



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2016/12/16
(87) Date publication PCT/PCT Publication Date: 2017/06/22
(85) Entrée phase nationale/National Entry: 2018/05/31
(86) N° demande PCT/PCT Application No.: US 2016/067393
(87) N° publication PCT/PCT Publication No.: 2017/106799
(30) Priorités/Priorities: 2015/12/17 (US62/269,089);
2015/12/17 (US62/269,092); 2015/12/30 (US62/273,112);
2015/12/30 (US62/273,108); 2016/01/04 (US62/274,727);
2016/01/04 (US62/274,726); 2016/01/04 (US62/274,733);
2016/01/04 (US62/274,722); 2016/05/18 (US62/338,478);
2016/05/18 (US62/338,456); 2016/10/17 (US62/409,343)

(51) Cl.Int./Int.Cl. *C12N 9/90* (2006.01),
A61K 31/7088 (2006.01), *A61K 9/127* (2006.01),
A61P 3/00 (2006.01), *C12N 15/61* (2006.01)
(71) Demandeur/Applicant:
MODERNATX, INC., US
(72) Inventeurs/Inventors:
MARTINI, PAOLO, US;
PRESNYAK, VLADIMIR, US
(74) Agent: SMART & BIGGAR

(54) Titre : POLYNUCLEOTIDES CODANT POUR LA METHYLMALONYL-COA MUTASE
(54) Title: POLYNUCLEOTIDES ENCODING METHYLMALONYL-COA MUTASE

(57) **Abrégé/Abstract:**

The disclosure relates to polynucleotides comprising an open reading frame of linked nucleosides encoding human methylmalonyl-CoA mutase precursor, human methylmalonyl-CoA mutase (MCM) mature form, or functional fragments thereof. In some embodiments, the disclosure includes methods of treating methylmalonic acidemia in a subject in need thereof comprising administering an mRNA encoding an MCM polypeptide.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 June 2017 (22.06.2017)

(10) International Publication Number
WO 2017/106799 A1

- (51) International Patent Classification:
C12N 9/90 (2006.01) A61K 31/7088 (2006.01)
C12N 15/61 (2006.01) A61K 9/127 (2006.01)
A61P 3/00 (2006.01)

(21) International Application Number:
PCT/US2016/067393

(22) International Filing Date:
16 December 2016 (16.12.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/269,089 17 December 2015 (17.12.2015) US
62/269,092 17 December 2015 (17.12.2015) US
62/273,112 30 December 2015 (30.12.2015) US
62/273,108 30 December 2015 (30.12.2015) US
62/274,727 4 January 2016 (04.01.2016) US
62/274,722 4 January 2016 (04.01.2016) US
62/274,733 4 January 2016 (04.01.2016) US
62/274,726 4 January 2016 (04.01.2016) US
62/338,478 18 May 2016 (18.05.2016) US
62/338,456 18 May 2016 (18.05.2016) US
62/409,343 17 October 2016 (17.10.2016) US

(71) Applicant: MODERNATX, INC. [US/US]; 200 Technology Square, Cambridge, Massachusetts 02139 (US).

(72) Inventors: MARTINI, Paolo; C/o MODERNATX, INC., 200 Technology Square, Cambridge, Massachusetts 02139 (US). PRESNYAK, Vladimir; 123 Webster Avenue, #3, Cambridge, MA 02141 (US).

(74) Agents: KIM, Ji-Eun et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C, 1100 New York Avenue, NW, Washington, District of Columbia 20005 (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, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:
— with international search report (Art. 21(3))
— with sequence listing part of description (Rule 5.2(a))
- (54) Title: POLYNUCLEOTIDES ENCODING METHYLMALONYL-CoA MUTASE
- (57) Abstract: The disclosure relates to polynucleotides comprising an open reading frame of linked nucleosides encoding human methylmalonyl-CoA mutase precursor, human methylmalonyl-CoA mutase (MCM) mature form, or functional fragments thereof. In some embodiments, the disclosure includes methods of treating methylmalonic acidemia in a subject in need thereof comprising administering an mRNA encoding an MCM polypeptide.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 3
CONTENANT LES PAGES 1 À 163

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 3
CONTAINING PAGES 1 TO 163

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

POLYNUCLEOTIDES ENCODING METHYLMALONYL-CoA MUTASE

BACKGROUND OF THE DISCLOSURE

[0001] Methylmalonic acidemia (MMA) is a metabolic disorder characterized by the abnormal buildup of the metabolic byproduct methylmalonic acid in patients. MMA causes developmental delay, intellectual disability, kidney disease, coma, or even death. MMA is also referred to as methylmalonic aciduria. It has an estimated incidence of 1 in 50,000 to 100,000. Current treatment for MMA is primarily via dietary control to limit the usage of metabolic pathways that lead to methylmalonic acid formation. In serious cases, kidney and liver transplants have also been performed to provide a new reservoir of cells that can properly metabolize and remove the methylmalonic acid. However, none of these treatments completely or reliably controls the disorder. As such there is a need for improved therapy to treat MMA.

[0002] The principal gene associated with MMA is methylmalonyl-CoA mutase (NM_000255; NP_000246; also referred to as MCM or MUT). MCM is a metabolic enzyme (E.C. 5.4.99.2) that plays a critical role in the catabolism of various amino acids, fatty acids, and cholesterol. MCM's biological function is to isomerize L-methylmalonyl-CoA into succinyl-CoA, a Krebs cycle intermediate. MCM localizes to the mitochondria of cells, exists as a homodimer in its native form and is adenosylcobalamin-dependent. The precursor form of human MCM is 750 amino acids, while its mature form is 718 amino acids – a 32 amino acid leader sequence is cleaved off by mitochondrial importation and processing machinery. This leader sequence is variously referred to as MCM's mitochondrial targeting peptide, mitochondrial targeting sequence, or mitochondrial transit peptide.

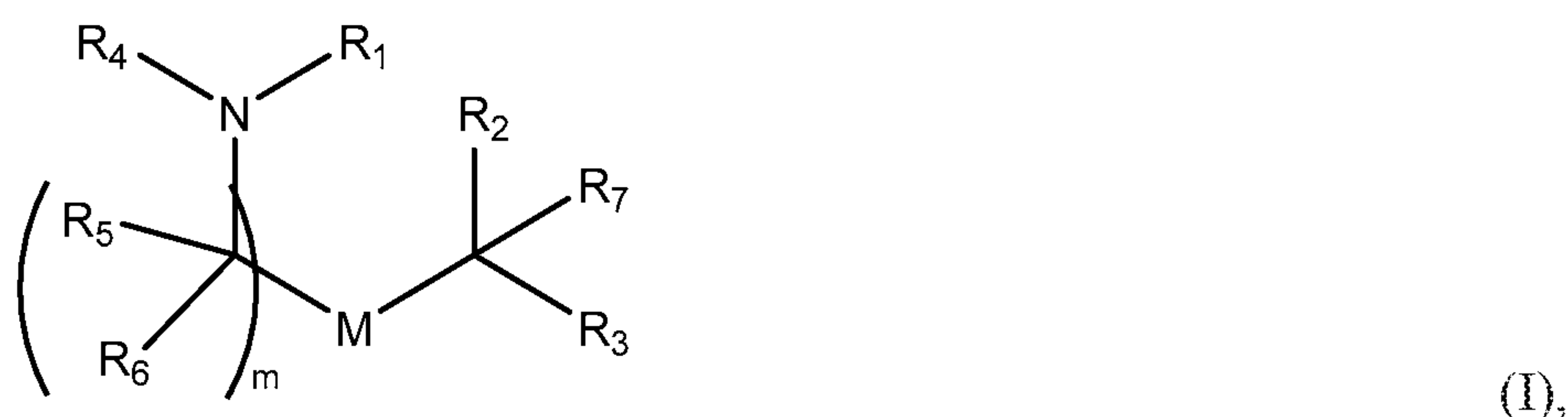
[0003] A complete or partial loss of MCM function leads to buildup of abnormal metabolites and metabolic intermediates upstream of MCM, such as methylmalonic acid, propionyl-carnitine, acetyl-carnitine, propionyl-CoA, D-methylmalonyl-CoA and L-methylmalonyl-CoA. For example, loss of MCM has been reported to lead to a 1000-fold increase in the methylmalonic acid. Nonetheless, there is no currently available therapeutic to treat MMA.

SUMMARY OF THE DISCLOSURE

[0004] The present disclosure provides methods of treating methylmalonic acidemia in a subject, the methods comprising administering to the subject an effective amount of a polynucleotide comprising an mRNA encoding an MCM polypeptide, wherein the

administration alleviates the symptoms of methylmalonic acidemia in the subject. The present disclosure also provides compositions comprising a polynucleotide sequence encoding an MCM polypeptide. In some embodiments, the compositions include a delivery agent.

[0005] In some embodiments, the composition comprises a polynucleotide that comprises an open reading frame (ORF) encoding an MCM polypeptide and a delivery agent, wherein the delivery agent comprises a compound having the formula (I)



or a salt or stereoisomer thereof, wherein

R_1 is selected from the group consisting of C_{5-20} alkyl, C_{5-20} alkenyl, $-R^*YR''$, $-YR''$, and $-R''M'R'$;

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, $-R^*YR''$, $-YR''$, and $-R^*OR''$, or R_2 and R_3 , together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a carbocycle, heterocycle,

$-OR$, $-O(CH_2)_nN(R)_2$, $-C(O)OR$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-N(R)_2$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, and $-C(R)N(R)_2C(O)OR$, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13; and

provided when R₄ is -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, or -CQ(R)₂, then (i) Q is not -N(R)₂ when n is 1, 2, 3, 4 or 5, or (ii) Q is not 5, 6, or 7-membered heterocycloalkyl when n is 1 or 2.

[0006] In some embodiments, the delivery agent further comprises a phospholipid, a structural lipid, a PEG lipid, or any combination thereof.

[0007] In some embodiments, the polynucleotides comprise an ORF having significant sequence similarity to a polynucleotide selected from the group of SEQ ID NOs: 1-207, 732-765, and 772, wherein the ORF encodes an MCM polypeptide. In some embodiments, the polynucleotides comprise an ORF having significant sequence similarity to a polynucleotide selected from the group of SEQ ID NOs: 151, 152, 153, 154, 732, 733, and 734 (FIGS. 9-15), wherein the ORF encodes an MCM polypeptide. In some embodiments, the polynucleotide comprises an ORF having significant sequence similarity to SEQ ID NO: 734 (FIG. 11), wherein the ORF encodes an MCM polypeptide.

[0008] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 99% or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 734, wherein the ORF encodes an MCM polypeptide.

[0009] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF sequence having at least 98%, at least 99%, or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 732, wherein the ORF encodes an MCM polypeptide.

[0010] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group

consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 182, 733, and 741, wherein the ORF encodes an MCM polypeptide.

[0011] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 735, 736, 738, 743, 744, 748, 749, 750, 754, 755, 758, 762, and 765, wherein the ORF encodes an MCM polypeptide.

[0012] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF sequence having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 180, 187, 737, 739, 740, 742, 745, 746, 747, 751, 752, 753, 757, 759, 760, 761, 763, and 764, wherein the ORF encodes an MCM polypeptide.

[0013] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 181 and 756, wherein the ORF encodes an MCM polypeptide.

[0014] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175, wherein the ORF encodes an MCM polypeptide.

[0015] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 164, 166, 167, 168, 169, 170, 172, 177, 178, 179, 195, and 204, wherein the ORF encodes an MCM polypeptide.

[0016] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at

least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 156, 157, 158, 159, 160, 161, 162, 174 and 176, wherein the ORF encodes an MCM polypeptide.

[0017] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 155 and 203, wherein the ORF encodes an MCM polypeptide.

[0018] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 64, 66, 71, 91, and 128, wherein the ORF encodes an MCM polypeptide.

[0019] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 9, 11, 18, 19, 21, 22, 23, 24, 32, 33, 37, 39, 40, 44, 45, 47, 50, 51, 52, 55, 57, 61, 65, 70, 79, 84, 86, 88, 90, 92, 98, 100, 115, 117, 126, 129, 135, 136, 137, 144, 148, 150, 184, 190, 191, and 206, wherein the ORF encodes an MCM polypeptide.

[0020] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 3, 5, 6, 8, 10, 12, 14, 16, 17, 20, 27, 28, 29, 31, 34, 35, 36, 38, 41, 42, 43, 46, 48, 49, 53, 54, 56, 58, 60, 63, 68, 69, 74, 77, 78, 80, 83, 85, 87, 93, 95, 96, 97, 99, 102, 103, 104, 105, 107, 110, 112, 113, 114, 116, 119, 120, 122, 123, 124, 125, 127, 131, 132,

133, 134, 138, 139, 140, 141, 142, 143, 147, 149, 183, 186, 188, 189, 192, 193, 194, 196, 197, 198, 199, 200, 201, 202, 205, and 207, wherein the ORF encodes an MCM polypeptide.

[0021] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 1, 2, 4, 7, 13, 15, 25, 26, 30, 59, 62, 67, 72, 73, 75, 76, 81, 82, 89, 94, 101, 106, 108, 109, 111, 118, 121, 130, 145, 146, and 185, wherein the ORF encodes an MCM polypeptide.

[0022] In some embodiments, the polynucleotides further comprise a nucleotide sequence encoding a transit peptide, *e.g.*, mitochondrial transit peptide. The mitochondrial transit peptide can be any peptide that facilitates the transport of MCM to mitochondria or localization of MCM in mitochondria. In some embodiments, the polynucleotide comprises a nucleotide sequence encoding a mitochondrial transit peptide selected from the group listed in **Table 1** (SEQ ID NOs: 251 to 265). In some embodiments, the polynucleotide comprises a nucleotide sequence encoding a mitochondrial transit peptide selected from the group consisting of SEQ ID NOs: 270 to 719.

[0023] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 99% or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 734, wherein the ORF encodes an MCM polypeptide.

[0024] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 98%, at least 99%, or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 732, wherein the ORF encodes an MCM polypeptide.

[0025] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 182 and 733, wherein the ORF encodes an MCM polypeptide.

[0026] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the

group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 735, 741, 743, 744, 748, 758, 762, and 765, wherein the ORF encodes an MCM polypeptide.

[0027] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 180, 181, 736, 738, 739, 740, 742, 746, 747, 749, 750, 751, 752, 753, 754, 755, 757, 759, 760, 761, and 763, wherein the ORF encodes an MCM polypeptide.

[0028] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 745, 756, and 764, wherein the ORF encodes an MCM polypeptide.

[0029] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175, wherein the ORF encodes an MCM polypeptide.

[0030] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 166, 167, 168, 169, 170, 172, 177, 178, 179, 187, and 204, wherein the ORF encodes an MCM polypeptide.

[0031] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 156, 157, 158, 159, 160, 162, 164, 174, 176, 195, and 737, wherein the ORF encodes an MCM polypeptide.

[0032] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at

least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 155, 161, and 203, wherein the ORF encodes an MCM polypeptide.

[0033] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 71 and 128, wherein the ORF encodes an MCM polypeptide.

[0034] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 4, 6, 8, 9, 11, 19, 22, 23, 24, 32, 33, 37, 40, 44, 45, 47, 51, 61, 64, 65, 66, 79, 84, 86, 90, 91, 92, 100, 101, 112, 115, 117, 126, 129, 135, 136, 146, 148, 184, 190, and 191, wherein the ORF encodes an MCM polypeptide.

[0035] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 2, 3, 5, 7, 10, 12, 13, 14, 15, 16, 18, 20, 21, 26, 27, 28, 29, 31, 34, 36, 38, 39, 41, 42, 43, 46, 48, 49, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 68, 69, 70, 72, 73, 74, 76, 77, 80, 83, 85, 88, 95, 96, 97, 98, 102, 104, 105, 106, 107, 108, 109, 110, 113, 114, 120, 121, 122, 123, 124, 127, 131, 132, 133, 134, 137, 138, 139, 140, 141, 142, 144, 145, 147, 149, 150, 186, 188, 189, 192, 193, 194, 196, 198, 199, 200, 202, 205, 206, and 207, wherein the ORF encodes an MCM polypeptide.

[0036] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at

least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 1, 17, 25, 30, 35, 50, 63, 67, 75, 78, 81, 82, 87, 89, 93, 94, 99, 103, 111, 116, 118, 119, 125, 130, 143, 183, 185, 197, and 201, wherein the ORF encodes an MCM polypeptide.

[0037] In some embodiments, the disclosure is directed to polynucleotides that encode functional MCMs or fragments thereof. In some embodiments, the disclosure provides polynucleotides that encode functional human MCMs (SEQ ID NO: 208, SEQ ID NO: 209, SEQ ID NO: 210, SEQ ID NO: 211, SEQ ID NO: 212, and SEQ ID NO: 213). In some embodiments, the disclosure provides polynucleotides that encode functional MCM polypeptides having at least one point mutation in the MCM sequence, while still retaining MCM enzymatic activity. In some embodiments, the encoded MCM polypeptide comprises one or more of the point mutations V69, T499, H532, A598, and V671, as defined by the polypeptide sequences in SEQ ID NO: 209, SEQ ID NO: 210, SEQ ID NO: 211, SEQ ID NO: 212, and SEQ ID NO: 213, respectively. In some embodiments, the polynucleotides are fully or partially modified (*e.g.*, chemically and/or structurally) in a manner as to avoid the deficiencies of other molecules of the art. The polynucleotides of the disclosure can be synthesized as an IVT polynucleotide, chimeric polynucleotide or a circular polynucleotide and such embodiments are contemplated.

[0038] In some embodiments, the polynucleotide is a DNA or RNA that comprises at least one chemically modified nucleoside. In some embodiments, the at least one chemically modified nucleoside is selected from any of those described herein.

[0039] In some embodiments, the polynucleotide further comprises or encodes a 5' UTR. In other embodiments, the polynucleotide further comprises or encodes a 3' UTR. In some embodiments, the UTR comprises or encodes a miRNA (*e.g.*, miR-142-3p, miR-142-5p, miR-126-3p, and/or miR-126-5p). In some embodiments, the polynucleotide further comprises a 5' terminal cap. In some embodiments, the polynucleotide further comprises or encodes a 3' polyA tail.

[0040] In some embodiments, the polynucleotide is RNA, *e.g.*, mRNA. In some embodiments, the mRNA comprises the sequences listed in SEQ ID NOs: 766-771.

[0041] In some embodiments, the polynucleotide is an RNA polynucleotide that is formulated in a lipid nanoparticle (LNP) carrier.

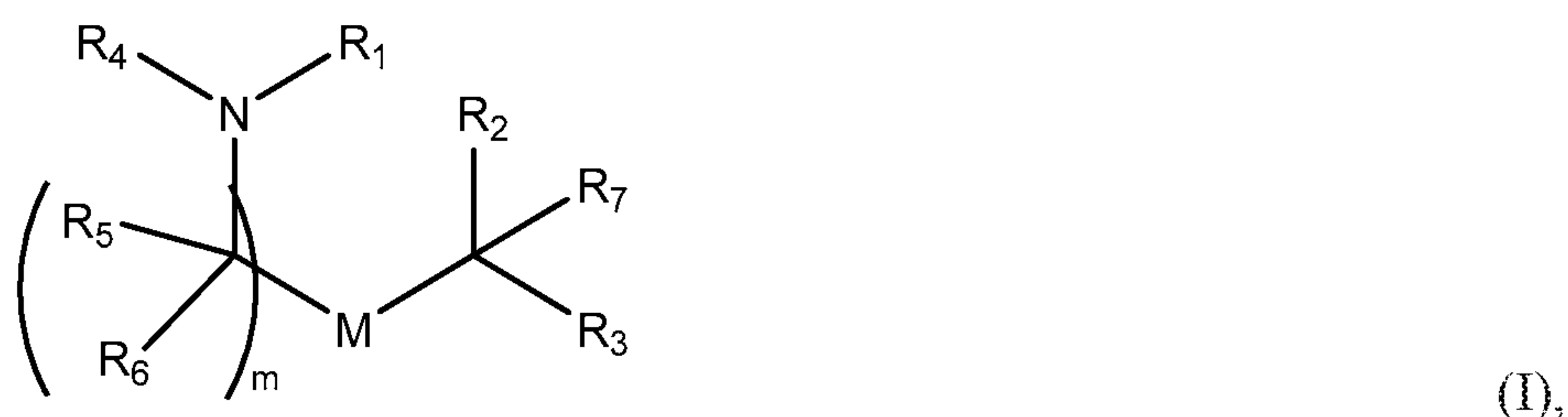
[0042] The disclosure is also directed to a method of treating methylmalonic acidemia in a subject, the method comprising administering to the subject an effective amount of a polynucleotide comprising an mRNA encoding an MCM polypeptide, wherein the administration alleviates the symptoms of methylmalonic acidemia in the subject. In some embodiments, the polynucleotide useful for the disclosure is any one of the polynucleotides encoding an MCM polypeptide described herein or is formulated as any one of the compositions described herein.

[0043] In some embodiments, the disclosure includes a method of reducing the level of a metabolite associated with methylmalonic acidemia in a subject in need thereof, the method comprising administering to the subject an effective amount of a polynucleotide comprising an mRNA encoding an MCM polypeptide. In some embodiments, the polynucleotide is a polynucleotide described elsewhere herein, or is formulated as a composition described herein. In certain embodiments, the polynucleotide reduces the level of methylmalonic acid present in the subject after the administration by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%. In other embodiments, after the administration, the polynucleotide reduces the level of propionyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%. In yet other embodiments, the polynucleotide reduces the level of acetyl-carnitine present in the subject after the administration by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%. In certain embodiments, one or more metabolites associated with methylmalonic acidemia are reduced after the administration within one day, within two days, within three days, within four days, within five days, within seven days, within one week, within two weeks, within three weeks, or within one month of the administration of the polynucleotide.

[0044] The details of various embodiments of the disclosure are set forth in the description below. Other features, objects, and advantages of the disclosure will be apparent from the description and the drawings, and from the claims.

EMBODIMENTS

[0045] E1. A composition comprising a polynucleotide that comprises an open reading frame (ORF) encoding an MCM polypeptide and a delivery agent, wherein the delivery agent comprises a compound having the formula (I)



or a salt or stereoisomer thereof, wherein

R_1 is selected from the group consisting of C_{5-20} alkyl, C_{5-20} alkenyl, $-R^*YR''$, $-YR''$, and $-R''M'R'$;

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, $-R^*YR''$, $-YR''$, and $-R^*OR''$, or R_2 and R_3 , together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a carbocycle, heterocycle,

$-OR$, $-O(CH_2)_nN(R)_2$, $-C(O)OR$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-N(R)_2$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, and $-C(R)N(R)_2C(O)OR$, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13; and

provided when R₄ is -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, or -CQ(R)₂, then (i) Q is not -N(R)₂ when n is 1, 2, 3, 4 or 5, or (ii) Q is not 5, 6, or 7-membered heterocycloalkyl when n is 1 or 2.

[0046] E2. The composition of embodiment 1, wherein the ORF has at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 1 to 207, 732 to 765, and 772.

