagcugcgac	780
guggcgggaag	ggaaauacaa	agaaaaagg	aucccgcg	ucaucgaaaa	cgaccaccac	840
augaacagca	ucaggaaaau	caucugggaa	gacaucuucc	cgaaacucua	acucugggaa	900
uucuuccaag	uggacgugaa	caaagcggug	gaacaauuca	ggaggcuccu	cacgcaagaa	960
aacaggagg	ugacgaaaag	cgacccgaac	caacaccuca	cgaucaucca	agacccggaa	1020
uacaggaggu	ucgggugcac	gguggacaug	aacaucgcgc	ucacgacguu	caucccgcac	1080
gacaaaggcg	cggcggcgau	cgaagaagc	ugcaacuggu	uccacaaaag	gauggaagaa	1140
cucaacagcg	aaaaacacag	gcucaucaac	uaccaccaag	aacaagcggg	gaacugccuc	1200
cucgggaacg	uguucuacga	aaggcucgcg	gggcacgggc	cgaaacucgg	gccggugacg	1260
aggaaacacc	cgcucgugac	gagguacuuc	acguucccg	ucgaagaaau	cgacuucagc	1320
auggaagaaa	gcaugaucca	ccucccgaac	aaagcgugcu	uccucauggc	gcacaacggg	1380
ugggugaugg	gggacgaccc	gcucaggaac	uucgcggaac	cggggagcga	aguguaccuc	1440
aggagggaac	ucaucugcug	gggggacagc	gugaaacuca	gguacgggaa	caaaccggaa	1500
gacugcccgu	accucugggc	gcacaugaaa	aaauacacgg	aaaucacggc	gacguacuuc	1560
caagggguga	ggcucgacaa	cugccacagc	acgccgcucc	acguggcgga	auacaugcuc	1620
gacgcggcg	ggaaccucca	accgaaccuc	uacguggugg	cggaaacucu	cacggggagc	1680
gaagaccucg	acaacguguu	cgugacgagg	cucgggauca	gcagccucau	cagggaagcg	1740
augagcgcgu	acaacagcca	cgaagaagg	aggcucgugu	acagguacgg	gggggaaccg	1800

ULPI_041_01W0_SeqList_ST25.txt

gugggggagcu ucgugcaacc gugccucagg ccgcucaugc cggcgauvcg gcacgcgcug	1860
uucauggaca ucacgcacga caacgaauvc ccgaucgugc acaggagvcg guacgcgcg	1920
cucccgagca cgacgaucgu gagcauggcg ugcugvcgca gcgggagcac gaggggguac	1980
gacgaacucg ugccgcacca aaucagvcg gugagvcgaag aaagguucua cacgaaaugg	2040
aacccggaag cgcucccgag caacacgggg gaagugaacu uccaaagvcg gaucaucvcg	2100
gcgaggugvcg cgauvcgaa acuccaccaa gaacucgggg cgaaaggguu cauccaagug	2160
uacguggacc aaguggacga agacaucvcg gcggugacga ggcacagccc gagcauccac	2220
caaagvcugg uggcggugag caggacggcg uucaggaacc cgaaaacgag cuucuacagc	2280
aaagaagugc cgcaaaugug caucccgggg aaaaucgaag aaguggugcu cgaagvcagg	2340
acgaucgaaa ggaacacgaa accguacagg aaagacgaaa acagcaucua cgggacgccc	2400
gacaucacgg uggaaaucag ggaacacauc caacucaacg aaagcaaaau cgugaaacaa	2460
gcgggggugg cgacgaaagg gccgaacgaa uacauccaag aaucgaaau cgaaaaccuc	2520
agcccgggga gcgugaucau cuucagggug agccucgacc cgcacvcgca aguggcggug	2580
gggauccuca ggaaccaccu cacgcaauuc agcccgcacu ucaaaagvcg gagccucvcg	2640
guggacaacg cggacccgau ccucaaaauc ccguucvcga gccucvcgag caggcucacg	2700
cucgcggaac ucaaccaaau ccucuacagg ugcgaaagvc aagaaaaaga agacgggggg	2760
gggugcuacg acaucccgaa cuggagvcg cucaaaucg cggggcuca agggcucaug	2820
agcgvvcucg cggaaaucag gccgaaaaac gaccucgggc acccgvucug caacaaccuc	2880
aggagvcggg acuggaugau cgacuacvcg agcaacaggc uaucagvcg gagcgggacg	2940
aucgcggaag uggggaaaug gcuccaagvc auguucuuc accucaaaca aauccgagg	3000
uaccucaucc cgugcuacuu cgacvcgauc cucaucgggg cguaacvcg gcuccvcgac	3060
acggcgvuga aaaaaugag cagcuucvcg caaacggga gcacgvucgu gaaacaccuc	3120
agccucggga gcgugcaacu cugcggggug gggaaaucc cgagccucc gauccvcagc	3180
ccggcvcuca uggacgvcc guacaggvc aacgaaauc cgaaagaaaa agaacaavc	3240