[0047] E3. The composition of embodiment 1 or 2, wherein the ORF has

(i) at least 99% or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 734,

(ii) at least 98%, at least 99%, or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 732,

(iii) at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 182, 733, and 741;

(iv) at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 735, 736, 738, 743, 744, 748, 749, 750, 754, 755, 758, 762, and 765;

(v) at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 180, 187, 737, 739, 740, 742, 745, 746, 747, 751, 752, 753, 757, 759, 760, 761, 763, and 764;

(vi) at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 181 and 756;

(vii) at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175;

(viii) at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 164, 166, 167, 168, 169, 170, 172, 177, 178, 179, 195, and 204;

(ix) at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 156, 157, 158, 159, 160, 161, 162, 174 and 176;

(x) at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 155 and 203;

(xi) at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 64, 66, 71, 91, and 128;

(xii) at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 9, 11, 18, 19, 21, 22, 23, 24, 32, 33, 37, 39, 40, 44, 45, 47, 50, 51, 52, 55, 57, 61, 65, 70, 79, 84, 86, 88, 90, 92, 98, 100, 115, 117, 126, 129, 135, 136, 137, 144, 148, 150, 184, 190, 191, and 206;

(xiii) at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 3, 5, 6, 8, 10, 12, 14, 16, 17, 20, 27, 28, 29, 31, 34, 35, 36, 38, 41, 42, 43, 46, 48, 49, 53, 54, 56, 58, 60, 63, 68, 69, 74, 77, 78, 80, 83, 85, 87, 93, 95, 96, 97, 99, 102, 103, 104, 105, 107, 110, 112, 113, 114, 116, 119, 120, 122, 123, 124, 125, 127, 131, 132, 133, 134, 138, 139, 140, 141, 142, 143, 147, 149, 183, 186, 188, 189, 192, 193, 194, 196, 197, 198, 199, 200, 201, 202, 205, and 207; or

(xiv) at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 1, 2, 4, 7, 13, 15, 25, 26, 30, 59, 62, 67, 72, 73, 75, 76, 81, 82, 89, 94, 101, 106, 108, 109, 111, 118, 121, 130, 145, 146, and 185.

[0048] E4. The composition of any one of embodiments 1 to 3, wherein the ORF further comprises a nucleic acid sequence encoding a transit peptide.

[0049] E5. The composition of embodiment 4, wherein the transit peptide comprises a mitochondrial transit peptide.

[0050] E6. The composition of embodiment 5, wherein the mitochondrial transit peptide is derived from a protein selected from the group consisting of SEQ ID NOs: 251 to 265 and 270 to 719.

[0051] E7. The composition of any one of embodiments 4 to 6, wherein the nucleic acid sequence encoding a transit peptide has at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least

about 99%, or about 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to 96 of SEQ ID NOs: 1 to 207, 732 to 765, and 772.

[0052] E8. The composition of any one of embodiments 1 to 7, wherein the ORF has

(i) at least 99% or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 734;

(ii) at least 98%, at least 99%, or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 732;

(iii) at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 182 and 733;

(iv) at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 735, 741, 743, 744, 748, 758, 762, and 765;

(v) at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 180, 181, 736, 738, 739, 740, 742, 746, 747, 749, 750, 751, 752, 753, 754, 755, 757, 759, 760, 761, and 763;

(vi) at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 745, 756, and 764;

(vii) at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175;

(viii) at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to

nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 166, 167, 168, 169, 170, 172, 177, 178, 179, 187, and 204;

(ix) at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 156, 157, 158, 159, 160, 162, 164, 174, 176, 195, and 737;

(x) at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 155, 161, and 203;

(xi) at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 71 and 128;

(xii) at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 4, 6, 8, 9, 11, 19, 22, 23, 24, 32, 33, 37, 40, 44, 45, 47, 51, 61, 64, 65, 66, 79, 84, 86, 90, 91, 92, 100, 101, 112, 115, 117, 126, 129, 135, 136, 146, 148, 184, 190, and 191;

(xiii) at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 2, 3, 5, 7, 10, 12, 13, 14, 15, 16, 18, 20, 21, 26, 27, 28, 29, 31, 34, 36, 38, 39, 41, 42, 43, 46, 48, 49, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 68, 69, 70, 72, 73, 74, 76, 77, 80, 83, 85, 88, 95, 96, 97, 98, 102, 104, 105, 106, 107, 108, 109, 110, 113, 114, 120, 121, 122, 123, 124, 127, 131, 132, 133, 134, 137, 138, 139, 140, 141, 142, 144, 145, 147, 149, 150, 186, 188, 189, 192, 193, 194, 196, 198, 199, 200, 202, 205, 206, and 207; or

(xiv) at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 1, 17, 25, 30, 35, 50, 63, 67, 75, 78, 81, 82, 87, 89, 93, 94, 99, 103, 111, 116, 118, 119, 125, 130, 143, 183, 185, 197, and 201.

[0053] E9. The composition of any one of embodiments 1 to 8, wherein the MCM polypeptide comprises an amino acid sequence at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identical to SEQ ID NO: 208, and wherein the MCM polypeptide retains methylmalonyl-CoA mutase activity.

[0054] E10. The composition of embodiment 9, wherein the MCM polypeptide comprises SEQ ID NO: 209.

[0055] E11. The composition of embodiment 9, wherein the MCM polypeptide comprises SEQ ID NO: 210.

[0056] E12. The composition of embodiment 9, wherein the MCM polypeptide comprises SEQ ID NO: 211.

[0057] E13. The composition of embodiment 9, wherein the MCM polypeptide comprises SEQ ID NO: 212.

[0058] E14. The composition of embodiment 9, wherein the MCM polypeptide comprises SEQ ID NO: 213.

[0059] E15. The composition of any one of embodiments 1-14, wherein the polynucleotide comprises at least one chemically modified nucleobase, sugar, backbone, or any combination thereof.

[0060] E16. The composition of embodiment 15, wherein the at least one chemically modified nucleobase is selected from the group consisting of pseudouracil (ψ), N1-methylpseudouracil (m1 ψ), 2-thiouracil (s2U), 4'-thiouracil, 5-methylcytosine, 5-methyluracil, and any combination thereof.

[0061] E17. The composition of embodiment 16, wherein the at least one chemically modified nucleoside is 5-methoxyuracil.

[0062] E18. The composition of any one of embodiments 1-17, wherein the nucleosides in the polynucleotide sequence are chemically modified by at least 10%, at least

20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0063] E19. The composition of any one of embodiments 15-18, wherein the chemically modified nucleosides in the polynucleotide sequence are selected from the group consisting of uridine, adenine, cytosine, guanine, and any combination thereof.

[0064] E20. The composition of any one of embodiments 1-19, wherein the uridine nucleosides in the polynucleotide sequence are chemically modified by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0065] E21. The composition of any one of embodiments 1-20, wherein the adenine nucleosides in the polynucleotide sequence are chemically modified by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0066] E22. The composition of any one of embodiments 1-21, wherein the cytosine nucleosides in the polynucleotide sequence are chemically modified by at least at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0067] E23. The composition of any one of embodiments 1-22, wherein the guanine nucleosides in the polynucleotide sequence are chemically modified by at least at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0068] E24. The composition of any one of embodiments 1-23, wherein the polynucleotide further comprises a 5' UTR.

[0069] E25. The composition of embodiment 24, wherein the 5' UTR comprises a nucleic acid sequence at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a sequence selected from SEQ ID NOs: 215-231, 266, and 725-731.

[0070] E26. The composition of any one of embodiments 1 to 25, wherein the polynucleotide further comprises a miRNA binding site.

[0071] E27. The composition of embodiment 26, wherein the miRNA binding site comprises one or more polynucleotide sequences selected SEQ ID NOs: 720, 721, 722, 723, and 724.

[0072] E28. The composition of embodiment 26, wherein the miRNA binding site binds to miR-142 or miR-126.

[0073] E29. The composition of embodiment 26, wherein the miRNA binding site binds to miR-142-3p, miR-142-5p, miR-126-3p, or miR-126-5p.

[0074] E30. The composition of embodiment 24, wherein the 5'UTR comprises a sequence selected from SEQ ID NOs: 725, 726, 727, 728, 729, 730, and 731.

[0075] E31. The composition of any one of embodiments 24-30, wherein the 5' UTR is sequence optimized.

[0076] E32. The composition of any one of embodiments 1-31, wherein the polynucleotide further comprises a 3' UTR.

[0077] E33. The composition of embodiment 32, wherein the 3' UTR comprises a nucleic acid sequence at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a sequence selected from SEQ ID NO: 232-248 and 267.

[0078] E34. The composition of embodiment 32 or 33, wherein the 3' UTR is codon optimized.

[0079] E35. The composition of any one of embodiments 1-34, wherein the polynucleotide further comprises a 5' terminal cap.

[0080] E36. The composition of embodiment 35, wherein the 5' terminal cap is a Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azidoguanosine, Cap2, Cap4, 5' methylG cap, or an analog thereof.

[0081] E37. The composition of any one of embodiments 1-36, wherein the polynucleotide further comprises a 3' polyA tail.

[0082] E38. The composition of any one of embodiments 1-37, wherein the polynucleotide is RNA.

[0083] E39. The composition of embodiment 38, wherein the RNA is mRNA.

[0084] E40. The composition of any one of embodiments 1-39, wherein the polynucleotide is *in vitro* transcribed (IVT).

[0085] E41. The composition of any one of embodiments 1-40, wherein the polynucleotide is chimeric.

[0086] E42. The composition of any one of embodiments 1-41, wherein the polynucleotide is circular.

[0087] E43. The composition of any one of embodiments 1-42, wherein the polynucleotide is purified by strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), liquid chromatography-mass spectrometry (LCMS), capillary electrophoresis (CE) and capillary gel electrophoresis (CGE).

[0088] E44. The composition of any one of embodiments 1-43, wherein the compound is of Formula (IA):



or a salt or stereoisomer thereof, wherein

l is selected from 1, 2, 3, 4, and 5;

m is selected from 5, 6, 7, 8, and 9;

M₁ is a bond or M';

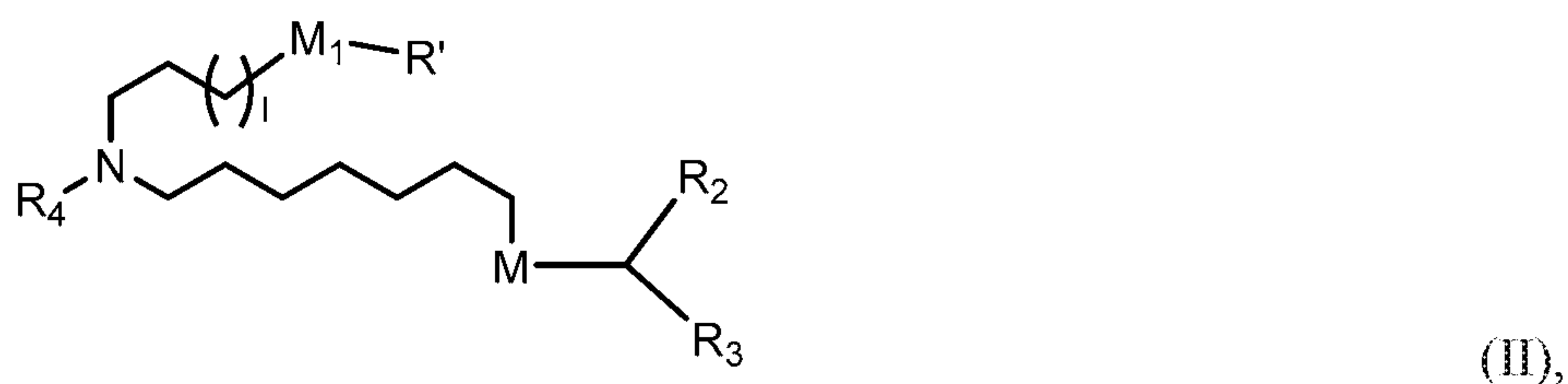
R₄ is unsubstituted C₁₋₃ alkyl, or -(CH₂)_nQ, in which n is 1, 2, 3, 4, or 5 and Q is OH, -NHC(S)N(R)₂, or -NHC(O)N(R)₂;

M and M' are independently selected from C(O)O, OC(O), C(O)N(R'), P(O)(OR')O, an aryl group, and a heteroaryl group; and

R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl.

[0089] E45. The composition of any one of embodiments 1 to 44, wherein m is 5, 7, or 9.

[0090] E46. The composition of any one of embodiments 1 to 45, wherein the compound is of Formula (II):



or a salt or stereoisomer thereof, wherein

l is selected from 1, 2, 3, 4, and 5;

M_1 is a bond or M' ;

R_4 is unsubstituted C_{1-3} alkyl, or $-(CH_2)_nQ$, in which n is 2, 3, or 4 and Q is OH, $-NHC(S)N(R)_2$, or $-NHC(O)N(R)_2$;

M and M' are independently selected from $-C(O)O$, $OC(O)$, $C(O)N(R')$, $P(O)(OR')O$, an aryl group, and a heteroaryl group; and

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, and C_{2-14} alkenyl.

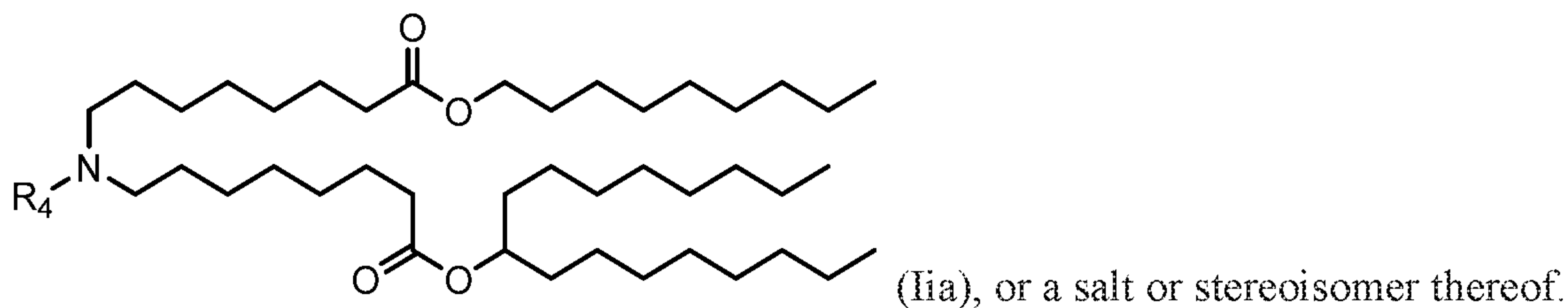
[0091] E47. The composition of any one of embodiments 44 to 46, wherein M_1 is M' .

[0092] E48. The composition of embodiment 47, wherein M and M' are independently $-C(O)O-$ or $-OC(O)-$.

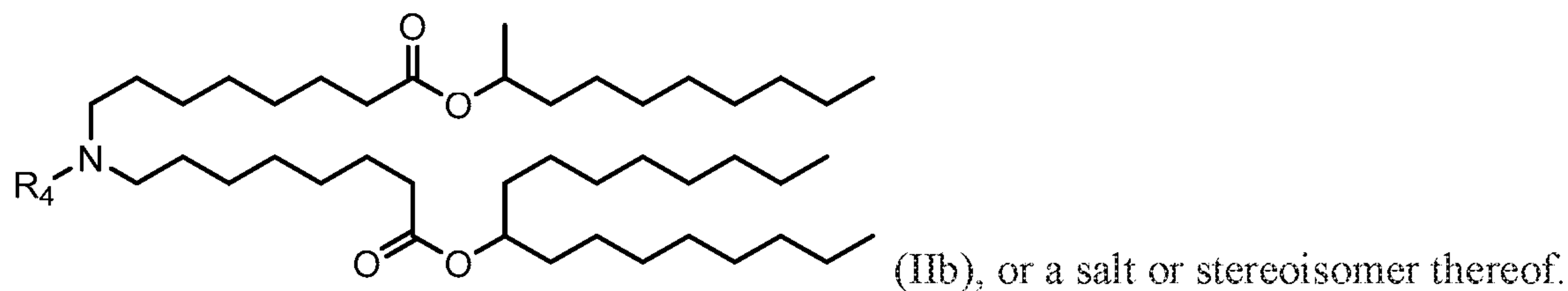
[0093] E49. The composition of any one of embodiments 44 to 48, wherein l is 1, 3, or 5.

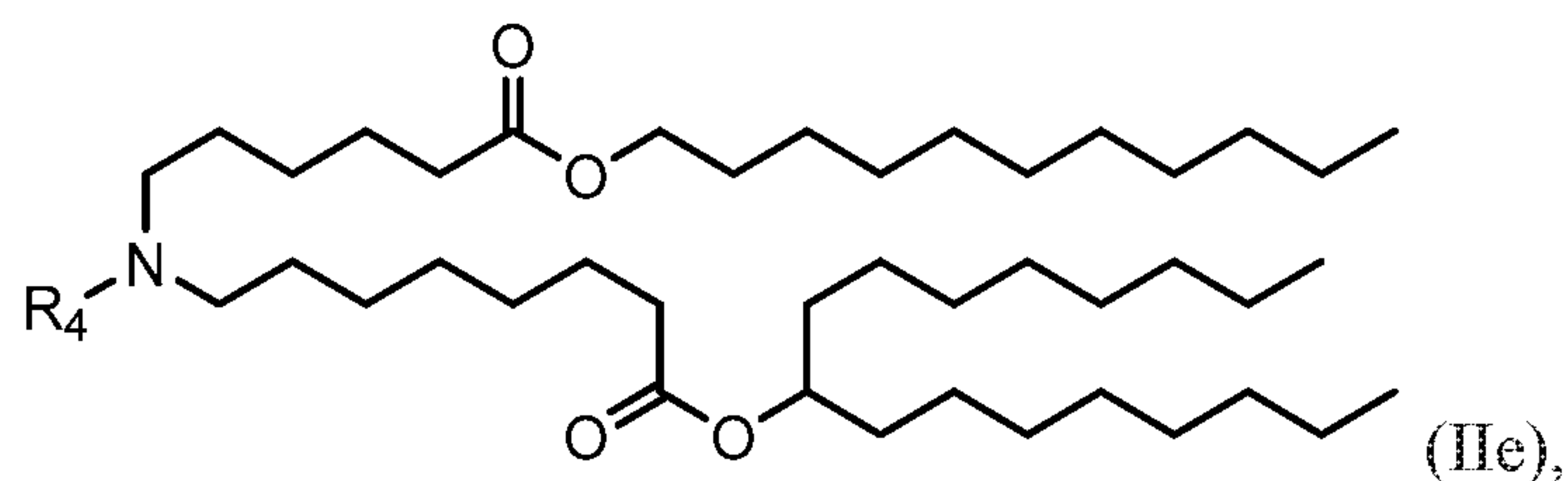
[0094] E50. The composition of any one of embodiments 1 to 43, wherein the compound is selected from the group consisting of Compound 1 to Compound 147, salts and stereoisomers thereof, and any combination thereof.

[0095] E51. The composition of any one of embodiments 1 to 43, wherein the compound is of the Formula (IIa),

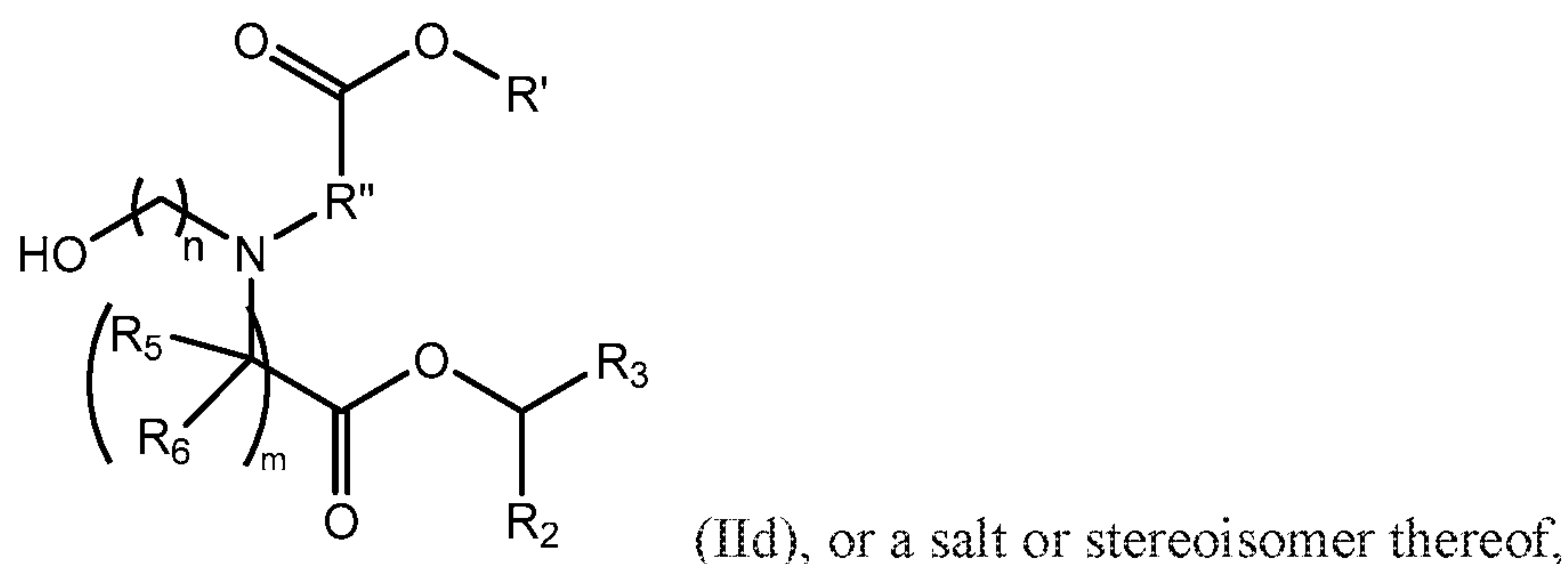


[0096] E52. The composition of any one of embodiments 1 to 43, wherein the compound is of the Formula (IIb),



CCCCCCCCCOC(=O)CCCCN(R4)CCCCC(=O)OCCCCCCCC (IIc)

[0099] E55. The composition of any one of embodiments 1 to 43, wherein the compound is of the Formula (IIId),



[0102] E58. The composition of any one of embodiments 55 to 57, wherein m is 5, 7, or 9.

- [0103] E59. The composition of any one of embodiments 55 to 58, wherein each R₅ is H.
- [0104] E60. The composition of embodiment 59, wherein each R₆ is H.
- [0105] E61. The composition of any one of embodiments 1 to 60, wherein the composition is a nanoparticle composition.
- [0106] E62. The composition of embodiment 61, wherein the delivery agent further comprises a phospholipid.
- [0107] E63. The composition of embodiment 62, wherein the phospholipid is selected from the group consisting of 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16:0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and any mixtures thereof.
- [0108] E64. The composition of any one of embodiments 1 to 63, wherein the delivery agent further comprises a structural lipid.

[0109] E65. The composition of embodiment 64, wherein the structural lipid is selected from the group consisting of cholesterol, fecosterol, sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, tomatidine, ursolic acid, alpha-tocopherol, and any mixtures thereof.

[0110] E66. The composition of any one of embodiments 1 to 65, wherein the delivery agent further comprises a PEG lipid.

[0111] E67. The composition of embodiment 66, wherein the PEG lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and any mixtures thereof.

[0112] E68. The composition of any one of embodiments 1 to 67, wherein the delivery agent further comprises an ionizable lipid selected from the group consisting of 3-(didodecylamino)-N1,N1,4-tridodecyl-1-piperazineethanamine (KL10), N1-[2-(didodecylamino)ethyl]-N1,N4,N4-tridodecyl-1,4-piperazinediethanamine (KL22), 14,25-ditridecyl-15,18,21,24-tetraaza-octatriacontane (KL25), 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLin-DMA), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (DLin-MC3-DMA), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-KC2-DMA), 1,2-dioleyloxy-N,N-dimethylaminopropane (DODMA), 2-((8-[(3 β)-cholest-5-en-3-yloxy]octyl)oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yl oxy]propan-1-amine (Octyl-CLinDMA), (2R)-2-((8-[(3 β)-cholest-5-en-3-yloxy]octyl)oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2R)), and (2S)-2-((8-[(3 β)-cholest-5-en-3-yloxy]octyl)oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2S)).

[0113] E69. The composition of any one of embodiments 1 to 68, wherein the delivery agent further comprises a phospholipid, a structural lipid, a PEG lipid, or any combination thereof.

[0114] E70. The composition of any one of embodiments 1-69, wherein the composition is formulated for *in vivo* delivery.

[0115] E71. The composition of embodiment 70 which is formulated for intramuscular, subcutaneous, or intradermal delivery.

[0116] E72. The composition of any one of embodiments 1-71 which increases cellular expression of MCM.

[0117] E73. The composition of embodiment 72, wherein the cellular expression of MCM is increased by at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%.

[0118] E74. A method of administering the composition of any one of embodiments 1-73, wherein the administration alleviates the symptoms of methylmalonic acidemia in the subject.

[0119] E75. The method of embodiment 74, wherein the administration results in a reduction of the level of a metabolite associated with methylmalonic acidemia in a subject in need thereof.

[0120] E76. The method of embodiment 74 or 75, further comprising measuring the level of the metabolite in the subject or in a sample obtained from the subject before and/or after the administering.

[0121] E77. The method of embodiment 76, wherein the sample is taken from the subject's blood, urine, cerebrospinal fluid, or any combination thereof.

[0122] E78. The method of any of embodiments 74 to 77, wherein the administration reduces the level of methylmalonic acid present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

[0123] E79. The method of any of embodiments 74 to 78, wherein the polynucleotide reduces the level of propionyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

[0124] E80. The method of any of embodiments 74 to 79, wherein the polynucleotide reduces the level of acetyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

[0125] E81. The method of any one of embodiments 74 to 80, wherein one or more metabolites associated with methylmalonic acidemia are reduced within one day, within two days, within three days, within four days, within five days, within seven days, within one week,

within two weeks, within three weeks, or within one month of the administration of the polynucleotide.

[0126] E82. A method of treating methylmalonic acidemia in a subject in need thereof, the method comprising administering to the subject an effective amount of a polynucleotide comprising an mRNA that comprises an ORF encoding an MCM polypeptide, wherein the administration alleviates the symptoms of methylmalonic acidemia in the subject.

[0127] E83. The method of embodiment 82, wherein the polynucleotide comprises the polynucleotide in the composition of any one of embodiments 1 to 73.

[0128] E84. The method of embodiment 82 or 83, wherein the administration reduces the level of a metabolite associated with methylmalonic acidemia in a subject in need thereof.

[0129] E85. The method of any one of embodiments 82-84, further comprising measuring the level of the metabolite in the subject or in a sample obtained from the subject before and/or after the administering.

[0130] E86. The method of embodiment 85, wherein the sample is taken from the subject's blood, urine, cerebrospinal fluid, or any combination thereof.

[0131] E87. The method of any of embodiments 82 to 86, wherein the polynucleotide reduces the level of methylmalonic acid present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

[0132] E88. The method of any of embodiments 82 to 87, wherein the polynucleotide reduces the level of propionyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

[0133] E89. The method of any of embodiments 82 to 88, wherein the polynucleotide reduces the level of acetyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

[0134] E90. The method of any one of embodiments 82 to 89, wherein one or more metabolites associated with methylmalonic acidemia are reduced within one day, within two days, within three days, within four days, within five days, within seven days, within one week, within two weeks, within three weeks, or within one month of the administration of the polynucleotide.

[0135] E91. The method of any one of embodiments 82 to 90, wherein the nucleotide is administered as a nanoparticle composition.

[0136] E92. The method of embodiment 91, wherein the composition further comprises a delivery agent.

[0137] E93. The method of embodiment 92, wherein the delivery agent comprises a phospholipid.

[0138] E94. The method of embodiment 93, wherein the phospholipid is selected from the group consisting of 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosaheptaenoyl-sn-glycero-3-phosphocholine, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16:0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosaheptaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and any mixtures thereof.

[0139] E95. The method of any one of embodiments 92 to 94, wherein the delivery agent further comprises a structural lipid.

[0140] E96. The method of embodiment 95, wherein the structural lipid is selected from the group consisting of cholesterol, fecosterol, sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, tomatidine, ursolic acid, alpha-tocopherol, and any mixtures thereof.

[0141] E97. The method of any one of embodiments 92 to 96, wherein the delivery agent further comprises a PEG lipid.

[0142] E98. The method of embodiment 97, wherein the PEG lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and any mixtures thereof.

[0143] E99. The method of any one of embodiments 92 to 98, wherein the delivery agent further comprises an ionizable lipid selected from the group consisting of 3-(didodecylamino)-N1,N1,4-tridodecyl-1-piperazineethanamine (KL10), N1-[2-(didodecylamino)ethyl]-N1,N4,N4-tridodecyl-1,4-piperazinediethanamine (KL22), 14,25-ditridecyl-15,18,21,24-tetraaza-octatriacontane (KL25), 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLin-DMA), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (DLin-MC3-DMA), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-KC2-DMA), 1,2-dioleyloxy-N,N-dimethylaminopropane (DODMA), 2-((8-[(3 β)-cholest-5-en-3-yloxy]octyl)oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yl oxy]propan-1-amine (Octyl-CLinDMA), (2R)-2-((8-[(3 β)-cholest-5-en-3-yloxy]octyl)oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2R)), and (2S)-2-((8-[(3 β)-cholest-5-en-3-yloxy]octyl)oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2S)).

[0144] E100. The method of any one of embodiments 92 to 99, wherein the delivery agent further comprises a phospholipid, a structural lipid, a PEG lipid, or any combination thereof.

[0145] E101. The method of any one of embodiments 92-100, wherein the composition is formulated for *in vivo* delivery.

[0146] E102. The method of embodiment 101 which is formulated for intramuscular, subcutaneous, or intradermal delivery.

[0147] E103. The method of any one of embodiments 82-102 which increases cellular expression of MCM.

[0148] E104. The method of embodiment 103, wherein the cellular expression of MCM is increased by at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%.

[0149] E105. The method of any one of embodiments 74 to 104, wherein the polynucleotide is administered at a dose of 0.1 mg/kg to 1.0 mg/kg, 0.1 mg/kg to 10 mg/kg, 0.1 mg/kg to 2 mg/kg, 0.1 mg/kg to 5 mg/kg, 1 mg/kg to 5 mg/kg, or 1 mg/kg to 3 mg/kg.

[0150] E106. The method of any one of embodiments 74 to 105, wherein the plasma MMA level after the administration is reduced at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% compared to the plasma MMA level prior to the administration.

[0151] E107. The method of embodiment 106, wherein the plasma MMA level is reduced about 75% to 85% compared to the plasma MMA level prior to the administration.

[0152] E108. The method of any one of embodiments 74 to 107, wherein the plasma MMA level after the administration is lower than about 5 $\mu\text{mol/L}$, about 4.5 $\mu\text{mol/L}$, about 4 $\mu\text{mol/L}$, about 3.5 $\mu\text{mol/L}$, about 3 $\mu\text{mol/L}$, about 2.5 $\mu\text{mol/L}$, about 2 $\mu\text{mol/L}$, about 1.5 $\mu\text{mol/L}$, about 1 $\mu\text{mol/L}$, about 0.9 $\mu\text{mol/L}$, about 0.8 $\mu\text{mol/L}$, about 0.7 $\mu\text{mol/L}$, about 0.6 $\mu\text{mol/L}$, about 0.5 $\mu\text{mol/L}$, about 0.4 $\mu\text{mol/L}$, about 0.3 $\mu\text{mol/L}$, or 0.27 $\mu\text{mol/L}$.

[0153] E109. The method of embodiments 74 to 108, wherein the urinary MMA level is less than 2000 mmol/mol creatinine, less than 1900 mmol/mol creatinine, less than 1800 mmol/mol creatinine, less than 1700 mmol/mol creatinine, less than 1600 mmol/mol creatinine, less than 1500 mmol/mol creatinine, less than 1400 mmol/mol creatinine, less than 1300 mmol/mol creatinine, less than 1200 mmol/mol creatinine, less than 1100 mmol/mol creatinine, less than 1000 mmol/mol creatinine, 900 mmol/mol creatinine, 800 mmol/mol creatinine, 700 mmol/mol creatinine, 600 mmol/mol creatinine, 500 mmol/mol creatinine, 400 mmol/mol creatinine, 300 mmol/mol creatinine, 200 mmol/mol creatinine, 100 mmol/mol creatinine, 90 mmol/mol creatinine, 80 mmol/mol creatinine, 70 mmol/mol creatinine, 60 mmol/mol creatinine, 50 mmol/mol creatinine, 40 mmol/mol creatinine, 30 mmol/mol creatinine, 20 mmol/mol creatinine, 10 mmol/mol creatinine, 9 mmol/mol creatinine, 8 mmol/mol creatinine, 7 mmol/mol creatinine, 6 mmol/mol creatinine, 5 mmol/mol creatinine, 4 mmol/mol creatinine, 3 mmol/mol creatinine, 2 mmol/mol creatinine, or 1 mmol/mol creatinine.

BRIEF DESCRIPTION OF THE DRAWINGS

[0154] **FIG. 1** is a Western blot analysis of endogenous methylmalonyl-CoA mutase expression in a mouse liver mitochondrial extract, mouse cells (Hepal-6), and human cells (HepG2, SNU423, and HeLa). The upper band (thin arrow) shows Mouse α -Methylmalonyl-CoA mutase, and the lower band (thick arrow) shows Rabbit α -Citrate synthetase.

[0155] **FIGS. 2A-2C** show immunofluorescence analyses of the localization of endogenous methylmalonyl-CoA mutase in HeLa cells. **FIG. 2A** shows the location of mitochondria using Mitotracker and the nucleus using DAPI. **FIG. 2B** shows the immunostaining of hMCM protein using a murine monoclonal anti-MCM antibody and the location of the nucleus using DAPI. **FIG. 2C** shows the merged picture of **FIGS. 2A** and **2B**.

[0156] **FIG. 3** is a Western blot analysis comparing methylmalonyl-CoA mutase expression in (i) HeLa cells transfected with a control GFP expression construct, (ii) HeLa cells transfected with a construct for expressing methylmalonyl-CoA mutase, and (iii) a mouse liver mitochondrial extract.

[0157] **FIG. 4** is a comparison of methylmalonyl-CoA mutase enzymatic activity in (i) HeLa cells transfected with a control GFP expression construct, (ii) HeLa cells transfected with a construct for expressing methylmalonyl-CoA mutase, and (iii) a mouse liver mitochondrial extract.

[0158] **FIG. 5** is a Western blot analysis of methylmalonyl-CoA mutase expression in Hepal-6 cells, fibroblasts from normal human subjects (NHDF), and fibroblasts from MMA patients (GM50 and GM1573) that were transfected with control mRNA, human MCM mRNA, or mouse MCM mRNA.

[0159] **FIGS. 6A-6D** show immunofluorescence analyses of the localization of exogenously expressed methylmalonyl-CoA mutase in human fibroblasts transfected with eGFP mRNAs or MCM mRNAs (also referred to as "MUT"). The left panels show the location of mitochondria using Mitotracker and the nucleus using DAPI, the middle panels show shows the location of mitochondria using MCM protein and the nucleus using DAPI, and the right panels show merged images of the left panel and the right panel. **FIGS. 6A** and **6C** are images taken of patient fibroblasts transfected with mRNA encoding eGFP. **FIGS. 6B** and **6D** are images taken of patient fibroblasts transfected with mRNA encoding hMCM.

[0160] FIG. 7 is measurement of methylmalonyl-CoA mutase activity in Hepal-6 cells, fibroblasts from normal human subjects (NHDF), and fibroblasts from MMA patients (GM50 and GM1573) that were transfected with control mRNA, human MCM mRNA, or mouse MCM mRNA.

[0161] FIGS. 8A-8B is an analysis of *in vivo* treatment with mRNA encoding methylmalonyl-CoA mutase. For FIGS. 8A and 8B, C57B/L6 mice were injected intravenously with either control mRNA (NT-FIX) or MCM mRNA at 0.5 mg mRNA/kg body weight ("mpk"). Mice were sacrificed after 24 or 48 hours and MCM protein in mitochondria from livers were determined by capillary electrophoresis (CE). The upper panel (FIG. 8A) shows injection of MCM mRNA increased MCM protein expression after 24 and 48 hours, while the lower panel (FIG. 8B) shows the expression of the control protein citrate synthase.

[0162] FIGS. 9-15 show exemplary codon optimized MCM sequences that encode methylmalonyl-CoA mutase. The illustrated sequences in FIGS. 9-15 are SEQ ID NOs: 732, 733, 734, 151, 152, 153, and 154, respectively.

[0163] FIGS. 16A-C show analysis of MMA levels and body weight in the MCK mouse model. FIG. 16A shows plasma levels of methylmalonic acid (MMA) in μ M measured by LC-MS/MS over time in mice treated weekly with control mRNA (NT-FIX) at 0.1 mg/kg, codon optimized MCM mRNA (encoding SEQ ID NO:734) formulated in lipid nanoparticles at 0.16 or 0.2 mg/kg for 5 injections, or codon optimized MCM mRNA (encoding SEQ ID NO:734) formulated in lipid nanoparticles at 0.2 mg/kg for 2 injections. FIG. 16B shows the body weight of the mice over time (measured twice a week). *** $p < 0.001$; P-values obtained from repeated measures ANOVA. FIG. 16C shows the increase in body weight over time in mice injected weekly with codon optimized MCM mRNA.

[0164] FIG. 17 shows MCM expression in liver of wild-type CD1 mice dosed with codon optimized MCM mRNA (SEQ ID NO: 734) formulated in lipid nanoparticles at 0.2 mg/kg compared to endogenous human MCM and endogenous mouse MCM.

[0165] FIGS. 18A-C show a time course of the effects of injection of codon optimized MCM mRNA. FIG. 18A shows levels of lipid nanoparticles after single dose injection of codon optimized MCM mRNA. FIG. 18B shows Hepatic hMut mRNA levels in mouse liver after single dose injection of codon optimized MCM mRNA. FIG. 18C shows MCM protein levels after single dose injection of codon optimized MCM mRNA.

[0166] FIGS. 19A-B show MCM expression in level of wild-type CD mice dosed with codon optimized MCM mRNAs, where the mRNA is formulated either with MC3 or Compound 18. NTFIX mRNA was used as a control. FIG. 19A shows a Western blot of expression after dosing with different formulations, and FIG. 19B shows a quantification of that Western blot.

[0167] FIGS. 20A-B show the effects of administering codon optimized MCM mRNA to mice on the plasma levels of MMA (FIG. 20A) and the body weight of the mice (FIG. 20B).

DETAILED DESCRIPTION

[0168] The present disclosure provides polynucleotide sequences that encode a sequence-optimized nucleic acid encoding a methylmalonyl-CoA mutase polypeptide ("MCM" or "MUT"). MCM is the principal gene associated with methylmalonic acidemia ("MMA," also referred to as methylmalonic aciduria). Wild type nucleic acid and amino acid sequences for human methylmalonyl-CoA mutase (MCM) are described in NCBI sequence records gi296010795 (reference sequence NM_000255.3, "Homo sapiens methylmalonyl-CoA mutase (MUT), mRNA"; *see also*, SEQ ID NO: 214) and gi156105689 (reference sequence NP_000246.2, "methylmalonyl-CoA mutase, mitochondrial precursor [Homo sapiens]"; *see also*, SEQ ID NO: 208), respectively. Accession numbers and the associated sequences are found at the National Center for Biotechnology Information (NCBI) website.

[0169] MCM is a metabolic enzyme (E.C. 5.4.99.2), the biological function of which is to isomerize L-methylmalonyl-CoA into succinyl-CoA, a Krebs cycle intermediate. MCM localizes to the mitochondria of cells, exists as a homodimer in its native form, and is adenosylcobalamin-dependent. The precursor form of human MCM is 750 amino acids, while its mature form is 718 amino acids – a 32 amino acid leader sequence is cleaved off by mitochondrial importation and processing machinery.

I. Composition

Polynucleotides Encoding MCM

[0170] In certain aspects, the present disclosure provides nucleic acid molecules, specifically polynucleotides that encode one or more MCM polypeptides. The MCM polypeptides that are encoded can be mammalian MCM polypeptides, for example, human MCM peptides, or functional fragments thereof.

[0171] In some embodiments, the polynucleotides described herein encode at least one methylmalonyl-CoA mutase protein, functional fragment, or variant thereof. MCM catalyzes enzymatic transformation of methylmalonyl-CoA into succinyl-CoA, and also comprises a cobalamin-binding domain. MCM's enzymatic activity is dependent on its binding to its cofactor, deosylcobalamin.

[0172] MCM plays a critical role in the catabolism of fat and protein, specifically in disposing of methylmalonyl-CoA created during metabolism. For example, methylmalonyl-CoA is an intermediate in the catabolism of amino acids such as isoleucine, methionine, and threonine. Methylmalonyl-CoA is also an intermediate in the catabolism of cholesterol and fatty acids. Defects in the activity of this enzyme lead to inefficient metabolism and buildup of potentially toxic metabolic intermediates such as methylmalonic acid. The lack of MCM causes the disorder known as methylmalonic acidemia (MMA).

[0173] Replacement of MCM has been theorized to be a cure of this form of MMA. In some embodiments, the polynucleotides disclosed herein comprise one or more sequences encoding a methylmalonyl-CoA mutase protein, functional fragment, or variant thereof that is suitable for use in such gene replacement therapy. In certain aspects, the present application addresses the problem of the lack of methylmalonyl-CoA mutase by providing a polynucleotide, *e.g.*, mRNA, that encodes methylmalonyl-CoA mutase or functional fragment thereof, wherein the polynucleotide is sequence-optimized. In some embodiments, the polynucleotide, *e.g.*, mRNA, increases MCM expression levels in cells when introduced into those cells, *e.g.*, by at least 20%, at least 20%, at least 25%, at least 35%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 100%.

[0174] In some embodiments, the polynucleotides of the disclosure encode functional MCM polypeptides or fragments thereof. In some embodiments, the polynucleotides of the disclosure encode an MCM protein or variant thereof that is full length (*i.e.*, it includes a mitochondrial transit peptide, either native or heterologous to that in native full-length MCM), while in other embodiments polynucleotides of the disclosure encode a functional MCM protein or variant thereof that is mature (*i.e.*, it lacks the mitochondrial transit peptide). In some embodiments, the polynucleotides encode a human MCM, or variant thereof, linked to a heterologous or homologous mitochondrial transit peptide.

[0175] In some embodiments, the polynucleotides of the disclosure encode functional human MCM (SEQ ID NO: 208, SEQ ID NO: 209, SEQ ID NO: 210, SEQ ID NO: 211, SEQ ID NO:

212, and SEQ ID NO: 213) or fragments thereof. In some embodiments, the polynucleotides of the disclosure encode mutant MCM. In some embodiments, the polynucleotides encode an MCM polypeptide that comprises at least one point mutation in the MCM sequence, while still retaining MCM enzymatic activity. In some embodiments, the polynucleotides encode a functional MCM polypeptide with mutations that do not alter the function of MCM. Such functional MCM can be referred to as function-neutral. In some embodiments, the encoded MCM polypeptide comprises one or more of the function-neutral point mutations V69, T499, H532, A598, and V671. In some embodiments, the polynucleotides of the disclosure encode the polypeptide sequences in SEQ ID NO: 209, SEQ ID NO: 210, SEQ ID NO: 211, SEQ ID NO: 212, and SEQ ID NO: 213, which contain the function-neutral mutants V69, T499, H532, A598, and V671, respectively. In particular embodiments, the encoded MCM polypeptide is a V671 mutant (SEQ ID NO: 213). Polynucleotides encoding MCM polypeptides are listed in SEQ ID NOs: 1 to 207, 214, 732 to 765, and 772.

[0176] In some embodiments, the polynucleotides comprise a nucleotide sequence having significant sequence similarity to a polynucleotide selected from the group of SEQ ID NOs: 1-207, 732-765, and 772, wherein the ORF encodes an MCM polypeptide. In some embodiments, the polynucleotide comprises a nucleotide sequence having significant sequence similarity to SEQ ID NOs: 151, 152, 153, 154, 732, 733, and 734 (FIGS. 9-15). In some embodiments, the polynucleotide comprises a nucleotide sequence having significant sequence similarity to SEQ ID NO: 734 (FIG. 11).

[0177] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 99% or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 734, wherein the ORF encodes an MCM polypeptide.

[0178] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 98%, at least 99%, or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 732, wherein the ORF encodes an MCM polypeptide.

[0179] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 182, 733, and 741, wherein the ORF encodes an MCM polypeptide.

[0180] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 735, 736, 738, 743, 744, 748, 749, 750, 754, 755, 758, 762, and 765, wherein the ORF encodes an MCM polypeptide.

[0181] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 180, 187, 737, 739, 740, 742, 745, 746, 747, 751, 752, 753, 757, 759, 760, 761, 763, and 764, wherein the ORF encodes an MCM polypeptide.

[0182] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 181 and 756, wherein the ORF encodes an MCM polypeptide.

[0183] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175, wherein the ORF encodes an MCM polypeptide.

[0184] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 164, 166, 167, 168, 169, 170, 172, 177, 178, 179, 195, and 204, wherein the ORF encodes an MCM polypeptide.

[0185] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of

nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 156, 157, 158, 159, 160, 161, 162, 174 and 176, wherein the ORF encodes an MCM polypeptide.

[0186] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 155 and 203, wherein the ORF encodes an MCM polypeptide.

[0187] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 64, 66, 71, 91, and 128, wherein the ORF encodes an MCM polypeptide.

[0188] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 9, 11, 18, 19, 21, 22, 23, 24, 32, 33, 37, 39, 40, 44, 45, 47, 50, 51, 52, 55, 57, 61, 65, 70, 79, 84, 86, 88, 90, 92, 98, 100, 115, 117, 126, 129, 135, 136, 137, 144, 148, 150, 184, 190, 191, and 206, wherein the ORF encodes an MCM polypeptide.

[0189] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 3, 5, 6, 8, 10, 12, 14, 16, 17, 20, 27, 28, 29, 31, 34, 35, 36, 38, 41, 42, 43, 46, 48, 49, 53, 54, 56, 58, 60, 63, 68, 69, 74, 77, 78, 80, 83, 85, 87, 93, 95, 96, 97, 99, 102, 103, 104, 105, 107, 110, 112, 113, 114, 116, 119, 120, 122, 123, 124, 125, 127, 131, 132, 133, 134, 138, 139, 140, 141, 142, 143, 147, 149, 183, 186, 188, 189, 192, 193, 194, 196, 197, 198, 199, 200, 201, 202, 205, and 207, wherein the ORF encodes an MCM polypeptide.