ULPI_041_01W0_SeqList_ST25.txt

ugcgugagcc ucgcggcggg gcucccgcac uucagcagcg ggaucuucag gugcuggggg	3300
agggacacgu ucaucgcgcu caggggggauc cuccucauca cggggaggua cguggaagcg	3360
aggaacauca uccucgcguu cgcgggggacg cucaggcacg ggcucauccc gaaccuccuc	3420
ggggaaggga ucuacgcgag guacaacugc agggacgcgg ugugguggug gcuccaaugc	3480
auccaagacu acugcaaaau ggugccgaac gggcucgaca uccucaaauug cccggugagc	3540
aggauguacc cgacggacga cagcgcgccg cucccggcgg ggacgcucga ccaaccgcug	3600
uucgaaguga uccaagaagc gaugcaaaaa cacaugcaag ggauccaauu cagggaaagg	3660
aacgcggggc cgcaaaucga caggaacaug aaagacgaag gguucaacau cacggcgggg	3720
guggacgaag aaacgggguu cguguacggg gggaacaggu ucaacugcgg gacguggaug	3780
gacaaaaugg gggaaagcga cagggcgagg aacaggggga ucccggcgac gccgagggac	3840
gggagcgcgg ugghaaucgu ggggcucagc aaaagcgcgg ugagguggcu ccucgaacuc	3900
agcaaaaaaa acaucuuccc guaccacgaa gugacgguga aaaggcacgg gaaagcgauc	3960
aaagugagcu acgacgaau gaaacaggaaa auccaagaca acuuacgaaa acucuuccac	4020
gugagcgaag acccgagcga ccucaacgaa aaacacccga accucgugca caaaaggggg	4080
aucuacaaaag acagcuacgg ggcgagcagc ccguggugcg acuaccaacu caggccgaac	4140
uucacgaucg cgaugguggu ggcgccggaa cucuucacga cggaaaaagc guggaaagcg	4200
cucgaaaucg cggaaaaaaa acuccucggg ccgcucggga ugaaaaacgu cgacccggac	4260
gacauggugu acugcgggau cuacgacaac gcgcucgaca acgacaacua caaccucgcg	4320
aaaggguuca acuaccacca agggccggaa uggcucuggc cgauagggua cuuccucagg	4380
gcgaaacucu acuuacgag gcucaugggg ccggaaacga cggcgaaaac gaucgugcuc	4440
gugaaaaacg ugcucagcag gcacuacgug caccucgaaa ggagcccug gaaagggcuc	4500
ccggaacuca cgaacgaaaa cgcgcaauac ugcccguuca gcugcgaaac gcaagcgugg	4560
agcaucgcga cgauccucga aacgcucuac gaccucuga	4599