[0190] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 1, 2, 4, 7, 13, 15, 25, 26, 30, 59, 62, 67, 72, 73, 75, 76, 81, 82, 89, 94, 101, 106, 108, 109, 111, 118, 121, 130, 145, 146, and 185, wherein the ORF encodes an MCM polypeptide.

[0191] The polynucleotides of the disclosure can also encode additional features that facilitate trafficking of the polypeptides to therapeutically relevant sites. One such feature that aids in protein trafficking is the signal sequence, or targeting sequence. The peptides encoded by these signal sequences are known by a variety of names, including targeting peptides, transit peptides, and signal peptides. The disclosure also includes a polynucleotide comprising a sequence that encodes a mitochondrial transit peptide operably linked to the polynucleotide described herein, *i.e.*, linked to a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0192] As used herein, a "signal sequence" or "signal peptide" is a polynucleotide or polypeptide, respectively, which is from about 9 to 200 nucleotides (3-70 amino acids) in length that, in some embodiments, is incorporated at the 5' (or N-terminus) of the coding region or polypeptide encoded, respectively. Addition of these sequences result in trafficking of the encoded polypeptide to a desired site, such as the endoplasmic reticulum or the mitochondria through one or more targeting pathways. Some signal peptides are cleaved from the protein, for example by a signal peptidase after the proteins are transported.

[0193] For example, human MCM's precursor protein comprises a 32-amino acid mitochondrial transit peptide, also referred to as an MCM mitochondrial targeting sequence or mitochondrial targeting peptide, that facilitates delivery of the MCM protein to, and localization in, mitochondria. The present disclosure comprises both polynucleotides that encode a homologous targeting sequence (*i.e.*, MCM's native mitochondrial transit sequence) and polynucleotides that encode a heterologous mitochondrial transit sequences (*i.e.*, a mitochondrial transit peptide that is not the native targeting peptide for the operably linked MCM protein). In some embodiments, the alternate targeting sequences facilitate delivery of MCM to mitochondria.

[0194] Exemplary sequences of known mitochondrial transit peptides include MLSLRQSIRFFKPATRTLCSRYLL (SEQ ID NO: 251), MALLRGVVFVAAKRTP (SEQ ID NO: 252) and MLRIPVRKALVGLSKSSKGCVRT (SEQ ID NO: 253). Non-limiting examples of the mitochondrial transit peptides are listed below in **Table 1** (SEQ ID NOs: 251-265). Further examples of mitochondrial transit peptides are provided as SEQ ID NOs: 270-719. Additional mitochondrial transit peptides that can be utilized in the present disclosure can be identified using predictive tools known in the art. For example, mitochondrial targeting can be analyzed using the methods described in Fukusawa *et al.*, Molecular and Cellular Proteomics 14:1113-1126 (2015), the contents of which are incorporated herein in their entirety.

Table 1. Mitochondrial Transit Peptides

ID (SEQ ID NO)	Name of the protein	Sequence
COX4 (SEQ ID NO: 251)	Saccharomyces cerevisiae mitochondrial cytochrome c oxidase subunit IV	MLSLRQSIRFFKPATRTLCSRYLL
ACAA2 (SEQ ID NO: 252)	Mitochondrial 3-ketoacyl-coa thiolase	MALLRGVVFVAAKRTP
NDUFS1 (SEQ ID NO: 253)	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial isoform 2	MLRIPVRKALVGLSKSSKGCVRT
A6NK58 (SEQ ID NO: 254)	Putative lipoyltransferase 2, mitochondrial (EC 2.3.1.181) (Lipoate-protein ligase B) (Lipoyl/octanoyl transferase) (Octanoyl-[acyl-carrier-protein]-protein N-octanoyltransferase)	MRQPAVRLVRLGRVPYAELLGLQDRWLR RLQ
A8K5M9 (SEQ ID NO: 255)	Uncharacterized protein C15orf62, mitochondrial	METWRKGSFRN
A8MUP2 (SEQ ID NO: 256)	Methyltransferase-like protein 12, mitochondrial (EC 2.1.1.-)	MAALRRMLHLPSLMMGTCRPFAGSLADS
O00142 (SEQ ID NO: 257)	Thymidine kinase 2, mitochondrial (EC 2.7.1.21) (Mt-TK)	MLLWPLRGWAARALRCFGPGSRGSPASG PGPRR
O00217 (SEQ ID NO: 258)	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-23kD) (CI-23kD) (NADH-ubiquinone oxidoreductase 23 kDa subunit) (TYKY subunit)	MRCLTTFMLLRALAQAARAGPPGGRSLH SSAVAA
O00330 (SEQ ID NO: 259)	Pyruvate dehydrogenase protein X component, mitochondrial (Dihydrolipoamide	MAASWRLGCDPRLRLRYLVGFPGRRSVGL VKGALGWSVSRGANWRWFHSTQWLR

259)	dehydrogenase-binding protein of pyruvate dehydrogenase complex) (E3-binding protein) (E3BP) (Lipoyl-containing pyruvate dehydrogenase complex component X) (proX)	
O00411 (SEQ ID NO: 260)	DNA-directed RNA polymerase, mitochondrial (MtRPOL) (EC 2.7.7.6)	MSALCWGRGAAGLKRALRPCGRPGLPGK EGTAGGVCGPRRS
O00746 (SEQ ID NO: 261)	Nucleoside diphosphate kinase, mitochondrial (NDK) (NDP kinase, mitochondrial) (EC 2.7.4.6) (Nucleoside diphosphate kinase D) (NDPKD) (nm23-H4)	MGGLEWRSALRGLRCGPRAPGPSLLVRH GSGGP
O14521 (SEQ ID NO: 262)	Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial (CybS) (CII-4) (QPs3) (Succinate dehydrogenase complex subunit D) (Succinate- ubiquinone oxidoreductase cytochrome b small subunit) (Succinate-ubiquinone reductase membrane anchor subunit)	MAVLWRLSAVCGALGGRALLLRTPVVRP AHI SAFLQDRPI PEWCGVQHIHLSPSHH
O14548 (SEQ ID NO: 263)	Cytochrome c oxidase subunit 7A-related protein, mitochondrial (COX7a-related protein) (Cytochrome c oxidase subunit VIIa- related protein) (EB1)	MYYKFSGFTQKLAGAWASEAYSPQGLKP VVSTEAPPIIFATPTKLTSDSTVYDYA
mMCM (SEQ ID NO:264)	Mouse methylmalonyl-CoA mutase	MLRAKNQLFLLSPHYLKQLNIPASRWK RL
hMCM (SEQ ID NO:265)	Human methylmalonyl-CoA mutase	MLRAKNQLFLLSPHYLRQVKESGSRLLI QQRL

[0195] In some embodiments, the nucleic acid sequence encoding a mitochondrial transit peptide has at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% sequence identity to a sequence selected from the group of nucleotides 1 to 96 of SEQ ID NOs: 1-207, 732-765, and 772, wherein the transit peptide is capable of targeting or carrying the MCM polypeptide into the mitochondria.

[0196] In some embodiments, the nucleic acid sequence encoding a mitochondrial transit peptide has at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% sequence identity to a sequence in **Table 1** (SEQ ID NOs: 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, or 265), wherein the transit peptide is capable of targeting or carrying the MCM polypeptide into the mitochondria. In some embodiments, the nucleic acid

sequence encoding a mitochondrial transit peptide has at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 270 to 719, wherein the transit peptide is capable of targeting or carrying the MCM polypeptide into the mitochondria.

[0197] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 99% or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 734, wherein the ORF encodes an MCM polypeptide.

[0198] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 98%, at least 99%, or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 732, wherein the ORF encodes an MCM polypeptide.

[0199] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 182 and 733, wherein the ORF encodes an MCM polypeptide.

[0200] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 735, 741, 743, 744, 748, 758, 762, and 765, wherein the ORF encodes an MCM polypeptide.

[0201] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 180, 181, 736, 738, 739, 740, 742, 746, 747, 749, 750, 751, 752, 753, 754, 755, 757, 759, 760, 761, and 763, wherein the ORF encodes an MCM polypeptide.

[0202] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 745, 756, and 764, wherein the ORF encodes an MCM polypeptide.

[0203] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175, wherein the ORF encodes an MCM polypeptide.

[0204] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 166, 167, 168, 169, 170, 172, 177, 178, 179, 187, and 204, wherein the ORF encodes an MCM polypeptide.

[0205] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 156, 157, 158, 159, 160, 162, 164, 174, 176, 195, and 737, wherein the ORF encodes an MCM polypeptide.

[0206] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 155, 161, and 203, wherein the ORF encodes an MCM polypeptide.

[0207] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 71 and 128, wherein the ORF encodes an MCM polypeptide.

[0208] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at

least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 4, 6, 8, 9, 11, 19, 22, 23, 24, 32, 33, 37, 40, 44, 45, 47, 51, 61, 64, 65, 66, 79, 84, 86, 90, 91, 92, 100, 101, 112, 115, 117, 126, 129, 135, 136, 146, 148, 184, 190, and 191, wherein the ORF encodes an MCM polypeptide.

[0209] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 2, 3, 5, 7, 10, 12, 13, 14, 15, 16, 18, 20, 21, 26, 27, 28, 29, 31, 34, 36, 38, 39, 41, 42, 43, 46, 48, 49, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 68, 69, 70, 72, 73, 74, 76, 77, 80, 83, 85, 88, 95, 96, 97, 98, 102, 104, 105, 106, 107, 108, 109, 110, 113, 114, 120, 121, 122, 123, 124, 127, 131, 132, 133, 134, 137, 138, 139, 140, 141, 142, 144, 145, 147, 149, 150, 186, 188, 189, 192, 193, 194, 196, 198, 199, 200, 202, 205, 206, and 207, wherein the ORF encodes an MCM polypeptide.

[0210] In some embodiments, the disclosure is directed to a polynucleotide comprising an ORF having at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 1, 17, 25, 30, 35, 50, 63, 67, 75, 78, 81, 82, 87, 89, 93, 94, 99, 103, 111, 116, 118, 119, 125, 130, 143, 183, 185, 197, and 201, wherein the ORF encodes an MCM polypeptide.

[0211] In some embodiments, the polynucleotide includes from about 1500 to about 100,000 nucleotides (*e.g.*, from about 1500 to 2500, from about 1800 to about 2600, from about 1900 to about 2600, from about 2000 to about 2700, from 2154 to 2,750, from 2154 to 3,000, from 2154 to 5,000, from 2154 to 7,000, from 2154 to 10,000, from 2154 to 25,000, from 2154 to 50,000, from 2154 to 70,000, from 2154 to 100,000, from 2250 to 2750, from 2250 to 3,000, from 2250 to 5,000, from 2250 to 7,000, from 2250 to 10,000, from 2250 to 25,000, from 2250 to 50,000, from 2250 to 70,000, and from 2250 to 100,000 nucleotides).

[0212] In some embodiments, the polynucleotides of the present disclosure can further comprise at least one nucleic acid sequence that is non-coding.

[0213] In some embodiments, the length of a region encoding at least one polypeptide of interest is greater than about 2154 nucleotides in length (*e.g.*, at least or greater than about 2154, 2,250, 2,500, 3,000, 4,000, 4,100, 4,200, 4,300, 4,400, 4,500, 4,600, 4,700, 4,800, 4,900, 5,000, 5,100, 5,200, 5,300, 5,400, 5,500, 5,600, 5,700, 5,800, 5,900, 6,000, 7,000, 8,000, 9,000, 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000 or up to and including 100,000 nucleotides). As used herein, such a region can be referred to as a "coding region" or "region encoding."

[0214] In some embodiments, the polynucleotides of the present disclosure are, or function as, a messenger RNA (mRNA). As used herein, the term "messenger RNA" (mRNA) refers to any polynucleotide that encodes at least one polypeptide of interest and that is capable of being translated to produce the encoded polypeptide of interest *in vitro*, *in vivo*, *in situ* or *ex vivo*. Exemplary mRNAs that can be used are listed in SEQ ID NOs: 776-771.

Optimized Polynucleotides Encoding MCM

[0215] The polynucleotides of the disclosure, their regions or parts or subregions are sequence-optimized. Sequence optimization methods are known in the art and can be useful to achieve one or more desired results. These results include to match codon frequencies in target and host organisms to ensure proper folding, bias GC content to increase mRNA stability or reduce secondary structures, minimize tandem repeat codons or base runs that can impair gene construction or expression, customize transcriptional and translational control regions, insert or remove protein trafficking sequences, remove/add post translation modification sites in encoded protein (*e.g.*, glycosylation sites), add, remove or shuffle protein domains, insert or delete restriction sites, modify ribosome binding sites and mRNA degradation sites, to adjust translational rates to allow the various domains of the protein to fold properly, or to reduce or eliminate problem secondary structures within the polynucleotide. Sequence optimization tools, algorithms and services are known in the art, non-limiting examples include services from GeneArt (Life Technologies), DNA2.0 (Menlo Park CA) and/or proprietary methods. In some embodiments, the ORF sequence is optimized using optimization algorithms. Codon options for each amino acid are given in Table 2.

Table 2. Codon Options

Amino Acid	Single Letter Code	Codon Options
Isoleucine	I	ATT, ATC, ATA
Leucine	L	CTT, CTC, CTA, CTG, TTA, TTG
Valine	V	GTT, GTC, GTA, GTG
Phenylalanine	F	TTT, TTC
Methionine	M	ATG
Cysteine	C	TGT, TGC
Alanine	A	GCT, GCC, GCA, GCG
Glycine	G	GGT, GGC, GGA, GGG
Proline	P	CCT, CCC, CCA, CCG
Threonine	T	ACT, ACC, ACA, ACG
Serine	S	TCT, TCC, TCA, TCG, AGT, AGC
Tyrosine	Y	TAT, TAC
Tryptophan	W	TGG
Glutamine	Q	CAA, CAG
Asparagine	N	AAT, AAC
Histidine	H	CAT, CAC
Glutamic acid	E	GAA, GAG
Aspartic acid	D	GAT, GAC
Lysine	K	AAA, AAG
Arginine	R	CGT, CGC, CGA, CGG, AGA, AGG
Selenocysteine	Sec	UGA in mRNA in presence of Selenocysteine insertion element (SECIS)
Stop codons	Stop	TAA, TAG, TGA

[0216] In some embodiments, the percentage of uracil or thymine nucleobases in a sequence-optimized nucleotide sequence (e.g., encoding an MCM polypeptide, a functional fragment, or a variant thereof) is modified (e.g., reduced) with respect to the percentage of uracil or thymine nucleobases in the reference wild-type nucleotide sequence. Such a sequence is referred to as a uracil-modified or thymine-modified sequence. The percentage of uracil or thymine content in a nucleotide sequence can be determined by dividing the number of uracils or thymines in a sequence by the total number of nucleotides and multiplying by 100. In some embodiments, the sequence-optimized nucleotide sequence has a lower uracil or thymine content than the uracil or thymine content in the reference wild-type sequence. In some embodiments, the uracil or thymine content in a sequence-optimized nucleotide sequence of the disclosure is greater than the uracil or thymine content in the reference wild-type sequence and still maintain beneficial effects, e.g., increased expression and/or reduced Toll-Like Receptor (TLR) response when compared to the reference wild-type sequence.

[0217] The uracil or thymine content of wild-type MCM is about 26.67%. In some embodiments, the uracil or thymine content of a uracil- or thymine- modified sequence encoding an MCM polypeptide is less than 26.67%. In some embodiments, the uracil or thymine content

of a uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure is less than 19%, less than 18%, less than 17%, less than 16%, less than 15%, less than 14%, less than 13%, less than 12%, less than 11%, or less than 10%. In some embodiments, the uracil or thymine content is not less than 18%, 17%, 16%, 15%, 14%, 13%, 12%, or 11%. The uracil or thymine content of a sequence disclosed herein, *i.e.*, its total uracil or thymine content, is abbreviated herein as %U_{TL} or %T_{TL}.

[0218] In some embodiments, the uracil or thymine content (%U_{TL} or %T_{TL}) of a uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure is between 11% and 26%, between 12% and 25%, between 12% and 24%, between 13% and 23%, between 13% and 22%, between 14% and 21%, between 14% and 20%, between 14% and 19%, between 14% and 18%, between 14% and 17%, or between 14% and 16%.

[0219] In some embodiments, the uracil or thymine content (%U_{TL} or %T_{TL}) of a uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure is between 13% and 17%, between 13% and 16%, or between 14% and 16%.

[0220] In a particular embodiment, the uracil or thymine content (%U_{TL} or %T_{TL}) of a uracil- or thymine modified sequence encoding an MCM polypeptide of the disclosure is between about 14% and about 16%, *e.g.*, between 14% and 15%.

[0221] A uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure can also be described according to its uracil or thymine content relative to the uracil or thymine content in the corresponding wild-type nucleic acid sequence (%U_{WT} or %T_{WT}), or according to its uracil or thymine content relative to the theoretical minimum uracil or thymine content of a nucleic acid encoding the wild-type protein sequence (%U_{TM} or (%T_{TM}).

[0222] The phrases "uracil or thymine content relative to the uracil or thymine content in the wild type nucleic acid sequence," refers to a parameter determined by dividing the number of uracils or thymines in a sequence-optimized nucleic acid by the total number of uracils or thymines in the corresponding wild-type nucleic acid sequence and multiplying by 100. This parameter is abbreviated herein as %U_{WT} or %T_{WT}.

[0223] In some embodiments, the %U_{WT} or %T_{WT} of a uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure is above 50%, above 55%, above 60%, above 65%, above 70%, above 75%, above 80%, above 85%, above 90%, or above 95%.

[0224] In some embodiments, the %U_{WT} or %T_{WT} of a uracil- or thymine modified sequence encoding an MCM polypeptide of the disclosure is between 42% and 68%, between 43% and

67%, between 44% and 66%, between 45% and 65%, between 46% and 64%, between 47% and 63%, between 48% and 62%, between 49% and 61%, between 50% and 60%, between 51% and 59%, or between 52% and 58%.

[0225] In some embodiments, the %U_{WT} or %T_{WT} of a uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure is between 51% and 60%, between 51% and 59%, between 52% and 59%, between 52% and 58%, or between 53% and 58%.

[0226] In a particular embodiment, the %U_{WT} or %T_{WT} of a uracil- or thymine-modified sequence encoding an MCM polypeptide of the disclosure is between about 53% and about 58%.

[0227] Uracil- or thymine- content relative to the uracil or thymine theoretical minimum, refers to a parameter determined by dividing the number of uracils or thymines in a sequence-optimized nucleotide sequence by the total number of uracils or thymines in a hypothetical nucleotide sequence in which all the codons in the hypothetical sequence are replaced with synonymous codons having the lowest possible uracil or thymine content and multiplying by 100. This parameter is abbreviated herein as %U_{TM} or %T_{TM}.

[0228] For DNA it is recognized that thymine is present instead of uracil, and one would substitute T where U appears. Thus, all the disclosures related to, e.g., %U_{TM}, %U_{WT}, or %U_{TL}, with respect to RNA are equally applicable to %T_{TM}, %T_{WT}, or %T_{TL} with respect to DNA.

[0229] In some embodiments, the %U_{TM} of a uracil-modified sequence encoding an MCM polypeptide of the disclosure is below 300%, below 295%, below 290%, below 285%, below 280%, below 275%, below 270%, below 265%, below 260%, below 255%, below 250%, below 245%, below 240%, below 235%, below 230%, below 225%, below 220%, below 215%, below 200%, below 195%, below 190%, below 185%, below 180%, below 175%, below 170%, below 165%, below 160%, below 155%, below 150%, below 145%, below 140%, below 139%, below 138%, below 137%, below 136%, below 135%, below 134%, below 133%, below 132%, below 131%, below 130%, below 129%, below 128%, below 127%, below 126%, below 125%, below 124%, below 123%, below 122%, below 121%, below 120%, below 119%, below 118%, or below 117%.

[0230] In some embodiments, the %U_{TM} of a uracil-modified sequence encoding an MCM polypeptide of the disclosure is above 100%, above 101%, above 102%, above 103%, above 104%, above 105%, above 106%, above 107%, above 108%, above 109%, above 110%, above 111%, above 112%, above 113%, above 114%, above 115%, above 116%, above 117%, above

118%, above 119%, above 120%, above 121%, above 122%, above 123%, above 124%, above 125%, or above 126%, above 127%, above 128%, or above 129%.

[0231] In some embodiments, the %U_{TM} of a uracil-modified sequence encoding an MCM polypeptide of the disclosure is between 123% and 125%, between 122% and 126%, between 121% and 127%, between 120% and 128%, between 119% and 129%, between 118% and 130%, between 117% and 131%, between 116% and 132%, between 115% and 133%, between 114% and 134%, between 113% and 135%, between 112% and 136%, or between 111% and 137%.

[0232] In some embodiments, the %U_{TM} of a uracil-modified sequence encoding an MCM polypeptide of the disclosure is between about 118% and about 129%.

[0233] In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has a reduced number of consecutive uracils with respect to the corresponding wild-type nucleic acid sequence. For example, two consecutive leucines can be encoded by the sequence CUUUUG, which includes a four uracil cluster. Such a subsequence can be substituted, e.g., with CUGCUC, which removes the uracil cluster.

[0234] Phenylalanine can be encoded by UUC or UUU. Thus, even if phenylalanines encoded by UUU are replaced by UUC, the synonymous codon still contains a uracil pair (UU). Accordingly, the number of phenylalanines in a sequence establishes a minimum number of uracil pairs (UU) that cannot be eliminated without altering the number of phenylalanines in the encoded polypeptide. For example, if the polypeptide, e.g., wild type MCM, has 27, 28, 29, or 30 phenylalanines, the absolute minimum number of uracil pairs (UU) in that uracil-modified sequence encoding the polypeptide, e.g., wild type MCM, can contain is 27, 28, 29, or 30, respectively.

[0235] Wild type MCM contains 82 uracil pairs (UU), and 29 uracil triplets (UUU). In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has a reduced number of uracil triplets (UUU) with respect to the wild-type nucleic acid sequence. In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure contains 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 or no uracil triplets (UUU).

[0236] In some embodiments, a uracil-modified sequence encoding an MCM polypeptide has a reduced number of uracil pairs (UU) with respect to the number of uracil pairs (UU) in the wild-type nucleic acid sequence. In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has a number of uracil pairs (UU) corresponding to the

minimum possible number of uracil pairs (UU) in the wild-type nucleic acid sequence, e.g., 28 uracil pairs in the case of wild type MCM.

[0237] In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, or 54 uracil pairs (UU) less than the number of uracil pairs (UU) in the wild-type nucleic acid sequence. In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has between 20 and 35 uracil pairs (UU).

[0238] The phrase "uracil pairs (UU) relative to the uracil pairs (UU) in the wild type nucleic acid sequence," refers to a parameter determined by dividing the number of uracil pairs (UU) in a sequence-optimized nucleotide sequence by the total number of uracil pairs (UU) in the corresponding wild-type nucleotide sequence and multiplying by 100. This parameter is abbreviated herein as %UU_{wt}.

[0239] In some embodiments, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has a %UU_{wt} less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 40%, less than 30%, or less than 20%.

[0240] In some embodiments, a uracil-modified sequence encoding an MCM polypeptide has a %UU_{wt} between 20% and 50%. In a particular embodiment, a uracil-modified sequence encoding an MCM polypeptide of the disclosure has a %UU_{wt} between 24% and 43%.

[0241] In some embodiments, the polynucleotide of the disclosure comprises a uracil-modified sequence encoding an MCM polypeptide disclosed herein. In some embodiments, the uracil-modified sequence encoding an MCM polypeptide comprises at least one chemically modified nucleobase, e.g., 5-methoxyuracil. In some embodiments, at least 95% of a nucleobase (e.g., uracil) in a uracil-modified sequence encoding an MCM polypeptide of the disclosure are modified nucleobases. In some embodiments, at least 95% of uracil in a uracil-modified sequence encoding an MCM polypeptide is 5-methoxyuracil.

[0242] In some embodiments, the "guanine content of the sequence optimized ORF encoding MCM with respect to the theoretical maximum guanine content of a nucleotide sequence encoding the MCM polypeptide," abbreviated as %G_{TMX} is at least 69%, at least 70%, at least 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100%. In some embodiments, the %G_{TMX} is between about 70% and about 80%, between about

71% and about 79%, between about 71% and about 78%, between about 71% and about 77% or between about 71% and about 76%.

[0243] In some embodiments, the "cytosine content of the ORF relative to the theoretical maximum cytosine content of a nucleotide sequence encoding the MCM polypeptide," abbreviated as %C_{TMX}, is at least about 68%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100%. In some embodiments, the %C_{TMX} is between about 68% and about 77%, between about 69% and about 76%, or between about 70% and about 75%.

[0244] In some embodiments, the "guanine and cytosine content (G/C) of the ORF relative to the theoretical maximum G/C content in a nucleotide sequence encoding the MCM polypeptide," abbreviated as %G/C_{TMX} is at least about 85%, at least about 90%, at least about 95%, or about 100%. The %G/C_{TMX} is between about 85% and about 100%, between about 89% and about 96%, between about 90% and about 95%, or between about 91% and about 94%.

[0245] In some embodiments, the "G/C content in the ORF relative to the G/C content in the corresponding wild-type ORF," abbreviated as %G/C_{WT} is at least 120%, at least 130%, at least 140%, at least 141%, at least 142%, at least 143%, at least 144%, at least 145%, at least 146%, at least 147%, at least 150%, or at least 155%.

[0246] In some embodiments, the average G/C content in the 3rd codon position in the ORF is at least 50%, at least 51%, at least 52%, at least 53%, at least 54%, at least 55%, at least 56%, at least 57%, at least 58%, at least 59%, or at least 60% higher than the average G/C content in the 3rd codon position in the corresponding wild-type ORF.

[0247] In some embodiments, the polynucleotide of the disclosure comprises an open reading frame (ORF) encoding an MCM polypeptide, wherein the ORF has been sequence optimized, and wherein each of %U_{TL}, %U_{WT}, %U_{TM}, %G_{TL}, %G_{WT}, %G_{TMX}, %C_{TL}, %C_{WT}, %C_{TMX}, %G/C_{TL}, %G/C_{WT}, or %G/C_{TMX}, alone or in a combination thereof is in a range between (i) a maximum corresponding to the parameter's maximum value (MAX) plus about 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, or 10 standard deviations (STD DEV), and (ii) a minimum corresponding to the parameter's minimum value (MIN) less 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, or 10 standard deviations (STD DEV).

[0248] Features, which can be considered beneficial in some embodiments of the present disclosure, can be encoded by regions of the polynucleotide and such regions can be upstream (5') or downstream (3') to, or within, a region that encodes a polypeptide. These regions can be

incorporated into the polynucleotide before and/or after sequence optimization of the protein encoding region or open reading frame (ORF). It is not required that a polynucleotide contain both a 5' and 3' flanking region. Examples of such features include, but are not limited to, untranslated regions (UTRs), Kozak sequences, an oligo(dT) sequence, and detectable tags and can include multiple cloning sites that can have XbaI recognition.

[0249] In some embodiments, a 5' UTR and/or a 3' UTR region can be provided as flanking regions. Multiple 5' or 3' UTRs can be included in the flanking regions and can be the same or of different sequences. Any portion of the flanking regions, including none, can be sequence-optimized and any can independently contain one or more different structural or chemical modifications, before and/or after sequence optimization.

[0250] In some embodiments, the polynucleotide of the disclosure comprises, consists essentially or, or consists of the sequence set forth as SEQ ID NO: 769, wherein thymidine is changed to uridine. In other embodiments, the polynucleotide of the disclosure comprises, consists essentially or, or consists of the sequence set forth as SEQ ID NO: 770, wherein thymidine is changed to uridine. In other embodiments, the polynucleotide does not comprise a polyC.

[0251] After optimization (if desired), the polynucleotides components are reconstituted and transformed into a vector such as, but not limited to, plasmids, viruses, cosmids, and artificial chromosomes. For example, the optimized polynucleotide can be reconstituted and transformed into chemically competent *E. coli*, yeast, neurospora, maize, drosophila, etc. where high copy plasmid-like or chromosome structures occur by methods described herein.

[0252] Synthetic polynucleotides and their nucleic acid analogs play an important role in the research and studies of biochemical processes. Various enzyme-assisted and chemical-based methods have been developed to synthesize polynucleotides and nucleic acids.

[0253] Enzymatic methods include *in vitro* transcription that uses RNA polymerases to synthesize the polynucleotides of the present disclosure. Enzymatic methods and RNA polymerases for transcription are described in International Patent Application No. PCT/US2014/53907, the contents of which are herein incorporated by reference in its entirety.

[0254] Solid-phase chemical synthesis can be used to manufacture the polynucleotides described herein or portions thereof. Solid-phase chemical synthesis manufacturing of the polynucleotides described herein are described in International Patent Application No. PCT/US2014/53907, the contents of which are herein incorporated by reference in its entirety.

[0255] Liquid phase chemical synthesis can be used to manufacture the polynucleotides described herein or portions thereof. Liquid phase chemical synthesis manufacturing of the polynucleotides described herein are described in International Patent Application No. PCT/US2014/53907, the contents of which are herein incorporated by reference in its entirety.

[0256] Combinations of different synthetic methods can be used to manufacture the polynucleotides described herein or portions thereof. These combinations are described in International Patent Application No. PCT/US2014/53907, the contents of which are herein incorporated by reference in its entirety.

[0257] Small region synthesis can be used for regions or subregions of the polynucleotides of the present disclosure. These synthesis methods are described in International Patent Application No. PCT/US2014/53907, the contents of which are herein incorporated by reference in its entirety.

[0258] Ligation of polynucleotide regions or subregions can be used to prepare the polynucleotides described herein. These ligation methods are described in International Patent Application No. PCT/US2014/53907, the contents of which are herein incorporated by reference in its entirety.

Polypeptides Encoded by the Polynucleotides of the Disclosure

[0259] In some embodiments, the MCM polypeptides encoded by polynucleotides of the disclosure peptide are functional MCM. As used herein, the term "MCM" protein is used interchangeably with "MUT" protein. Therefore, human MCM protein can be written as human MUT or hMUT, and murine MCM protein can be written as murine MUT or mMUT. In some embodiments, the MCM polypeptides encoded by polynucleotides of the disclosure peptide are variants, peptides or polypeptides containing substitutions, insertions and/or additions, deletions and covalent modifications with respect to an MCM peptide sequence. For example, sequence tags or amino acids, such as one or more lysines, can be added to the peptide sequences encoded by the polynucleotides of the disclosure (*e.g.*, at the N-terminal or C-terminal ends). Sequence tags can be used for peptide purification or localization. Lysines can be used to increase peptide solubility or to allow for biotinylation. In some embodiments, amino acid residues located at the carboxy and amino terminal regions of a polypeptide encoded by the polynucleotides of the disclosure can optionally be deleted providing for truncated sequences.

[0260] In some embodiments, the polynucleotides described herein encode a substitutional variant of an MCM protein. The substitutional variant can comprise one, two, three or more than three substitutions. "Substitutional variants" when referring to polypeptides are those that have at least one amino acid residue in a native or starting sequence removed and a different amino acid inserted in its place at the same position. The substitutions can be single, where only one amino acid in the molecule has been substituted, or they can be multiple, where two or more amino acids have been substituted in the same molecule.

[0261] In some embodiments, the polynucleotides described herein encode a variant of an MCM protein with one or more conservative amino acids substitutions. As used herein the term "conservative amino acid substitution" refers to the substitution of an amino acid that is normally present in the sequence with a different amino acid of similar size, charge, or polarity. Examples of conservative substitutions include the substitution of a non-polar (hydrophobic) residue such as isoleucine, valine and leucine for another non-polar residue. Likewise, examples of conservative substitutions include the substitution of one polar (hydrophilic) residue for another such as between arginine and lysine, between glutamine and asparagine, and between glycine and serine. Additionally, the substitution of a basic residue such as lysine, arginine or histidine for another, or the substitution of one acidic residue such as aspartic acid or glutamic acid for another acidic residue are additional examples of conservative substitutions. Examples of non-conservative substitutions include the substitution of a non-polar (hydrophobic) amino acid residue such as isoleucine, valine, leucine, alanine, methionine for a polar (hydrophilic) residue such as cysteine, glutamine, glutamic acid or lysine and/or a polar residue for a non-polar residue.

[0262] In other embodiments, the polynucleotides encode an insertional MCM variant. "Insertional variants" when referring to polypeptides are those with one or more amino acids inserted immediately adjacent to an amino acid at a particular position in a native or starting sequence. "Immediately adjacent" to an amino acid means connected to either the alpha-carboxy or alpha-amino functional group of the amino acid.

[0263] In other embodiments, the polynucleotides of the disclosure encode a deletional MCM variant. "Deletional variants" when referring to polypeptides are those with one or more amino acids in the native or starting amino acid sequence removed. Ordinarily, deletional variants will have one or more amino acids deleted in a particular region of the molecule.

[0264] In some embodiments, the polynucleotides of the disclosure encode a covalent derivative. "Covalent derivatives" when referring to polypeptides include modifications of a native or starting protein with an organic proteinaceous or non-proteinaceous derivatizing agent, and/or post-translational modifications. Covalent modifications are traditionally introduced by reacting targeted amino acid residues of the protein with an organic derivatizing agent that is capable of reacting with selected side-chains or terminal residues, or by harnessing mechanisms of post-translational modifications that function in selected recombinant host cells. The resultant covalent derivatives are useful in programs directed at identifying residues important for biological activity, for immunoassays, or for the preparation of anti-protein antibodies for immunoaffinity purification of the recombinant glycoprotein. Such modifications are within the ordinary skill in the art and are performed without undue experimentation.

[0265] Certain post-translational modifications are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues can be present in the polypeptides produced in accordance with the present disclosure.

[0266] Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (T. E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)).

[0267] "Features," when referring to polypeptides, are defined as distinct amino acid sequence-based components of a molecule. Features of the polypeptides encoded by the polynucleotides of the present disclosure include surface manifestations, local conformational shape, folds, loops, half-loops, domains, half-domains, sites, termini or any combination thereof.

[0268] As used herein, when referring to polypeptides, the term "domain" refers to a motif of a polypeptide having one or more identifiable structural or functional characteristics or properties (e.g., binding capacity, serving as a site for protein-protein interactions).

[0269] As used herein, when referring to polypeptides, the terms "site" as it pertains to amino acid based embodiments is used synonymously with "amino acid residue" and "amino acid side chain." A site represents a position within a peptide or polypeptide that can be modified, manipulated, altered, derivatized or varied within the polypeptide based molecules of the present disclosure.

[0270] As used herein the terms "termini" or "terminus," when referring to polypeptides, refers to an extremity of a peptide or polypeptide. Such extremity is not limited only to the first or final site of the peptide or polypeptide but can include additional amino acids in the terminal regions. The polypeptide based molecules of the present disclosure can be characterized as having both an N-terminus (terminated by an amino acid with a free amino group (NH₂)) and a C-terminus (terminated by an amino acid with a free carboxyl group (COOH)). Proteins of the disclosure are in some cases made up of multiple polypeptide chains brought together by disulfide bonds or by non-covalent forces (multimers, oligomers). These sorts of proteins will have multiple N- and C-termini. Alternatively, the termini of the polypeptides can be modified such that they begin or end, as the case can be, with a non-polypeptide based moiety such as an organic conjugate.

[0271] Once any of the features have been identified or defined as a desired component of a polypeptide to be encoded by the polynucleotide of the disclosure, any of several manipulations and/or modifications of these features can be performed by moving, swapping, inverting, deleting, randomizing or duplicating. Furthermore, it is understood that manipulation of features can result in the same outcome as a modification to the molecules of the disclosure. For example, a manipulation that involved deleting a domain would result in the alteration of the length of a molecule just as modification of a nucleic acid to encode less than a full length molecule would.

[0272] Modifications and manipulations can be accomplished by methods known in the art such as, but not limited to, site directed mutagenesis or *a priori* incorporation during chemical synthesis. The resulting modified molecules can then be tested for activity using *in vitro* or *in vivo* assays such as those described herein or any other suitable screening assay known in the art.

[0273] According to the present disclosure, the polypeptides can comprise a consensus sequence that is discovered through rounds of experimentation. As used herein a "consensus" sequence is a single sequence that represents a collective population of sequences allowing for variability at one or more sites.

[0274] As recognized by those skilled in the art, protein fragments, functional protein domains, and homologous proteins are also considered to be within the scope of polypeptides of interest of this disclosure. For example, provided herein is any protein fragment (meaning a polypeptide sequence at least one amino acid residue shorter than a reference polypeptide sequence but otherwise identical) of a reference protein 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or greater than 100 amino acids in length. In another example, any protein that includes a stretch of

about 20, about 30, about 40, about 50, or about 100 amino acids that are about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, or about 100% identical to any of the sequences described herein can be utilized in accordance with the disclosure.

[0275] In certain embodiments, a polypeptide encoded by the polynucleotide of the disclosure includes 2, 3, 4, 5, 6, 7, 8, 9, 10, or more mutations as shown in any of the sequences provided or referenced herein.

[0276] In some embodiments, the encoded polypeptide variant has the same or a similar activity as the reference polypeptide. Alternatively, the variant has an altered activity (*e.g.*, increased or decreased) relative to a reference polypeptide. Generally, variants of a particular polynucleotide or polypeptide of the disclosure will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% but less than 100% sequence identity to that particular reference polynucleotide or polypeptide as determined by sequence alignment programs and parameters described herein and known to those skilled in the art. Such tools for alignment include those of the BLAST suite (Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Res.* 25:3389-3402.) Other tools are described herein, specifically in the definition of "Identity."

[0277] Default parameters in the BLAST algorithm include, for example, an expect threshold of 10, Word size of 28, Match/Mismatch Scores 1, -2, Gap costs Linear. Any filter can be applied as well as a selection for species specific repeats, *e.g.*, *Homo sapiens*.

[0278] According to the present disclosure, the protein is encoded by a polynucleotide that can comprise at least a first region of linked nucleosides encoding at least one polypeptide of interest. Some polypeptides encoded by the polynucleotides of interest of the present disclosure are listed in **Table 3** below. In particular, **Table 3** shows human MCM wild type and mutant amino acid sequences.

Table 3. MCM Polypeptides and Polynucleotides

SEQ ID No	Gene	Sequence
SEQ ID NO: 208	Functional Human MCM	MLRAKNQLFLLSPHYLRQVKESSGSRLIQQRLLHQOPLHPEWAALAKKQLKGKNPED LIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWTIRQYAGFSTV EESNKFYKDNIKAGQQGLSVAFDLATHRGYDSDNPRVRGVDVGMAGVAIDTVEDTKILF DGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPEKEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIIADIFEYTAKHMPKFNSISISGYHMQEAGADAILELAYTLADGLEYSRT

		GLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLR AHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIAR NTQIIIIQEESGIPKVADPWGGSYMECLTNDVYDAALKLINEIEEMGGMAKAVAEGIP KLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAI DNTSVRNRQIEKLKKIK SSRDQALAERCLAALTECAASGDGNILALAVDASRARCTVGEITDALKKVFGGEHKAND RMVSGAYRQEFGESKEITSAIKRVHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGF ADLGFVDVIGPLFQTPREVAQQAVDADVHAVGISTLAAGHKTLVPELIKELNSLGRPD ILVMCGGVIPPQDYEFLFEVGVS NVFGPGTRI PKAAVQVLDDIEKCLEKKQQSV
a.a. 33 to 750 of SEQ ID NO: 208	Mature human MCM	LHQOQPLHPEWAALAKKQLKGKNPEDLIWHTPEGISIKPLYSKRDTMDLPEELPGVKP FTRGPYPTMYTFRPWTIRQYAGFSTVEESNKFYKDNIKAGQQGLSVAFDLATHRGYDS DNPRVRGDVGMAGVAIDTVEDTKILFDGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQ GVPKEKLTGTIQNDILKEFMVRNTYIFPPEPSMKIIADI FEYTAKHMPKFNSISISGY HMQEAGADAI LELAYTLADGLEYSRTGLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLRAHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFG GTQSLHTNSFDEALGLPTVKSARIARNTQIIIIQEESGIPKVADPWGGSYMECLTNDV YDAALKLINEIEEMGGMAKAVAEGIPKLRIEECAARRQARIDSGSEVIVGVNKYQLEK EDAVEVLAI DNTSVRNRQIEKLKKIKSSRDQALAERCLAALTECAASGDGNILALAVD ASRARCTVGEITDALKKVFGGEHKANDRMVSGAYRQEFGESKEITSAIKRVHKFMEREG RRPRLLVAKMGQDGHDRGAKVIATGFADLGFVDVIGPLFQTPREVAQQAVDADVHAVG VSTLAAGHKTLVPELIKELNSLGRPDILVMCGGVIPPQDYEFLFEVGVS NVFGPGTRI PKAAVQVLDDIEKCLEKKQQSV
SEQ ID NO: 209	Function al Human MCM I69V	MLRAKNQLFLLSPHYLRQVKESSGSRLIQQRLLHQOQPLHPEWAALAKKQLKGKNPED LIWHTPEGISVKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWTIRQYAGFSTV EESNKFYKDNIKAGQQGLSVAFDLATHRGYDS DNPRVRGDVGMAGVAIDTVEDTKILF DGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPKEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIIADI FEYTAKHMPKFNSISISGYHMQEAGADAI LELAYTLADGLEYSRT GLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLR AHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIAR NTQIIIIQEESGIPKVADPWGGSYMECLTNDVYDAALKLINEIEEMGGMAKAVAEGIP KLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAI DNTSVRNRQIEKLKKIK SSRDQALAERCLAALTECAASGDGNILALAVDASRARCTVGEITDALKKVFGGEHKAND RMVSGAYRQEFGESKEITSAIKRVHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGF ADLGFVDVIGPLFQTPREVAQQAVDADVHAVGISTLAAGHKTLVPELIKELNSLGRPD ILVMCGGVIPPQDYEFLFEVGVS NVFGPGTRI PKAAVQVLDDIEKCLEKKQQSV
SEQ ID NO: 210	Function al Human MCM A499T	MLRAKNQLFLLSPHYLRQVKESSGSRLIQQRLLHQOQPLHPEWAALAKKQLKGKNPED LIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWTIRQYAGFSTV EESNKFYKDNIKAGQQGLSVAFDLATHRGYDS DNPRVRGDVGMAGVAIDTVEDTKILF DGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPKEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIIADI FEYTAKHMPKFNSISISGYHMQEAGADAI LELAYTLADGLEYSRT GLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLR AHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIAR NTQIIIIQEESGIPKVADPWGGSYMECLTNDVYDAALKLINEIEEMGGMAKAVAEGIP KLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDTVEVLAI DNTSVRNRQIEKLKKIK SSRDQALAERCLAALTECAASGDGNILALAVDASRARCTVGEITDALKKVFGGEHKAND RMVSGAYRQEFGESKEITSAIKRVHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGF ADLGFVDVIGPLFQTPREVAQQAVDADVHAVGISTLAAGHKTLVPELIKELNSLGRPD ILVMCGGVIPPQDYEFLFEVGVS NVFGPGTRI PKAAVQVLDDIEKCLEKKQQSV
SEQ ID NO: 211	Function al Human MCM R532H	MLRAKNQLFLLSPHYLRQVKESSGSRLIQQRLLHQOQPLHPEWAALAKKQLKGKNPED LIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWTIRQYAGFSTV EESNKFYKDNIKAGQQGLSVAFDLATHRGYDS DNPRVRGDVGMAGVAIDTVEDTKILF DGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPKEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIIADI FEYTAKHMPKFNSISISGYHMQEAGADAI LELAYTLADGLEYSRT GLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLR AHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIAR NTQIIIIQEESGIPKVADPWGGSYMECLTNDVYDAALKLINEIEEMGGMAKAVAEGIP KLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAI DNTSVRNRQIEKLKKIK SSRDQALAEHCLAALTECAASGDGNILALAVDASRARCTVGEITDALKKVFGGEHKAND RMVSGAYRQEFGESKEITSAIKRVHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGF

		ADLGFDVDIGPLFQTPREVAQQAVDADVHAVGISTLAAGHKTLVPELIKELNSLGRPD ILVMCGGVIPPQDYEFLEFVGVSNVFPGPTRIPKAAVQVLDDIEKCLEKKQQSV
SEQ ID NO: 212	Function al Human MCM T598A	MLRAKNQLFLLSPHYLRQVKESSESRLLIQQRLLHQQOPLHPEWAALAKKQLKGKNPED LIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWTIRQYAGFSTV EESNKFYKDNIKAGQQGLSVAFDLATHRGYDSNPRVRGDVGMAGVAIDTVEDTKILF DGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIIADIFEYTAKHMPKFNSISISGYHMQEAGADAI LELAYTLADGLEYSRT GLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLR AHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIAR NTQII IQEESGIPKVADPWGGSYMMECLTNDVYDAALKLINEIEEMGGMAKAVAEGIP KLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAI DNTSVRNRQIEKLLKKIK SSRDQALAERCLAALTECAASGDGNILALAVDASRARCTVGEITDALKKVFGHEHKAND RMVSGAYRQEFGESKEIASAI KRVHKFMEREGRRLVAKMGQDGHDRGAKVIATGF ADLGFDVDIGPLFQTPREVAQQAVDADVHAVGISTLAAGHKTLVPELIKELNSLGRPD ILVMCGGVIPPQDYEFLEFVGVSNVFPGPTRIPKAAVQVLDDIEKCLEKKQQSV
SEQ ID NO: 213	Function al Human MCM I671V	MLRAKNQLFLLSPHYLRQVKESSESRLLIQQRLLHQQOPLHPEWAALAKKQLKGKNPED LIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWTIRQYAGFSTV EESNKFYKDNIKAGQQGLSVAFDLATHRGYDSNPRVRGDVGMAGVAIDTVEDTKILF DGIPLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIIADIFEYTAKHMPKFNSISISGYHMQEAGADAI LELAYTLADGLEYSRT GLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLR AHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIAR NTQII IQEESGIPKVADPWGGSYMMECLTNDVYDAALKLINEIEEMGGMAKAVAEGIP KLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAI DNTSVRNRQIEKLLKKIK SSRDQALAERCLAALTECAASGDGNILALAVDASRARCTVGEITDALKKVFGHEHKAND RMVSGAYRQEFGESKEITSAI KRVHKFMEREGRRLVAKMGQDGHDRGAKVIATGF ADLGFDVDIGPLFQTPREVAQQAVDADVHAVGVSTLAAGHKTLVPELIKELNSLGRPD ILVMCGGVIPPQDYEFLEFVGVSNVFPGPTRIPKAAVQVLDDIEKCLEKKQQSV

II. Modified Polynucleotides

[0279] The disclosure also includes a modified polynucleotide comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide that contains polynucleotides that are chemically and/or structurally modified. When the polynucleotides of the present disclosure are chemically and/or structurally modified the polynucleotides can be referred to as "modified polynucleotides."