ULPI_041_01W0_SeqList_ST25.txt

<210> 44
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 44
 augggccaca gcaagcagau caggauccuc cuccucaacg aaauggaaaa gcucgaaaag 60
 acacucuuca ggcucgaaca gggcuacgaa cuccaguuca ggcucggccc aacacuccag 120
 ggcaaggccg ugacagugua cacaaacuac ccuuucccag gcgaaacauu caacagggaa 180
 aaguucagga gccucgacug ggaaaaccca acagaaaggg aagacgacag cgacaaguac 240
 ugcaagcuca accuccagca gagcggcagc uuccaguacu acuuccucca gggcaacgaa 300
 aagagcggcg gcggcuacau cgugguggac ccaauccuca ggguggggcg cgacaaccac 360
 gugcucccac ucgacugcgu gacacuccag acauuccucg ccaagugccu cggcccauuc 420
 gacgaauggg aaagcaggcu caggguggcc aaggaaagcg gcuacaacau gauccacuuc 480
 acaccacucc agacacucgg ccucagcagg agcugcuaca gccucgcca ccagcucgaa 540
 cucaaccag acuuacagcag gccaaacagg aaguacacau ggaacgacgu gggccagcuc 600
 guggaaaaagc ucaagaagga auggaacgug aucugcauca cagacguggu guacaaccac 660
 acagccgcca acagcaagug gauccaggaa caccagaau gcgccuacaa ccucgugaac 720
 agcccacacc ucaagccagc cugggugcuc gacagggccc ucuggagguu cagcugcgac 780
 guggccgaag gcaaguacaa ggaaaagggc aucccagccc ucaucgaaaa cgaccaccac 840
 augaacagca ucaggaagau caucugggaa gacaucuucc caaagcucaa gcucugggaa 900
 uucuuccagg uggacgugaa caaggccgug gaacaguuca ggaggcuccu cacacaggaa 960
 aacaggaggg ugacaaagag cgacccaaac cagcaccuca caaucaucca ggaccagaa 1020
 uacaggaggu ucggcugcac aguggacaug aacaucgccc ucacaacauu caucccacac 1080
 gacaagggcc cagccgccau cgaagaugc ugcaacuggu uccacaagag gauggaagaa 1140
 cucaacagcg aaaagcacag gcucaucaac uaccaccagg aacaggccgu gaacugccuc 1200

ULPI_041_01W0_SeqList_ST25.txt

cucggcaacg uguuucacga aaggcucgcc ggccacggcc caaagcucgg cccagugaca	1260
aggaagcacc cacucgugac aagguacuuc acauucccau ucgaagaaau cgacuucagc	1320
auggaagaaa gcaugaucca ccucccaaac aaggccugcu uccucauggc ccacaacggc	1380
ugggugaugg gcgacgaccc acucaggaac uucgccgaac caggcagcga aguguaccuc	1440
aggagggaac ucaucugcug gggcgacagc gugaagcuca gguacggcaa caagccagaa	1500
gacugcccau accucugggc ccacaugaag aaguacacag aaauacacagc cacauacuuc	1560
cagggcguga ggucgacaa cugccacagc acaccacucc acguggccga auacaugcuc	1620
gacgccgcca ggaaccucca gccaaaccuc uacguggugg ccgaacucuu cacaggcagc	1680
gaagaccucg acaacguguu cgugacaagg cucggcauca gcagccucau cagggaagcc	1740
augagcgccu acaacagcca cgaagaaggc aggcucgugu acagguacgg cggcgaacca	1800
gugggcagcu ucgugcagcc augccucagg ccacucaugc cagccaucgc ccacgcccuc	1860
uucauggaca ucacacacga caacgaugc ccaaucgugc acaggagcgc cuacgacgcc	1920
cucccaagca caacaucgu gagcauggcc ugcugcgcca gcggcagcac aaggggcuac	1980
gacgaacucg ugccacacca gaucagcgug gugagcgaag aaagguucua cacaaagugg	2040
aacccagaag cccuccaag caacacaggc gaagugaacu uccagagcgg caucaucgcc	2100
gccaggugcg ccaucagcaa gcuccaccag gaacucggcg ccaagggcuu cauccaggug	2160
uacguggacc agguggacga agacaucgug gccgugacaa ggcacagccc aagcauccac	2220
cagagcgugg uggccgugag caggacagcc uucaggaacc caaagacaag cuucuacagc	2280
aaggaagugc cacagaugug caucccaggc aagaucgaag aaguggugcu cgaagccagg	2340
acaaucgaaa ggaacacaaa gccauacagg aaggacgaaa acagcaucua cggcacacca	2400
gacaucacag uggaauucag ggaacacauc cagcucaacg aaagcaagau cgugaagcag	2460
gccggcgugg ccacaaaggg cccaaacgaa uacauccagg aaaucauuu cgaaaaccuc	2520
agcccaggca gcgugaucuu cuucagggug agccucgacc cacacgccc gguaggccgug	2580
ggcauccuca ggaaccaccu cacacaguuc agcccacacu ucaagagcgg cagccucgcc	2640

ULPI_041_01W0_SeqList_ST25.txt

guggacaacg ccgacccaau ccucaagauc ccuucgccg gccucgccag caggcucaca	2700
cucgccgaac ucaaccagau ccucuacagg ugcgaaagcg aagaaaagga agacggcggc	2760
ggcugcuacg acaucccaaa cuggagcgcc cucaaguacg ccggccucca gggccucaug	2820
agcgugcucg ccgaaucag gccaaagaac gaccucggcc acccauucug caacaaccuc	2880
aggagcggcg acuggaugau cgacuacgug agcaacaggc ucaucagcag gagcggcaca	2940
aucgccgaag ugggcaagug gcuccaggcc auguucuucu accucaagca gaucccaagg	3000
uaccucaucc caugcuacuu cgacgccauc cucaucggcg ccuacacaac acuccucgac	3060
acagccugga agcagaugag cagcuucgug cagaacggca gcacauucgu gaagcaccuc	3120
agccucggca gcgugcagcu cugcggcgug ggcaaguucc caagccuccc aauccucagc	3180
ccagcccuca uggacgugcc auacaggcuc aacgaaauc caaaggaaaa ggaacagugc	3240
ugcgugagcc ucgccgccgg ccucccacac uucagcagcg gcaucuucag gugcuggggc	3300
agggacacau ucaucgccc caggggcauc cuccucauca caggcaggua cguggaagcc	3360
aggaacauca uccucgccuu cgccggcaca cucaggcacg gccucaucc aaaccuccuc	3420
ggcgaaggca ucuacgccag guacaacugc agggacgccg ugugguggug gcuccagugc	3480
auccaggacu acugcaagau ggugccaaac ggccucgaca uccucaagug cccagugagc	3540
aggauguacc caacagacga cagcgcccca cuccagccg gcacacucga ccagccacuc	3600
uucgaaguga uccaggaagc caugcagaag cacaugcagg gcauccaguu cagggaagg	3660
aacgccggcc cacagaucga caggaacaug aaggacgaag gcuucaacau cacagccggc	3720
guggacgaag aaacaggcuu cguguacggc ggcaacaggu ucaacugcgg cacauggaug	3780
gacaagaugg gcgaaagcga cagggccagg aacaggggca ucccagccac accaagggac	3840
ggcagcgccg uggaaaucgu gggccucagc aagagcgccg ugagguggcu ccucgaacuc	3900
agcaagaaga acaucuucc auaccacgaa gugacaguga agaggcacgg caaggccauc	3960
aaggugagcu acgacgaug gaacaggaag auccaggaca acuucgaaaa gcuguuccac	4020
gugagcgaag acccaagcga ccucaacgaa aagcacccaa accucgugca caagaggggc	4080

ULPI_041_01W0_SeqList_ST25.txt

aucuacaagg acagcuacgg cgccagcagc ccauggugcg acuaccagcu caggccaaac	4140
uucacaaucg ccaugguggu ggccccagaa cucuucacaa cagaaaaggc cuggaaggcc	4200
cucgaaaucg ccgaaaagaa gcuccucggc ccacucggca ugaagacacu cgacccagac	4260
gacauggugu acugcggcau cuacgacaac gccucgaca acgacaacua caaccucgcc	4320
aagggcuuca acuaccacca gggcccagaa uggcucuggc caaucggcua cuuccucagg	4380
gccaagcucu acuucagcag gcucaugggc ccagaaacaa cagccaagac aaucgugcuc	4440
gugaagaacg ugcucagcag gcacuacgug caccucgaaa ggagcccaug gaagggccuc	4500
ccagaacuca caaacgaaaa cgcccaguac ugcccauuca gcugcgaaac acaggccugg	4560
agcaucgcca caauccucga aacacucuac gaccucuga	4599

<210> 45
 <211> 4599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Codon-Optimized AGL Coding Sequence