[0280] The present disclosure provides for modified nucleosides and nucleotides of a polynucleotide (*e.g.*, RNA polynucleotides, such as mRNA polynucleotides). A "nucleoside" refers to a compound containing a sugar molecule (*e.g.*, a pentose or ribose) or a derivative thereof in combination with an organic base (*e.g.*, a purine or pyrimidine) or a derivative thereof (also referred to herein as "nucleobase"). A nucleotide" refers to a nucleoside, including a phosphate group. Modified nucleotides can be synthesized by any useful method, such as, for example, chemically, enzymatically, or recombinantly, to include one or more modified or non-natural nucleosides. Polynucleotides can comprise a region or regions of linked nucleosides. Such regions can have variable backbone linkages. The linkages can be standard

phosphodioester linkages, in which case the polynucleotides would comprise regions of nucleotides.

[0281] The modifications can be various distinct modifications. In some embodiments, the regions can contain one, two, or more (optionally different) nucleoside or nucleotide modifications. In some embodiments, a modified polynucleotide, introduced to a cell can exhibit reduced degradation in the cell, as compared to an unmodified polynucleotide.

Structural Modifications

[0282] In some embodiments, the polynucleotides of the present disclosure are structurally modified. As used herein, a "structural" modification is one in which two or more linked nucleosides are inserted, deleted, duplicated, inverted or randomized in a polynucleotide without significant chemical modification to the nucleotides themselves. Because chemical bonds will necessarily be broken and reformed to effect a structural modification, structural modifications are of a chemical nature and hence are chemical modifications. However, structural modifications will result in a different sequence of nucleotides. For example, the polynucleotide "ATCG" can be chemically modified to "AT-5meC-G". The same polynucleotide can be structurally modified from "ATCG" to "ATCCCG". Here, the dinucleotide "CC" has been inserted, resulting in a structural modification to the polynucleotide.

Chemical Modifications

[0283] In some embodiments, the polynucleotides of the present disclosure are chemically modified. As used herein in reference to a polynucleotide, the terms "chemical modification" or, as appropriate, "chemically modified" refer to modification with respect to adenosine (A), guanosine (G), uridine (U), thymidine (T) or cytidine (C) ribo- or deoxyribonucleosides in one or more of their position, pattern, percent or population. Generally, herein, these terms are not intended to refer to the ribonucleotide modifications in naturally occurring 5'-terminal mRNA cap moieties.

[0284] In some embodiments, the polynucleotides of the present disclosure can have a uniform chemical modification of all or any of the same nucleoside type or a population of modifications produced by mere downward titration of the same starting modification in all or any of the same nucleoside type, or a measured percent of a chemical modification of all any of the same nucleoside type but with random incorporation, such as where all uridines are replaced

by a uridine analog, *e.g.*, pseudouridine or 5-methoxyuridine. In another embodiment, the polynucleotides can have a uniform chemical modification of two, three, or four of the same nucleoside type throughout the entire polynucleotide (such as all uridines and all cytosines, etc. are modified in the same way).

[0285] Modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases, wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two complementary non-standard base structures. One example of such non-standard base pairing is the base pairing between the modified nucleotide inosine and adenine, cytosine or uracil. Any combination of base/sugar or linker can be incorporated into polynucleotides of the present disclosure.

[0286] The skilled artisan will appreciate that, except where otherwise noted, polynucleotide sequences set forth in the instant application will recite "T"s in a representative DNA sequence but where the sequence represents RNA, the "T"s would be substituted for "U"s.

[0287] Modifications of polynucleotides (*e.g.*, RNA polynucleotides, such as mRNA polynucleotides) that are useful in the compositions, methods and synthetic processes of the present disclosure include, but are not limited to the following nucleotides, nucleosides, and nucleobases: 2-methylthio-N6-(*cis*-hydroxyisopentenyl)adenosine; 2-methylthio-N6-methyladenosine; 2-methylthio-N6-threonyl carbamoyladenosine; N6-glycinyln-carbamoyladenosine; N6-isopentenyladenosine; N6-methyladenosine; N6-threonyln-carbamoyladenosine; 1,2'-O-dimethyladenosine; 1-methyladenosine; 2'-O-methyladenosine; 2'-O-ribosyladenosine (phosphate); 2-methyladenosine; 2-methylthio-N6-isopentenyladenosine; 2-methylthio-N6-hydroxynorvalyl carbamoyladenosine; 2'-O-methyladenosine; 2'-O-ribosyladenosine (phosphate); Isopentenyladenosine; N6-(*cis*-hydroxyisopentenyl)adenosine; N6,2'-O-dimethyladenosine; N6,2'-O-dimethyladenosine; N6,N6,2'-O-trimethyladenosine; N6,N6-dimethyladenosine; N6-acetyladenosine; N6-hydroxynorvalylcarbamoyladenosine; N6-methyl-N6-threonyln-carbamoyladenosine; 2-methyladenosine; 2-methylthio-N6-isopentenyladenosine; 7-deaza-adenosine; N1-methyladenosine; N6, N6 (dimethyl)adenine; N6-*cis*-hydroxy-isopentenyl-adenosine; α -thio-adenosine; 2 (amino)adenine; 2 (aminopropyl)adenine; 2 (methylthio) N6 (isopentenyl)adenine; 2-(alkyl)adenine; 2-(aminoalkyl)adenine; 2-(aminopropyl)adenine; 2-(halo)adenine; 2-

(halo)adenine; 2-(propyl)adenine; 2'-Amino-2'-deoxy-ATP; 2'-Azido-2'-deoxy-ATP; 2'-Deoxy-2'-a-aminoadenosine TP; 2'-Deoxy-2'-a-azidoadenosine TP; 6 (alkyl)adenine; 6 (methyl)adenine; 6-(alkyl)adenine; 6-(methyl)adenine; 7 (deaza)adenine; 8 (alkenyl)adenine; 8 (alkynyl)adenine; 8 (amino)adenine; 8 (thioalkyl)adenine; 8-(alkenyl)adenine; 8-(alkyl)adenine; 8-(alkynyl)adenine; 8-(amino)adenine; 8-(halo)adenine; 8-(hydroxyl)adenine; 8-(thioalkyl)adenine; 8-(thiol)adenine; 8-azido-adenosine; aza adenine; deaza adenine; N6 (methyl)adenine; N6-(isopentyl)adenine; 7-deaza-8-aza-adenosine; 7-methyladenine; 1-Deazaadenosine TP; 2'Fluoro-N6-Bz-deoxyadenosine TP; 2'-OMe-2-Amino-ATP; 2'O-methyl-N6-Bz-deoxyadenosine TP; 2'-a-Ethynyladenosine TP; 2-aminoadenine; 2-Aminoadenosine TP; 2-Amino-ATP; 2'-a-Trifluoromethyladenosine TP; 2-Azidoadenosine TP; 2'-b-Ethynyladenosine TP; 2-Bromoadenosine TP; 2'-b-Trifluoromethyladenosine TP; 2-Chloroadenosine TP; 2'-Deoxy-2',2'-difluoroadenosine TP; 2'-Deoxy-2'-a-mercaptopadenosine TP; 2'-Deoxy-2'-a-thiomethoxyadenosine TP; 2'-Deoxy-2'-b-aminoadenosine TP; 2'-Deoxy-2'-b-azidoadenosine TP; 2'-Deoxy-2'-b-bromoadenosine TP; 2'-Deoxy-2'-b-chloroadenosine TP; 2'-Deoxy-2'-b-fluoroadenosine TP; 2'-Deoxy-2'-b-iodoadenosine TP; 2'-Deoxy-2'-b-mercaptopadenosine TP; 2'-Deoxy-2'-b-thiomethoxyadenosine TP; 2-Fluoroadenosine TP; 2-Iodoadenosine TP; 2-Mercaptopadenosine TP; 2-methoxy-adenine; 2-methylthio-adenine; 2-Trifluoromethyladenosine TP; 3-Deaza-3-bromoadenosine TP; 3-Deaza-3-chloroadenosine TP; 3-Deaza-3-fluoroadenosine TP; 3-Deaza-3-iodoadenosine TP; 3-Deazaadenosine TP; 4'-Azidoadenosine TP; 4'-Carbocyclic adenosine TP; 4'-Ethynyladenosine TP; 5'-Homo-adenosine TP; 8-Aza-ATP; 8-bromo-adenosine TP; 8-Trifluoromethyladenosine TP; 9-Deazaadenosine TP; 2-aminopurine; 7-deaza-2,6-diaminopurine; 7-deaza-8-aza-2,6-diaminopurine; 7-deaza-8-aza-2-aminopurine; 2,6-diaminopurine; 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine; 2-thiocytidine; 3-methylcytidine; 5-formylcytidine; 5-hydroxymethylcytidine; 5-methylcytidine; N4-acetylcytidine; 2'-O-methylcytidine; 2'-O-methylcytidine; 5,2'-O-dimethylcytidine; 5-formyl-2'-O-methylcytidine; Lysidine; N4,2'-O-dimethylcytidine; N4-acetyl-2'-O-methylcytidine; N4-methylcytidine; N4,N4-Dimethyl-2'-OMe-Cytidine TP; 4-methylcytidine; 5-aza-cytidine; Pseudo-iso-cytidine; pyrrolo-cytidine; α -thio-cytidine; 2-(thio)cytosine; 2'-Amino-2'-deoxy-CTP; 2'-Azido-2'-deoxy-CTP; 2'-Deoxy-2'-a-aminocytidine TP; 2'-Deoxy-2'-a-azidocytidine TP; 3 (deaza) 5 (aza)cytosine; 3 (methyl)cytosine; 3-(alkyl)cytosine; 3-(deaza) 5 (aza)cytosine; 3-(methyl)cytidine; 4,2'-O-dimethylcytidine; 5 (halo)cytosine; 5 (methyl)cytosine; 5 (propynyl)cytosine; 5 (trifluoromethyl)cytosine; 5-(alkyl)cytosine; 5-(alkynyl)cytosine; 5-

(halo)cytosine; 5-(propynyl)cytosine; 5-(trifluoromethyl)cytosine; 5-bromo-cytidine; 5-iodo-cytidine; 5-propynyl cytosine; 6-(azo)cytosine; 6-aza-cytidine; aza cytosine; deaza cytosine; N4 (acetyl)cytosine; 1-methyl-1-deaza-pseudoisocytidine; 1-methyl-pseudoisocytidine; 2-methoxy-5-methyl-cytidine; 2-methoxy-cytidine; 2-thio-5-methyl-cytidine; 4-methoxy-1-methyl-pseudoisocytidine; 4-methoxy-pseudoisocytidine; 4-thio-1-methyl-1-deaza-pseudoisocytidine; 4-thio-1-methyl-pseudoisocytidine; 4-thio-pseudoisocytidine; 5-aza-zebularine; 5-methyl-zebularine; pyrrolo-pseudoisocytidine; Zebularine; (E)-5-(2-Bromo-vinyl)cytidine TP; 2,2'-anhydro-cytidine TP hydrochloride; 2'Fluor-N4-Bz-cytidine TP; 2'Fluoro-N4-Acetyl-cytidine TP; 2'-O-Methyl-N4-Acetyl-cytidine TP; 2'O-methyl-N4-Bz-cytidine TP; 2'-a-Ethynylcytidine TP; 2'-a-Trifluoromethylcytidine TP; 2'-b-Ethynylcytidine TP; 2'-b-Trifluoromethylcytidine TP; 2'-Deoxy-2',2'-difluorocytidine TP; 2'-Deoxy-2'-a-mercaptocytidine TP; 2'-Deoxy-2'-a-thiomethoxycytidine TP; 2'-Deoxy-2'-b-aminocytidine TP; 2'-Deoxy-2'-b-azidocytidine TP; 2'-Deoxy-2'-b-bromocytidine TP; 2'-Deoxy-2'-b-chlorocytidine TP; 2'-Deoxy-2'-b-fluorocytidine TP; 2'-Deoxy-2'-b-iodocytidine TP; 2'-Deoxy-2'-b-mercaptocytidine TP; 2'-Deoxy-2'-b-thiomethoxycytidine TP; 2'-O-Methyl-5-(1-propynyl)cytidine TP; 3'-Ethynylcytidine TP; 4'-Azidocytidine TP; 4'-Carbocyclic cytidine TP; 4'-Ethynylcytidine TP; 5-(1-Propynyl)ara-cytidine TP; 5-(2-Chloro-phenyl)-2-thiocytidine TP; 5-(4-Amino-phenyl)-2-thiocytidine TP; 5-Aminoallyl-CTP; 5-Cyanocytidine TP; 5-Ethynylara-cytidine TP; 5-Ethynylcytidine TP; 5'-Homo-cytidine TP; 5-Methoxycytidine TP; 5-Trifluoromethyl-Cytidine TP; N4-Amino-cytidine TP; N4-Benzoyl-cytidine TP; Pseudoisocytidine; 7-methylguanosine; N2,2'-O-dimethylguanosine; N2-methylguanosine; Wyosine; 1,2'-O-dimethylguanosine; 1-methylguanosine; 2'-O-methylguanosine; 2'-O-ribosylguanosine (phosphate); 2'-O-methylguanosine; 2'-O-ribosylguanosine (phosphate); 7-aminomethyl-7-deazaguanosine; 7-cyano-7-deazaguanosine; Archaeosine; Methylwyosine; N2,7-dimethylguanosine; N2,N2,2'-O-trimethylguanosine; N2,N2,7-trimethylguanosine; N2,N2-dimethylguanosine; N2,7,2'-O-trimethylguanosine; 6-thio-guanosine; 7-deaza-guanosine; 8-oxo-guanosine; N1-methyl-guanosine; α -thio-guanosine; 2 (propyl)guanine; 2-(alkyl)guanine; 2'-Amino-2'-deoxy-GTP; 2'-Azido-2'-deoxy-GTP; 2'-Deoxy-2'-a-aminoguanosine TP; 2'-Deoxy-2'-a-azidoguanosine TP; 6 (methyl)guanine; 6-(alkyl)guanine; 6-(methyl)guanine; 6-methyl-guanosine; 7 (alkyl)guanine; 7 (deaza)guanine; 7 (methyl)guanine; 7-(alkyl)guanine; 7-(deaza)guanine; 7-(methyl)guanine; 8 (alkyl)guanine; 8 (alkynyl)guanine; 8 (halo)guanine; 8 (thioalkyl)guanine; 8-(alkenyl)guanine; 8-(alkyl)guanine; 8-(alkynyl)guanine; 8-(amino)guanine; 8-(halo)guanine; 8-(hydroxyl)guanine; 8-

(thioalkyl)guanine; 8-(thiol)guanine; aza guanine; deaza guanine; N (methyl)guanine; N-(methyl)guanine; 1-methyl-6-thio-guanosine; 6-methoxy-guanosine; 6-thio-7-deaza-8-azaguanosine; 6-thio-7-deaza-guanosine; 6-thio-7-methyl-guanosine; 7-deaza-8-azaguanosine; 7-methyl-8-oxo-guanosine; N2,N2-dimethyl-6-thio-guanosine; N2-methyl-6-thio-guanosine; 1-Me-GTP; 2'Fluoro-N2-isobutyl-guanosine TP; 2'O-methyl-N2-isobutyl-guanosine TP; 2'-a-Ethynylguanosine TP; 2'-a-Trifluoromethylguanosine TP; 2'-b-Ethynylguanosine TP; 2'-b-Trifluoromethylguanosine TP; 2'-Deoxy-2',2'-difluoroguanosine TP; 2'-Deoxy-2'-a-mercaptoguanosine TP; 2'-Deoxy-2'-a-thiomethoxyguanosine TP; 2'-Deoxy-2'-b-aminoguanosine TP; 2'-Deoxy-2'-b-azidoguanosine TP; 2'-Deoxy-2'-b-bromoguanosine TP; 2'-Deoxy-2'-b-chloroguanosine TP; 2'-Deoxy-2'-b-fluoroguanosine TP; 2'-Deoxy-2'-b-iodoguanosine TP; 2'-Deoxy-2'-b-mercaptoguanosine TP; 2'-Deoxy-2'-b-thiomethoxyguanosine TP; 4'-Azidoguanosine TP; 4'-Carbocyclic guanosine TP; 4'-Ethynylguanosine TP; 5'-Homo-guanosine TP; 8-bromoguanosine TP; 9-Deazaguanosine TP; N2-isobutyl-guanosine TP; 1-methylinosine; Inosine; 1,2'-O-dimethylinosine; 2'-O-methylinosine; 7-methylinosine; 2'-O-methylinosine; Epoxyqueuosine; galactosyl-queuosine; Mannosylqueuosine; Queuosine; allyamino-thymidine; aza thymidine; deaza thymidine; deoxy-thymidine; 2'-O-methyluridine; 2-thiouridine; 3-methyluridine; 5-carboxymethyluridine; 5-hydroxyuridine; 5-methyluridine; 5-taurinomethyl-2-thiouridine; 5-taurinomethyluridine; Dihydrouridine; Pseudouridine; (3-(3-amino-3-carboxypropyl)uridine; 1-methyl-3-(3-amino-5-carboxypropyl)pseudouridine; 1-methylpseudouridine; 1-ethylpseudouridine; 2'-O-methyluridine; 2'-O-methylpseudouridine; 2'-O-methyluridine; 2-thio-2'-O-methyluridine; 3-(3-amino-3-carboxypropyl)uridine; 3,2'-O-dimethyluridine; 3-Methyl-pseudo-Uridine TP; 4-thiouridine; 5-(carboxyhydroxymethyl)uridine; 5-(carboxyhydroxymethyl)uridine methyl ester; 5,2'-O-dimethyluridine; 5,6-dihydro-uridine; 5-aminomethyl-2-thiouridine; 5-carbamoylmethyl-2'-O-methyluridine; 5-carbamoylmethyluridine; 5-carboxyhydroxymethyluridine; 5-carboxyhydroxymethyluridine methyl ester; 5-carboxymethylaminomethyl-2'-O-methyluridine; 5-carboxymethylaminomethyl-2-thiouridine; 5-carboxymethylaminomethyl-2-thiouridine; 5-carboxymethylaminomethyluridine; 5-carboxymethylaminomethyluridine; 5-Carbamoylmethyluridine TP; 5-methoxycarbonylmethyl-2'-O-methyluridine; 5-methoxycarbonylmethyl-2-thiouridine; 5-methoxycarbonylmethyluridine; 5-methyluridine,); 5-methoxyuridine; 5-methyl-2-thiouridine; 5-methylaminomethyl-2-selenouridine; 5-methylaminomethyl-2-thiouridine; 5-methylaminomethyluridine; 5-Methyldihydrouridine; 5-Oxyacetic acid- Uridine TP; 5-Oxyacetic acid-methyl ester-Uridine TP;

N1-methyl-pseudo-uracil; N1-ethyl-pseudo-uracil; uridine 5-oxyacetic acid; uridine 5-oxyacetic acid methyl ester; 3-(3-Amino-3-carboxypropyl)-Uridine TP; 5-(iso-Pentenylaminomethyl)-2-thiouridine TP; 5-(iso-Pentenylaminomethyl)-2'-O-methyluridine TP; 5-(iso-Pentenylaminomethyl)uridine TP; 5-propynyl uracil; α -thio-uridine; 1 (aminoalkylaminocarbonylethylenyl)-2(thio)-pseudouracil; 1 (aminoalkylaminocarbonylethylenyl)-2,4-(dithio)pseudouracil; 1 (aminoalkylaminocarbonylethylenyl)-4 (thio)pseudouracil; 1 (aminoalkylaminocarbonylethylenyl)-pseudouracil; 1 (aminocarbonylethylenyl)-2(thio)-pseudouracil; 1 (aminocarbonylethylenyl)-2,4-(dithio)pseudouracil; 1 (aminocarbonylethylenyl)-4 (thio)pseudouracil; 1 (aminocarbonylethylenyl)-pseudouracil; 1 substituted 2(thio)-pseudouracil; 1 substituted 2,4-(dithio)pseudouracil; 1 substituted 4 (thio)pseudouracil; 1 substituted pseudouracil; 1-(aminoalkylamino-carbonylethylenyl)-2-(thio)-pseudouracil; 1-Methyl-3-(3-amino-3-carboxypropyl) pseudouridine TP; 1-Methyl-3-(3-amino-3-carboxypropyl)pseudo-UTP; 1-Methyl-pseudo-UTP; 1-Ethyl-pseudo-UTP; 2 (thio)pseudouracil; 2' deoxy uridine; 2' fluorouridine; 2-(thio)uracil; 2,4-(dithio)pseudouracil; 2' methyl, 2' amino, 2' azido, 2' fluoro-guanosine; 2'-Amino-2'-deoxy-UTP; 2'-Azido-2'-deoxy-UTP; 2'-Azido-deoxyuridine TP; 2'-O-methylpseudouridine; 2' deoxy uridine; 2' fluorouridine; 2'-Deoxy-2'-aminouridine TP; 2'-Deoxy-2'-a-azidouridine TP; 2-methylpseudouridine; 3 (3 amino-3 carboxypropyl)uracil; 4 (thio)pseudouracil; 4-(thio)pseudouracil; 4-(thio)uracil; 4-thiouracil; 5 (1,3-diazole-1-alkyl)uracil; 5 (2-aminopropyl)uracil; 5 (aminoalkyl)uracil; 5 (dimethylaminoalkyl)uracil; 5 (guanidiniumalkyl)uracil; 5 (methoxycarbonylmethyl)-2-(thio)uracil; 5 (methoxycarbonyl-methyl)uracil; 5 (methyl) 2 (thio)uracil; 5 (methyl) 2,4 (dithio)uracil; 5 (methyl) 4 (thio)uracil; 5 (methylaminomethyl)-2 (thio)uracil; 5 (methylaminomethyl)-2,4 (dithio)uracil; 5 (methylaminomethyl)-4 (thio)uracil; 5 (propynyl)uracil; 5 (trifluoromethyl)uracil; 5-(2-aminopropyl)uracil; 5-(alkyl)-2-(thio)pseudouracil; 5-(alkyl)-2,4 (dithio)pseudouracil; 5-(alkyl)-4 (thio)pseudouracil; 5-(alkyl)pseudouracil; 5-(alkyl)uracil; 5-(alkynyl)uracil; 5-(allylamino)uracil; 5-(cyanoalkyl)uracil; 5-(dialkylaminoalkyl)uracil; 5-(dimethylaminoalkyl)uracil; 5-(guanidiniumalkyl)uracil; 5-(halo)uracil; 5-(1,3-diazole-1-alkyl)uracil; 5-(methoxy)uracil; 5-(methoxycarbonylmethyl)-2-(thio)uracil; 5-(methoxycarbonyl-methyl)uracil; 5-(methyl) 2(thio)uracil; 5-(methyl) 2,4 (dithio)uracil; 5-(methyl) 4 (thio)uracil; 5-(methyl)-2-(thio)pseudouracil; 5-(methyl)-2,4 (dithio)pseudouracil; 5-(methyl)-4 (thio)pseudouracil; 5-(methyl)pseudouracil; 5-(methylaminomethyl)-2 (thio)uracil; 5-(methylaminomethyl)-2,4(dithio)uracil; 5-