<400> 45	
augggccaca gcaagcagau ccggaucuug cugcugaacg agauggagaa gcuggagaag	60
accuguuca ggcuggagca gggcuacgag cugcaguucc gguugggccc caccuugcag	120
ggcaaggccg ugaccgugua caccaacuac ccuuccccg gcgagacguu caacaggagg	180
aaguuccggu ccuuggacug ggagaacccc accgagaggg aggacgacuc cgacaaguac	240
ugcaagcuga accugcagca guccggcucc uuccaguacu acuuccugca gggcaacgag	300
aagaguggcg gcggcuacau cgugguggac cccaucuugc gggugggcg cgacaaccac	360
gugcugcccu uggacugcgu gaccucgag accuucuugg ccaagugcuu gggcccuuc	420
gacgaguggg agagcaggcu gaggguggcc aaggaguccg gcuacaacau gauccacuuc	480
accccuugc agaccuggg ccuguccagg uccugcuacu ccuuggccaa ccaguuggag	540
uugaaccccg acuucuccag gcccaacagg aaguacaccu ggaacgacgu gggccagcug	600

ULPI_041_01W0_SeqList_ST25.txt

guggagaagu ugaagaagga guggaacgug aucugcauca ccgacguggu guacaaccac	660
accgccgcca acagcaagug gauccaggag caccgccagu gcgccuacaa ccuggugaac	720
uccccccacu ugaagcccg cuggguguug gacagggccc uguggcgguu cuccugcgac	780
guggccgagg gcaaguacaa ggagaagggc aucccgccu ugaucgagaa cgaccaccac	840
augaacucca uccggaagau caucugggag gacaucuucc ccaagcugaa gcugugggag	900
uucuuccagg uggacugaa caaggccgug gagcaguuca ggaggcugcu gaccaggag	960
aacaggcggg ugaccaaguc cgacccaac cagcaccuga ccaucaucca ggaccccgag	1020
uacaggcggu ucggcugcac cguggacaug aacaucgcc ugaccaccuu caucacccac	1080
gacaagggcc ccgccgccau cgaggagugc ugcaacuggu uccacaagag gauggaggag	1140
uugaacuccg agaagcaccg gcugaucaac uaccaccagg agcaggccgu gaacugccug	1200
uugggcaacg uguucuacga gcggcuggcc ggccacggcc ccaagcuggg ccccgugacc	1260
aggaagcacc ccuuggugac cagguacuuc accuucccu ucgaggagau cgacuucucc	1320
auggaggagu ccaugaucca ccugcccaac aaggccugcu uccugauggc ccacaacggc	1380
ugggugaugg gcgacgacc ccugcggaac uucgccgagc ccggcuccga gguguaccug	1440
aggagggagc ugaucugcug gggcgacagc gugaaguugc gguacggcaa caagcccag	1500
gacugcccu accugugggc ccacaugaag aaguacaccg agaucaccgc caccuacuuc	1560
cagggcgugc ggcuggacaa cugccacucc acccccugc acguggccga guacauguug	1620
gacgccgcca ggaacuugca gcccaacuug uacguggugg ccgagcuguu caccggcagc	1680
gaggaccugg acaacguguu cgugaccagg cugggcauca gcuccuugau cagggaggcc	1740
augagcgccu acaacagcca cgaggagggc agguuggugu accgguacgg cggcgagccc	1800
gugggcuccu ucgugcagcc cugcuugagg cccuugaugc ccgccaucgc ccacgccug	1860
uucauggaca ucacccacga caacgagugc cccaucgugc acagguccgc cuacgacgcc	1920
cugcccagca ccaccaucgu guccauggcc ugcugcgcca gcggcagcac caggggcuac	1980
gacgaguugg ugccccacca gaucuccgug guguccgagg agcgguucua caccaagugg	2040