(methylaminomethyl)-4-(thio)uracil; 5-(propynyl)uracil; 5-(trifluoromethyl)uracil; 5-aminoallyl-uridine; 5-bromo-uridine; 5-iodo-uridine; 5-uracil; 6 (azo)uracil; 6-(azo)uracil; 6-aza-uridine; allyamino-uracil; aza uracil; deaza uracil; N3 (methyl)uracil; P pseudo-UTP-1-2-ethanoic acid; Pseudouracil; 4-Thio-pseudo-UTP; 1-carboxymethyl-pseudouridine; 1-methyl-1-deaza-pseudouridine; 1-propynyl-uridine; 1-aurinomethyl-1-methyl-uridine; 1-aurinomethyl-4-thio-uridine; 1-aurinomethyl-pseudouridine; 2-methoxy-4-thio-pseudouridine; 2-thio-1-methyl-1-deaza-pseudouridine; 2-thio-1-methyl-pseudouridine; 2-thio-5-aza-uridine; 2-thio-dihydropseudouridine; 2-thio-dihydrouridine; 2-thio-pseudouridine; 4-methoxy-2-thio-pseudouridine; 4-methoxy-pseudouridine; 4-thio-1-methyl-pseudouridine; 4-thio-pseudouridine; 5-aza-uridine; Dihydropseudouridine; ($\pm$)-1-(2-Hydroxypropyl)pseudouridine TP; (2R)-1-(2-Hydroxypropyl)pseudouridine TP; (2S)-1-(2-Hydroxypropyl)pseudouridine TP; (E)-5-(2-Bromo-vinyl)ara-uridine TP; (E)-5-(2-Bromo-vinyl)uridine TP; (Z)-5-(2-Bromo-vinyl)ara-uridine TP; (Z)-5-(2-Bromo-vinyl)uridine TP; 1-(2,2,2-Trifluoroethyl)-pseudo-UTP; 1-(2,2,3,3,3-Pentafluoropropyl)pseudouridine TP; 1-(2,2-Diethoxyethyl)pseudouridine TP; 1-(2,4,6-Trimethylbenzyl)pseudouridine TP; 1-(2,4,6-Trimethyl-benzyl)pseudo-UTP; 1-(2,4,6-Trimethyl-phenyl)pseudo-UTP; 1-(2-Amino-2-carboxyethyl)pseudo-UTP; 1-(2-Amino-ethyl)pseudo-UTP; 1-(2-Hydroxyethyl)pseudouridine TP; 1-(2-Methoxyethyl)pseudouridine TP; 1-(3,4-Bis-trifluoromethoxybenzyl)pseudouridine TP; 1-(3,4-Dimethoxybenzyl)pseudouridine TP; 1-(3-Amino-3-carboxypropyl)pseudo-UTP; 1-(3-Amino-propyl)pseudo-UTP; 1-(3-Cyclopropyl-prop-2-ynyl)pseudouridine TP; 1-(4-Amino-4-carboxybutyl)pseudo-UTP; 1-(4-Amino-benzyl)pseudo-UTP; 1-(4-Amino-butyl)pseudo-UTP; 1-(4-Amino-phenyl)pseudo-UTP; 1-(4-Azidobenzyl)pseudouridine TP; 1-(4-Bromobenzyl)pseudouridine TP; 1-(4-Chlorobenzyl)pseudouridine TP; 1-(4-Fluorobenzyl)pseudouridine TP; 1-(4-Iodobenzyl)pseudouridine TP; 1-(4-Methanesulfonylbenzyl)pseudouridine TP; 1-(4-Methoxybenzyl)pseudouridine TP; 1-(4-Methoxy-benzyl)pseudo-UTP; 1-(4-Methoxy-phenyl)pseudo-UTP; 1-(4-Methylbenzyl)pseudouridine TP; 1-(4-Methyl-benzyl)pseudo-UTP; 1-(4-Nitrobenzyl)pseudouridine TP; 1-(4-Nitro-benzyl)pseudo-UTP; 1-(4-Nitro-phenyl)pseudo-UTP; 1-(4-Thiomethoxybenzyl)pseudouridine TP; 1-(4-Trifluoromethoxybenzyl)pseudouridine TP; 1-(4-Trifluoromethylbenzyl)pseudouridine TP; 1-(5-Amino-pentyl)pseudo-UTP; 1-(6-Amino-hexyl)pseudo-UTP; 1,6-Dimethyl-pseudo-UTP; 1-[3-(2-{2-[2-(2-Aminoethoxy)-ethoxy]-ethoxy}-ethoxy)-propionyl]pseudouridine TP; 1-{3-[2-(2-Aminoethoxy)-ethoxy]-propionyl } pseudouridine TP; 1-Acetyl-pseudouridine TP; 1-Alkyl-6-(1-propynyl)-pseudo-UTP; 1-Alkyl-6-

(2-propynyl)-pseudo-UTP; 1-Alkyl-6-allyl-pseudo-UTP; 1-Alkyl-6-ethynyl-pseudo-UTP; 1-Alkyl-6-homoallyl-pseudo-UTP; 1-Alkyl-6-vinyl-pseudo-UTP; 1-Allylpseudouridine TP; 1-Aminomethyl-pseudo-UTP; 1-Benzoylpseudouridine TP; 1-Benzylloxymethylpseudouridine TP; 1-Benzyl-pseudo-UTP; 1-Biotinyl-PEG2-pseudouridine TP; 1-Biotinylpseudouridine TP; 1-Butyl-pseudo-UTP; 1-Cyanomethylpseudouridine TP; 1-Cyclobutylmethyl-pseudo-UTP; 1-Cyclobutyl-pseudo-UTP; 1-Cycloheptylmethyl-pseudo-UTP; 1-Cycloheptyl-pseudo-UTP; 1-Cyclohexylmethyl-pseudo-UTP; 1-Cyclohexyl-pseudo-UTP; 1-Cyclooctylmethyl-pseudo-UTP; 1-Cyclooctyl-pseudo-UTP; 1-Cyclopentylmethyl-pseudo-UTP; 1-Cyclopentyl-pseudo-UTP; 1-Cyclopropylmethyl-pseudo-UTP; 1-Cyclopropyl-pseudo-UTP; 1-Ethyl-pseudo-UTP; 1-Hexyl-pseudo-UTP; 1-Homoallylpseudouridine TP; 1-Hydroxymethylpseudouridine TP; 1-iso-propyl-pseudo-UTP; 1-Me-2-thio-pseudo-UTP; 1-Me-4-thio-pseudo-UTP; 1-Me-alpha-thio-pseudo-UTP; 1-Methanesulfonylmethylpseudouridine TP; 1-Methoxymethylpseudouridine TP; 1-Methyl-6-(2,2,2-Trifluoroethyl)pseudo-UTP; 1-Methyl-6-(4-morpholino)-pseudo-UTP; 1-Methyl-6-(4-thiomorpholino)-pseudo-UTP; 1-Methyl-6-(substituted phenyl)pseudo-UTP; 1-Methyl-6-amino-pseudo-UTP; 1-Methyl-6-azido-pseudo-UTP; 1-Methyl-6-bromo-pseudo-UTP; 1-Methyl-6-butyl-pseudo-UTP; 1-Methyl-6-chloro-pseudo-UTP; 1-Methyl-6-cyano-pseudo-UTP; 1-Methyl-6-dimethylamino-pseudo-UTP; 1-Methyl-6-ethoxy-pseudo-UTP; 1-Methyl-6-ethylcarboxylate-pseudo-UTP; 1-Methyl-6-ethyl-pseudo-UTP; 1-Methyl-6-fluoro-pseudo-UTP; 1-Methyl-6-formyl-pseudo-UTP; 1-Methyl-6-hydroxyamino-pseudo-UTP; 1-Methyl-6-hydroxy-pseudo-UTP; 1-Methyl-6-iodo-pseudo-UTP; 1-Methyl-6-iso-propyl-pseudo-UTP; 1-Methyl-6-methoxy-pseudo-UTP; 1-Methyl-6-methylamino-pseudo-UTP; 1-Methyl-6-phenyl-pseudo-UTP; 1-Methyl-6-propyl-pseudo-UTP; 1-Methyl-6-tert-butyl-pseudo-UTP; 1-Methyl-6-trifluoromethoxy-pseudo-UTP; 1-Methyl-6-trifluoromethyl-pseudo-UTP; 1-Morpholinomethylpseudouridine TP; 1-Pentyl-pseudo-UTP; 1-Phenyl-pseudo-UTP; 1-Pivaloylpseudouridine TP; 1-Propargylpseudouridine TP; 1-Propyl-pseudo-UTP; 1-propynylpseudouridine; 1-p-tolyl-pseudo-UTP; 1-tert-Butyl-pseudo-UTP; 1-Thiomethoxymethylpseudouridine TP; 1-Thiomorpholinomethylpseudouridine TP; 1-Trifluoroacetylpseudouridine TP; 1-Trifluoromethyl-pseudo-UTP; 1-Vinylpseudouridine TP; 2,2'-anhydro-uridine TP; 2'-bromo-deoxyuridine TP; 2'-F-5-Methyl-2'-deoxy-UTP; 2'-OMe-5-Me-UTP; 2'-OMe-pseudo-UTP; 2'-a-Ethynyluridine TP; 2'-a-Trifluoromethyluridine TP; 2'-b-Ethynyluridine TP; 2'-b-Trifluoromethyluridine TP; 2'-Deoxy-2',2'-difluorouridine TP; 2'-Deoxy-2'-a-mercaptopuridine TP; 2'-Deoxy-2'-a-thiomethoxyuridine TP; 2'-Deoxy-2'-b-aminouridine TP;

2'-Deoxy-2'-b-azidouridine TP; 2'-Deoxy-2'-b-bromouridine TP; 2'-Deoxy-2'-b-chlorouridine TP; 2'-Deoxy-2'-b-fluorouridine TP; 2'-Deoxy-2'-b-iodouridine TP; 2'-Deoxy-2'-b-mercaptopuridine TP; 2'-Deoxy-2'-b-thiomethoxyuridine TP; 2-methoxy-4-thio-uridine; 2-methoxyuridine; 2'-O-Methyl-5-(1-propynyl)uridine TP; 3-Alkyl-pseudo-UTP; 4'-Azidouridine TP; 4'-Carbocyclic uridine TP; 4'-Ethynyluridine TP; 5-(1-Propynyl)ara-uridine TP; 5-(2-Furanyl)uridine TP; 5-Cyanouridine TP; 5-Dimethylaminouridine TP; 5'-Homo-uridine TP; 5-iodo-2'-fluoro-deoxyuridine TP; 5-Phenylethynyluridine TP; 5-Trideuteromethyl-6-deuterouridine TP; 5-Trifluoromethyl-Uridine TP; 5-Vinylarauridine TP; 6-(2,2,2-Trifluoroethyl)-pseudo-UTP; 6-(4-Morpholino)-pseudo-UTP; 6-(4-Thiomorpholino)-pseudo-UTP; 6-(Substituted-Phenyl)-pseudo-UTP; 6-Amino-pseudo-UTP; 6-Azido-pseudo-UTP; 6-Bromo-pseudo-UTP; 6-Butyl-pseudo-UTP; 6-Chloro-pseudo-UTP; 6-Cyano-pseudo-UTP; 6-Dimethylamino-pseudo-UTP; 6-Ethoxy-pseudo-UTP; 6-Ethylcarboxylate-pseudo-UTP; 6-Ethyl-pseudo-UTP; 6-Fluoro-pseudo-UTP; 6-Formyl-pseudo-UTP; 6-Hydroxyamino-pseudo-UTP; 6-Hydroxy-pseudo-UTP; 6-Iodo-pseudo-UTP; 6-iso-Propyl-pseudo-UTP; 6-Methoxy-pseudo-UTP; 6-Methylamino-pseudo-UTP; 6-Methyl-pseudo-UTP; 6-Phenyl-pseudo-UTP; 6-Phenyl-pseudo-UTP; 6-Propyl-pseudo-UTP; 6-tert-Butyl-pseudo-UTP; 6-Trifluoromethoxy-pseudo-UTP; 6-Trifluoromethyl-pseudo-UTP; Alpha-thio-pseudo-UTP; Pseudouridine 1-(4-methylbenzenesulfonic acid) TP; Pseudouridine 1-(4-methylbenzoic acid) TP; Pseudouridine TP 1-[3-(2-ethoxy)]propionic acid; Pseudouridine TP 1-[3-{2-(2-[2-(2-ethoxy)-ethoxy]-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-[3-{2-(2-[2-{2(2-ethoxy)-ethoxy}-ethoxy]-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-[3-{2-(2-[2-ethoxy]-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-[3-{2-(2-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-methylphosphonic acid; Pseudouridine TP 1-methylphosphonic acid diethyl ester; Pseudo-UTP-N1-3-propionic acid; Pseudo-UTP-N1-4-butanoic acid; Pseudo-UTP-N1-5-pentanoic acid; Pseudo-UTP-N1-6-hexanoic acid; Pseudo-UTP-N1-7-heptanoic acid; Pseudo-UTP-N1-methyl-p-benzoic acid; Pseudo-UTP-N1-p-benzoic acid; Wybutosine; Hydroxywybutosine; Isowyosine; Peroxywybutosine; undermodified hydroxywybutosine; 4-demethylwyosine; 2,6-(diamino)purine; 1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 1,3-(diaz)-2-(oxo)-phenthiazin-1-yl; 1,3-(diaz)-2-(oxo)-phenoxazin-1-yl; 1,3,5-(triaz)-2,6-(diox)-naphthalene; 2-(amino)purine; 2,4,5-(trimethyl)phenyl; 2'-methyl, 2'-amino, 2'-azido, 2'-fluoro-cytidine; 2'-methyl, 2'-amino, 2'-azido, 2'-fluoro-adenine; 2'-methyl, 2'-amino, 2'-azido, 2'-fluoro-uridine; 2'-amino-2'-deoxyribose; 2-amino-6-Chloro-purine; 2-aza-inosinyl; 2'-azido-2'-deoxyribose; 2'-fluoro-2'-deoxyribose; 2'-fluoro-modified bases; 2'-O-methyl-ribose; 2-

oxo-7-aminopyridopyrimidin-3-yl; 2-oxo-pyridopyrimidine-3-yl; 2-pyridinone; 3 nitropyrrole; 3-(methyl)-7-(propynyl)isocarbostyryl; 3-(methyl)isocarbostyryl; 4-(fluoro)-6-(methyl)benzimidazole; 4-(methyl)benzimidazole; 4-(methyl)indolyl; 4,6-(dimethyl)indolyl; 5 nitroindole; 5 substituted pyrimidines; 5-(methyl)isocarbostyryl; 5-nitroindole; 6-(aza)pyrimidine; 6-(azo)thymine; 6-(methyl)-7-(aza)indolyl; 6-chloro-purine; 6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; 7-(aminoalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenthiazin-1-yl; 7-(aminoalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-(aminoalkylhydroxy)-1,3-(diaz)-2-(oxo)-phenoxazin-1-yl; 7-(aminoalkylhydroxy)-1,3-(diaz)-2-(oxo)-phenthiazin-1-yl; 7-(aminoalkylhydroxy)-1,3-(diaz)-2-(oxo)-phenoxazin-1-yl; 7-(aza)indolyl; 7-(guanidiniumalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-(guanidiniumalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenthiazin-1-yl; 7-(guanidiniumalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-(guanidiniumalkylhydroxy)-1,3-(diaz)-2-(oxo)-phenoxazin-1-yl; 7-(guanidiniumalkylhydroxy)-1,3-(diaz)-2-(oxo)-phenthiazin-1-yl; 7-(guanidiniumalkylhydroxy)-1,3-(diaz)-2-(oxo)-phenoxazin-1-yl; 7-(propynyl)isocarbostyryl; 7-(propynyl)isocarbostyryl, propynyl-7-(aza)indolyl; 7-deaza-inosinyl; 7-substituted 1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-substituted 1,3-(diaz)-2-(oxo)-phenoxazin-1-yl; 9-(methyl)-imidizopyridinyl; Aminoindolyl; Anthracenyl; bis-ortho-(aminoalkylhydroxy)-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; bis-ortho-substituted-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; Difluorotolyl; Hypoxanthine; Imidizopyridinyl; Inosinyl; Isocarbostyryl; Isoguanisine; N2-substituted purines; N6-methyl-2-amino-purine; N6-substituted purines; N-alkylated derivative; Napthalenyl; Nitrobenzimidazolyl; Nitroimidazolyl; Nitroindazolyl; Nitropirazolyl; Nubularine; O6-substituted purines; O-alkylated derivative; ortho-(aminoalkylhydroxy)-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; ortho-substituted-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; Oxoformycin TP; para-(aminoalkylhydroxy)-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; para-substituted-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; Pentacenyl; Phenanthracenyl; Phenyl; propynyl-7-(aza)indolyl; Pyrenyl; pyridopyrimidin-3-yl; pyridopyrimidin-3-yl, 2-oxo-7-amino-pyridopyrimidin-3-yl; pyrrolo-pyrimidin-2-on-3-yl; Pyrrolopyrimidinyl; Pyrrolopyrizinyl; Stilbenzyl; substituted 1,2,4-triazoles; Tetracenyl; Tubercidine; Xanthine; Xanthosine-5'-TP; 2-thio-zebularine; 5-aza-2-thio-zebularine; 7-deaza-2-amino-purine; pyridin-4-one ribonucleoside; 2-Amino-riboside-TP; Formycin A TP; Formycin B TP; Pyrrolosine TP; 2'-OH-ara-adenosine TP; 2'-OH-ara-cytidine TP; 2'-OH-ara-uridine TP; 2'-

OH-ara-guanosine TP; 5-(2-carbomethoxyvinyl)uridine TP; and N6-(19-Aminopentaoxanonadecyl)adenosine TP.

[0288] In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of the aforementioned modified nucleobases.

[0289] In some embodiments, the mRNA comprises at least one chemically modified nucleoside. In some embodiments, the at least one chemically modified nucleoside is selected from the group consisting of pseudouridine (ψ), 2-thiouridine (s2U), 4'-thiouridine, 5-methylcytosine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-pseudouridine, 2-thio-5-aza-uridine, 2-thio-dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-pseudouridine, 4-methoxy-2-thio-pseudouridine, 4-methoxy-pseudouridine, 4-thio-1-methyl-pseudouridine, 4-thio-pseudouridine, 5-aza-uridine, dihydropseudouridine, 5-methyluridine, 5-methoxyuridine, 2'-O-methyl uridine, 1-methyl-pseudouridine (m1 ψ), 1-ethyl-pseudouridine (e1 ψ), 5-methoxyuridine (mo5U), 5-methyl-cytidine (m5C), α -thio-guanosine, α -thio-adenosine, 5-cyano uridine, 4'-thio uridine 7-deaza-adenine, 1-methyl-adenosine (m1A), 2-methyl-adenine (m2A), N6-methyl-adenosine (m6A), and 2,6-Diaminopurine, (I), 1-methyl-inosine (m1I), wyosine (imG), methylwyosine (mimG), 7-deaza-guanosine, 7-cyano-7-deaza-guanosine (preQ0), 7-aminomethyl-7-deaza-guanosine (preQ1), 7-methyl-guanosine (m7G), 1-methyl-guanosine (m1G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 2,8-dimethyladenosine, 2-geranylthiouridine, 2-lysidine, 2-selenouridine, 3-(3-amino-3-carboxypropyl)-5,6-dihydrouridine, 3-(3-amino-3-carboxypropyl)pseudouridine, 3-methylpseudouridine, 5-(carboxyhydroxymethyl)-2'-O-methyluridine methyl ester, 5-aminomethyl-2-geranylthiouridine, 5-aminomethyl-2-selenouridine, 5-aminomethyluridine, 5-carbamoylhydroxymethyluridine, 5-carbamoylmethyl-2-thiouridine, 5-carboxymethyl-2-thiouridine, 5-carboxymethylaminomethyl-2-geranylthiouridine, 5-carboxymethylaminomethyl-2-selenouridine, 5-cyanomethyluridine, 5-hydroxycytidine, 5-methylaminomethyl-2-geranylthiouridine, 7-aminocarboxypropyl-demethylwyosine, 7-aminocarboxypropylwyosine, 7-aminocarboxypropylwyosine methyl ester, 8-methyladenosine, N4,N4-dimethylcytidine, N6-formyladenosine, N6-hydroxymethyladenosine, agmatidine, cyclic N6-threonylcarbamoyladenosine, glutamyl-queuosine, methylated undermodified hydroxywybutosine, N4,N4,2'-O-trimethylcytidine, geranylated 5-methylaminomethyl-2-thiouridine, geranylated 5-carboxymethylaminomethyl-2-thiouridine, Qbase, preQ0base, preQ1base, and two or more combinations thereof. In some

embodiments, the at least one chemically modified nucleoside is selected from the group consisting of pseudouridine, 1-methyl-pseudouridine, 1-ethyl-pseudouridine, 5-methylcytosine, 5-methoxyuridine, and a combination thereof. In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of the aforementioned modified nucleobases.

Base Modifications

[0290] In certain embodiments, the chemical modification is at nucleobases in the polynucleotides (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide). In some embodiments, modified nucleobases in the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) are selected from the group consisting of 1-methyl-pseudouridine (m1 ψ), 1-ethyl-pseudouridine (e1 ψ), 5-methoxy-uridine (mo5U), 5-methyl-cytidine (m5C), pseudouridine (ψ), α -thio-guanosine and α -thio-adenosine. In some embodiments, the polynucleotide includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of the aforementioned modified nucleobases.

[0291] In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises pseudouridine (ψ) and 5-methyl-cytidine (m5C). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 1-methyl-pseudouridine (m1 ψ). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 1-ethyl-pseudouridine (e1 ψ). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 1-methyl-pseudouridine (m1 ψ) and 5-methyl-cytidine (m5C). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 1-ethyl-pseudouridine (e1 ψ) and 5-methyl-cytidine (m5C). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 2-thiouridine (s2U). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 2-thiouridine and 5-methyl-cytidine (m5C). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises methoxy-uridine (mo5U). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 5-methoxy-uridine (mo5U) and 5-methyl-cytidine (m5C). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 2'-O-methyl uridine. In some embodiments, the

polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 2'-O-methyl uridine and 5-methyl-cytidine (m5C). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises N6-methyl-adenosine (m6A). In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises N6-methyl-adenosine (m6A) and 5-methyl-cytidine (m5C).