ULPI_041_01W0_SeqList_ST25.txt

aaccccgagg ccuugcccuc caacaccggc gaggugaacu uccagagcgg caucaucgcc	2100
gccaggugcg ccaucagcaa gcugcaccag gagcugggcg ccaaggguu cauccaggug	2160
uacguggacc agguggacga ggacaucgug gccgugacca ggcacucucc cagcauccac	2220
caguccgugg uggccguguc caggaccgcc uucaggaacc ccaagaccuc cuucuacagc	2280
aaggaggugc cccagaugug caucuccggc aagaucgagg agguggugcu ggaggccagg	2340
accaucgaga ggaacaccaa gcccuacagg aaggacgaga acuccaucaa cggcaccccc	2400
gacaucaccg uggagaucag ggagcacauc cagcugaacg agagcaagau cgugaagcag	2460
gccggcgugg ccaccaaggg cccaacgag uacauccagg agaucgaguu cgagaacuug	2520
ucccccgga gcgugaucau cuucagggug agccuggacc cccacgcca gguggccgug	2580
ggcauccugc ggaaccaccu gaccaguuc agccccacu ucaaguccgg cagccuggcc	2640
guggacaacg ccgaccccau cuugaagauc cccuucgccu cccuggccuc cagguugacc	2700
uuggccgagc ugaaccagau ccuguaccgg ugcgaguccg aggagaagga ggacggcggc	2760
ggcugcuacg acaucccaa cugguccgcc cugaaguacg ccggccugca gggcuugaug	2820
uccguguugg ccgagaucag gccaagaac gacuugggcc accccuucug caacaacuug	2880
agguccggcg acuggaugau cgacuacgug agcaaccggc ugaucucccg guccggcacc	2940
aucgccgagg ugggcaagug guugcaggcc auguucuucu accugaagca gaucggcg	3000
uaccugaucc ccugcuacuu cgacgccauc ugaucggcg ccuacaccac ccugcuggac	3060
accgccugga agcagauguc cagcuucgug cagaacggcu ccaccuucgu gaagcaccug	3120
uccuugggcu ccgugcagcu gugcggcgug ggcaaguucc ccuccugcc cauccugucc	3180
cccgcccuga uggacgugcc cuacagguug aacgagauca ccaaggagaa ggagcagugc	3240
ugcguguccc uggccgccgg cuugccccac uucuccuccg gcaucuuccg gugcuggggc	3300
agggacaccu ucaucgccc uaggggcauc cugcugauca ccggccgguu cguggaggcc	3360
aggaacauc ucuuggccuu cgccggcacc cugaggcacg gccugauccc caaccugcug	3420
ggcgagggca ucuacgccag guacaacugc cgggacgccg ugugguggug gcugcagugc	3480

ULPI_041_01W0_SeqList_ST25.txt

auccaggacu acugcaagau ggugcccaac ggccuggaca uccugaagug ccccgugucc	3540
aggauguacc ccaccgacga cuccgcccc uugcccgccg gcaccugga ccagcccuug	3600
uucgagguga uccaggaggc caugcagaag cacaugcagg gcauccaguu ccgggagagg	3660
aacgccggcc cccagaucga ccggaacaug aaggacgagg gcuucaacau caccgccggc	3720
guggacgagg agacuggcuu cguguacggc ggcaaccggu ucaacugcgg caccuggaug	3780
gacaagaugg gcgagagcga cagggccagg aacaggggca uccccgccac cccagggac	3840
ggcuccgccg uggagaucgu gggccugagc aaguccgccg ugcggugguu gcuggaguug	3900
uccaagaaga acaucuucc cuaccacgag gugaccguga agaggcacgg caaggccauc	3960
aagguguccu acgacgagug gaacaggaag auccaggaca acuucgagaa gcuguuccac	4020
guguccgagg accccuccga cuugaacgag aagcacccca accuggugca caagcggggc	4080
aucuacaagg acagcuacgg cgccuccagc cccuggugcg acuaccagcu gaggcccaac	4140
uucaccaucg ccaugguggu ggccccgag cuguucacca ccgagaaggc cuggaaggcc	4200
uuggagaucg ccgagaagaa guugcugggc cccugggca ugaagaccuu ggacccccgac	4260
gacauggugu acugcggcau cuacgacaac gccuuggaca acgacaacua caaccuggcc	4320
aagggcuuca acuaccacca gggccccgag uggcuguggc ccaucggcua cuuccugcgg	4380
gccaaguugu acuucuccag guugaugggc cccgagacga ccgccaagac caucguguug	4440
gugaagaacg ugcugucccg gcacuacgug caccuggaga gguccccug gaagggccug	4500
cccgagcuga ccaacgagaa cgcccaguac ugccccuua gcugcgagac gcaggccugg	4560
uccaucgcca ccauccugga gacgcuguac gacuuguag	4599