[0292] In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) is uniformly modified (*e.g.*, fully modified, modified throughout the entire sequence) for a particular modification. For example, a polynucleotide can be uniformly modified with 5-methyl-cytidine (m5C), meaning that all cytosine residues in the mRNA sequence are replaced with 5-methyl-cytidine (m5C). Similarly, a polynucleotide can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as any of those set forth above.

[0293] In some embodiments, the chemically modified nucleosides in the open reading frame are selected from the group consisting of uridine, adenine, cytosine, guanine, and any combination thereof.

[0294] In some embodiments, the modified nucleobase is a modified cytosine. Examples of nucleobases and nucleosides having a modified cytosine include N4-acetyl-cytidine (ac4C), 5-methyl-cytidine (m5C), 5-halo-cytidine (*e.g.*, 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm5C), 1-methyl-pseudoisocytidine, 2-thio-cytidine (s2C), 2-thio-5-methyl-cytidine.

[0295] In some embodiments, a modified nucleobase is a modified uridine. Example nucleobases and nucleosides having a modified uridine include 5-cyano uridine or 4'-thio uridine.

[0296] In some embodiments, a modified nucleobase is a modified adenine. Example nucleobases and nucleosides having a modified adenine include 7-deaza-adenine, 1-methyl-adenosine (m1A), 2-methyl-adenine (m2A), N6-methyl-adenine (m6A), and 2,6-Diaminopurine.

[0297] In some embodiments, a modified nucleobase is a modified guanine. Example nucleobases and nucleosides having a modified guanine include inosine (I), 1-methyl-inosine (m1I), wyosine (imG), methylwyosine (mimG), 7-deaza-guanosine, 7-cyano-7-deaza-guanosine (preQ0), 7-aminomethyl-7-deaza-guanosine (preQ1), 7-methyl-guanosine (m7G), 1-methyl-guanosine (m1G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine.

[0298] In some embodiments, the nucleobase modified nucleotides in the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) are 5-methoxyuridine.

[0299] In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of modified nucleobases.

[0300] In some embodiments, at least 95% of a type of nucleobases (*e.g.*, uracil) in a polynucleotide of the disclosure (*e.g.*, an mRNA polynucleotide encoding MCM) are modified nucleobases. In some embodiments, at least 95% of uracil in a polynucleotide of the present disclosure (*e.g.*, an mRNA polynucleotide encoding MCM) is 5-methoxyuracil.

[0301] In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) comprises 5-methoxyuridine (5mo5U) and 5-methyl-cytidine (m5C).

[0302] In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) is uniformly modified (*e.g.*, fully modified, modified throughout the entire sequence) for a particular modification. For example, a polynucleotide can be uniformly modified with 5-methoxyuridine, meaning that substantially all uridine residues in the mRNA sequence are replaced with 5-methoxyuridine. Similarly, a polynucleotide can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as any of those set forth above.

[0303] In some embodiments, the modified nucleobase is a modified cytosine.

[0304] In some embodiments, a modified nucleobase is a modified uracil. Example nucleobases and nucleosides having a modified uracil include 5-methoxyuracil.

[0305] In some embodiments, a modified nucleobase is a modified adenine.

[0306] In some embodiments, a modified nucleobase is a modified guanine.

[0307] In some embodiments, the nucleobases, sugar, backbone, or any combination thereof in the open reading frame encoding an MCM polypeptide are chemically modified by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0308] In some embodiments, the uridine nucleosides in the open reading frame encoding an MCM polypeptide are chemically modified by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0309] In some embodiments, the adenosine nucleosides in the open reading frame encoding an MCM polypeptide are chemically modified by at least 10%, at least 20%, at least 30%, at

least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0310] In some embodiments, the cytidine nucleosides in the open reading frame encoding an MCM polypeptide are chemically modified by at least at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0311] In some embodiments, the guanosine nucleosides in the open reading frame encoding an MCM polypeptide are chemically modified by at least at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or 100%.

[0312] In some embodiments, the polynucleotides can include any useful linker between the nucleosides. Such linkers, including backbone modifications, that are useful in the composition of the present disclosure include, but are not limited to the following: 3'-alkylene phosphonates, 3'-amino phosphoramidate, alkene containing backbones, aminoalkylphosphoramidates, aminoalkylphosphotriesters, boranophosphates, $-\text{CH}_2-\text{O}-\text{N}(\text{CH}_3)-\text{CH}_2-$, $-\text{CH}_2-\text{N}(\text{CH}_3)-\text{N}(\text{CH}_3)-\text{CH}_2-$, $-\text{CH}_2-\text{NH}-\text{CH}_2-$, chiral phosphonates, chiral phosphorothioates, formacetyl and thioformacetyl backbones, methylene (methyylimino), methylene formacetyl and thioformacetyl backbones, methyleneimino and methylenehydrazino backbones, morpholino linkages, $-\text{N}(\text{CH}_3)-\text{CH}_2-\text{CH}_2-$, oligonucleosides with heteroatom internucleoside linkage, phosphinates, phosphoramidates, phosphorodithioates, phosphorothioate internucleoside linkages, phosphorothioates, phosphotriesters, PNA, siloxane backbones, sulfamate backbones, sulfide sulfoxide and sulfone backbones, sulfonate and sulfonamide backbones, thionoalkylphosphonates, thionoalkylphosphotriesters, and thionophosphoramidates.

Modifications on the Sugar

[0313] The modified nucleosides and nucleotides (*e.g.*, building block molecules), which can be incorporated into a polynucleotide (*e.g.*, RNA or mRNA, as described herein), can be modified on the sugar of the ribonucleic acid. For example, the 2' hydroxyl group (OH) can be modified or replaced with a number of different substituents. Exemplary substitutions at the 2'-position include, but are not limited to, H, halo, optionally substituted C_{1-6} alkyl; optionally substituted C_{1-6} alkoxy; optionally substituted C_{6-10} aryloxy; optionally substituted C_{3-8}

cycloalkyl; optionally substituted C₃₋₈ cycloalkoxy; optionally substituted C₆₋₁₀ aryloxy; optionally substituted C₆₋₁₀ aryl-C₁₋₆ alkoxy, optionally substituted C₁₋₁₂ (heterocyclyl)oxy; a sugar (*e.g.*, ribose, pentose, or any described herein); a polyethyleneglycol (PEG), -O(CH₂CH₂O)_nCH₂CH₂OR, where R is H or optionally substituted alkyl, and n is an integer from 0 to 20 (*e.g.*, from 0 to 4, from 0 to 8, from 0 to 10, from 0 to 16, from 1 to 4, from 1 to 8, from 1 to 10, from 1 to 16, from 1 to 20, from 2 to 4, from 2 to 8, from 2 to 10, from 2 to 16, from 2 to 20, from 4 to 8, from 4 to 10, from 4 to 16, and from 4 to 20); "locked" nucleic acids (LNA) in which the 2'-hydroxyl is connected by a C₁₋₆ alkylene or C₁₋₆ heteroalkylene bridge to the 4'-carbon of the same ribose sugar, where exemplary bridges included methylene, propylene, ether, or amino bridges; aminoalkyl, as defined herein; aminoalkoxy, as defined herein; amino as defined herein; and amino acid, as defined herein

[0314] Generally, RNA includes the sugar group ribose, which is a 5-membered ring having an oxygen. Exemplary, non-limiting modified nucleotides include replacement of the oxygen in ribose (*e.g.*, with S, Se, or alkylene, such as methylene or ethylene); addition of a double bond (*e.g.*, to replace ribose with cyclopentenyl or cyclohexenyl); ring contraction of ribose (*e.g.*, to form a 4-membered ring of cyclobutane or oxetane); ring expansion of ribose (*e.g.*, to form a 6- or 7-membered ring having an additional carbon or heteroatom, such as for anhydrohexitol, altritol, mannitol, cyclohexanyl, cyclohexenyl, and morpholino that also has a phosphoramidate backbone); multicyclic forms (*e.g.*, tricyclo; and "unlocked" forms, such as glycol nucleic acid (GNA) (*e.g.*, R-GNA or S-GNA, where ribose is replaced by glycol units attached to phosphodiester bonds), threose nucleic acid (TNA, where ribose is replaced with α -L-threofuranosyl-(3'→2')) , and peptide nucleic acid (PNA, where 2-amino-ethyl-glycine linkages replace the ribose and phosphodiester backbone). The sugar group can also contain one or more carbons that possess the opposite stereochemical configuration than that of the corresponding carbon in ribose. Thus, a polynucleotide molecule can include nucleotides containing, *e.g.*, arabinose, as the sugar. Such sugar modifications are taught International Patent Publication No. WO2013052523 and International Patent Application No. PCT/US2013/75177, the contents of each of which are incorporated herein by reference in its entirety.

Combinations of Modifications

[0315] The polynucleotides of the disclosure can include a combination of modifications to the sugar, the nucleobase, and/or the internucleoside linkage. These combinations can include any one or more modifications described herein.

[0316] Examples of modified nucleotides and modified nucleotide combinations are provided below in **Table 4**. These combinations of modified nucleotides can be used to form the polynucleotides of the disclosure. Unless otherwise noted, the modified nucleotides can be completely substituted for the natural nucleotides of the polynucleotides of the disclosure. As a non-limiting example, the natural nucleotide uridine can be substituted with a modified nucleoside described herein. In another non-limiting example, the natural nucleotide uridine can be partially substituted (*e.g.*, about 0.1%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 99.9%) with at least one of the modified nucleoside disclosed herein. Any combination of base/sugar or linker can be incorporated into the polynucleotides of the disclosure and such modifications are taught in International Patent Publication No. WO2013052523 and International Patent Application No. PCT/US2013/75177, the contents of each of which are incorporated herein by reference in its entirety.

Table 4. Combinations

Uracil	Cytosine	Adenine	Guanine
5-methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	N4Ac-CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Trifluoromethyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Hydroxymethyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Bromo-CTP	ATP	GTP
5-Methoxy-UTP	N4Ac-CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Trifluoromethyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Hydroxymethyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Bromo-CTP	ATP	GTP
5-Methoxy-UTP	N4-Ac-CTP	ATP	GTP
5-Methoxy-UTP	5-Iodo-CTP	ATP	GTP

5-Methoxy-UTP	5-Bromo-CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	5-Methyl-CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	5-Methyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
5-Methoxy-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	CTP	Alpha-thio-ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	Alpha-thio-ATP	GTP
5-Methoxy-UTP	CTP	ATP	Alpha-thio-GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	Alpha-thio-GTP
5-Methoxy-UTP	CTP	N6-Me-ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	N6-Me-	GTP

		ATP	
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	5-Methyl-CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	5-Methyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
5-Methoxy-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Ethyl-CTP	ATP	GTP
5-Methoxy-UTP	5-Methoxy-CTP	ATP	GTP
5-Methoxy-UTP	5-Ethynyl-CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % 1-Methyl-pseudo-UTP	5-Methyl-CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % 1-Methyl-pseudo-UTP	5-Methyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % 1-Methyl-pseudo-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
5-Methoxy-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP

75 % 5-Methoxy-UTP + 25 % 1-Methyl-pseudo-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % 1-Methyl-pseudo-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % 1-Methyl-pseudo-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % 1-Methyl-pseudo-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % 1-Methyl-pseudo-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % 1-Methyl-pseudo-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % 1-Methyl-pseudo-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % 1-Methyl-pseudo-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % 1-Methyl-pseudo-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % 1-Methyl-pseudo-UTP	CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % 1-Methyl-pseudo-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % 1-Methyl-pseudo-UTP	CTP	ATP	GTP
5-methoxy-UTP	CTP	ATP	GTP
5-methoxy-UTP	CTP	ATP	GTP
5-methoxy-UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	5-Methyl-CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	5-Methyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
5-Methoxy-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP

25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
5-Methoxy-UTP	5-Methyl-CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	5-Methyl-CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	5-Methyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	5-Methyl-CTP	ATP	GTP
5-Methoxy-UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
5-Methoxy-UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	CTP	ATP	GTP
50 % 5-Methoxy-UTP + 50 % UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP

5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	50 % 5-Methyl-CTP + 50 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Fluoro-CTP	ATP	GTP
5-Methoxy-UTP	5-Phenyl-CTP	ATP	GTP
5-Methoxy-UTP	N4-Bz-CTP	ATP	GTP
5-Methoxy-UTP	CTP	N6-Isopentenyl-ATP	GTP
5-Methoxy-UTP	N4-Ac-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % N4-Ac-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % N4-Ac-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % N4-Ac-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % N4-Ac-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Hydroxymethyl-CTP	ATP	GTP

25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Hydroxymethyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Hydroxymethyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Hydroxymethyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Hydroxymethyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	N4-Methyl CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % N4-Methyl CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % N4-Methyl CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % N4-Methyl CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % N4-Methyl CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Trifluoromethyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Trifluoromethyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Trifluoromethyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Trifluoromethyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Trifluoromethyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Bromo-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Bromo-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Bromo-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Bromo-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Bromo-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Iodo-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Iodo-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Iodo-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Iodo-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Iodo-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Ethyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Ethyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Ethyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Ethyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Ethyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Methoxy-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Methoxy-CTP + 75 % CTP	ATP	GTP

25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Methoxy-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Methoxy-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Methoxy-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Ethynyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Ethynyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Ethynyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Ethynyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Ethynyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Pseudo-iso-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Pseudo-iso-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Pseudo-iso-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Pseudo-iso-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Pseudo-iso-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Formyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Formyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Formyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Formyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Formyl-CTP + 25 % CTP	ATP	GTP
5-Methoxy-UTP	5-Aminoallyl-CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	25 % 5-Aminoallyl-CTP + 75 % CTP	ATP	GTP
25 % 5-Methoxy-UTP + 75 % UTP	75 % 5-Aminoallyl-CTP + 25 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	25 % 5-Aminoallyl-CTP + 75 % CTP	ATP	GTP
75 % 5-Methoxy-UTP + 25 % UTP	75 % 5-Aminoallyl-CTP + 25 % CTP	ATP	GTP

III. Polynucleotide Architecture

[0317] Traditionally, the basic components of an mRNA molecule include at least a coding region, a 5'UTR, a 3'UTR, a 5' cap and a poly-A tail. The polynucleotides of the present disclosure can function as mRNA but are distinguished from wild-type mRNA in their functional and/or structural design features that serve, *e.g.*, to overcome existing problems of effective polypeptide production using nucleic-acid based therapeutics.

In Vitro Transcribed Polynucleotides

[0318] The disclosure also includes an *in vitro* transcribed polynucleotide comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0319] Polynucleotides which are made using only *in vitro* transcription (IVT) enzymatic synthesis methods are referred to as "IVT polynucleotides." Methods of making IVT polynucleotides are known in the art and are described, *e.g.*, in International Publication Nos. WO2013151666, WO2013151667, WO2013151668, WO2013151663, WO2013151669, WO2013151670, WO2013151664, WO2013151665, WO2013151671, WO2013151672 and WO2013151736; the contents of each of which are herein incorporated by reference in their entireties.

[0320] The shortest length of the first region of the primary construct of the IVT polynucleotide can be the length of a nucleic acid sequence that is sufficient to encode for MCM, a fragment thereof, or variant thereof. The length of the first region of the primary construct of the IVT polynucleotide encoding the polypeptide of interest can be greater than about 30 nucleotides in length (*e.g.*, at least or greater than about 2,154, 2,250, 2,500, and 3,000, 4,000, 4,100, 4,200, 4,300, 4,400, 4,500, 4,600, 4,700, 4,800, 4,900, 5,000, 5,100, 5,200, 5,300, 5,400, 5,500, 6,000, 7,000, 8,000, 9,000, 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000 or up to and including 100,000 nucleotides).

[0321] In some embodiments, the first and second flanking regions of the IVT polynucleotide can range independently from 15-1,000 nucleotides in length (*e.g.*, greater than 30, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 4,500, 5,000, 5,500 nucleotides or at least 30, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 4,500, 5,000, 5,500 nucleotides).

[0322] In some embodiments, the tailing sequence of the IVT polynucleotide can range from absent to 500 nucleotides in length (*e.g.*, at least 60, 70, 80, 90, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, or 500 nucleotides). Where the tailing region is a polyA tail, the length can be determined in units of or as a function of polyA Binding Protein binding. In this embodiment, the polyA tail is long enough to bind at least 4 monomers of PolyA Binding Protein. PolyA

Binding Protein monomers bind to stretches of approximately 38 nucleotides. As such, it has been observed that polyA tails of about 80 nucleotides and 160 nucleotides are functional.

[0323] In some embodiments, the capping region of the IVT polynucleotide can comprise a single cap or a series of nucleotides forming the cap. In this embodiment the capping region can be from 1 to 10, *e.g.*, 2-9, 3-8, 4-7, 1-5, 5-10, or at least 2, or 10 or fewer nucleotides in length. In some embodiments, the cap is absent.

[0324] In some embodiments, the first and second operational regions of the IVT polynucleotide can range from 3 to 40, *e.g.*, 5-30, 10-20, 15, or at least 4, or 30 or fewer nucleotides in length and can comprise, in addition to a Start and/or Stop codon, one or more signal and/or restriction sequences.

[0325] In some embodiments, the IVT polynucleotides can be structurally modified or chemically modified. When the IVT polynucleotides are chemically and/or structurally modified, the polynucleotides can be referred to as "modified IVT polynucleotides."

[0326] In some embodiments, if the IVT polynucleotides are chemically modified they can have a uniform chemical modification of all or any of the same nucleoside type or a population of modifications produced by mere downward titration of the same starting modification in all or any of the same nucleoside type, or a measured percent of a chemical modification of all any of the same nucleoside type but with random incorporation, such as where all uridines are replaced by a uridine analog, *e.g.*, pseudouridine or 5-methoxyuridine. In another embodiment, the IVT polynucleotides can have a uniform chemical modification of two, three, or four of the same nucleoside type throughout the entire polynucleotide (such as all uridines and all cytosines, etc. are modified in the same way).

[0327] In some embodiments, the IVT polynucleotide can encode MCM and at least one additional peptide or polypeptide of interest. In another embodiment, the IVT polynucleotide can encode MCM and two or more peptides or polypeptides of interest. Non-limiting examples of peptides or polypeptides of interest include an enzyme and its substrate, a label and its binding molecule, a second messenger and its enzyme or the components of multimeric proteins or complexes.

[0328] In some embodiments, the IVT polynucleotide encodes an MCM protein or a functional fragment thereof. In some embodiments, the IVT polynucleotides of the disclosure comprise any one of the human MCM nucleic acid sequences selected from SEQ ID NOs: 1 to 207, 732 to 765, and 772. In some embodiments, the IVT polynucleotide encodes a human MCM

or functional fragment thereof comprising at least one amino acid mutation from the wild type sequence. In some embodiments, the IVT polynucleotide encodes an MCM mutant comprising one or more of the point mutations V69, T499, H532, A598, and V671. In some embodiments, the expression of the encoded polypeptide is increased. In some embodiments, the IVT polynucleotide increases MCM expression levels in cells when introduced into those cells, *e.g.*, by 20-50%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%.

Chimeric Polynucleotide Architecture

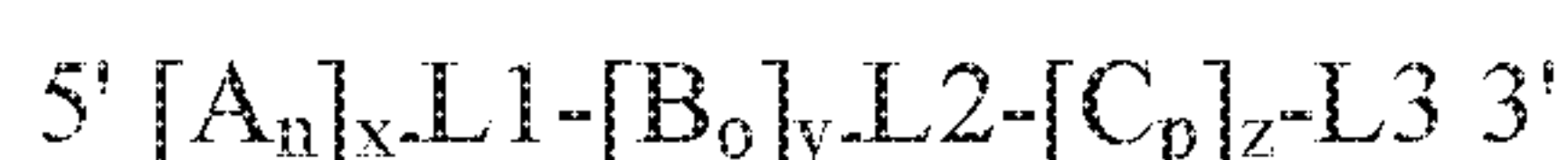
[0329] The disclosure also includes a chimeric polynucleotide comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0330] Polynucleotides which have portions or regions which differ in size and/or chemical modification pattern, chemical modification position, chemical modification percent or chemical modification population and combinations of the foregoing are known as "chimeric polynucleotides." A "chimera" according to the present disclosure is an entity having two or more incongruous or heterogeneous parts or regions. As used herein a "part" or "region" of a polynucleotide is defined as any portion of the polynucleotide which is less than the entire length of the polynucleotide. Chimeric polynucleotides which are modified mRNA molecules are termed "chimeric modified mRNA" or "chimeric mRNA."

[0331] Chimeric polynucleotides have portions or regions that differ in size and/or chemical modification pattern, chemical modification position, chemical modification percent or chemical modification population and combinations of the foregoing.

[0332] Examples of parts or regions, where the chimeric polynucleotide functions as an mRNA and encodes a polypeptide of interest include, but are not limited to, untranslated regions (UTRs, such as the 5' UTR or 3' UTR), coding regions, cap regions, polyA tail regions, start regions, stop regions, signal or target sequence regions, and combinations thereof.

[0333] In some embodiments, the chimeric polynucleotides of the disclosure have a structure comprising Formula I.



Formula I

wherein:

each of A and B independently comprise a region of linked nucleosides;

either A or B or both A and B encode MCM as described elsewhere herein;

C is an optional region of linked nucleosides;

at least one of regions A, B, or C is positionally modified, wherein said positionally modified region comprises at least two chemically modified nucleosides of one or more of the same nucleoside type of adenosine, thymidine, guanosine, cytidine, or uridine, and wherein at least two of the chemical modifications of nucleosides of the same type are different chemical modifications;

n, o and p are independently an integer between 15-10,000, representing the number of nucleosides in regions A, B, and C, respectively;

x and y are independently 1-20;

z is 0-5;

L1 and L2 are independently optional linker moieties, said linker moieties being either nucleic acid based or non-nucleic acid based; and

L3 is an optional conjugate or an optional linker moiety, said linker moiety being either nucleic acid based or non-nucleic acid based.

[0334] In some embodiments, at least one of the regions of linked nucleosides of A can comprise a sequence of linked nucleosides that can function as a 5' untranslated region (UTR). The sequence of linked nucleosides can be a natural or synthetic 5' UTR. As a non-limiting example, the chimeric polynucleotide can encode MCM and the sequence of linked nucleosides of A can encode the native 5' UTR of the MCM protein or a non-heterologous 5' UTR such as, but not limited to a synthetic UTR.

[0335] In another embodiment, at least one of the regions of linked nucleosides of A can be a cap region. The cap region can be located 5' to a region of linked nucleosides of A functioning as a 5'UTR. The cap region can comprise at least one cap such as, but not limited to, Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azido-guanosine, Cap2 and Cap4.

[0336] In some embodiments, the polynucleotide of the disclosure comprises a Cap1 5'UTR. In some embodiments, a polynucleotide comprises the Cap1 5'UTR, wherein the polynucleotide encodes human MCM or functional fragment thereof. In some embodiments, a polynucleotide comprising 5'UTR sequence, *e.g.*, Cap1, for encoding an MCM protein as disclosed herein increases expression of MCM compared to polynucleotides encoding MCM comprising a different 5'UTR (*e.g.*, Cap0, ARCA, inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-

guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azido-guanosine, Cap2 or Cap4). In some embodiments, polynucleotide comprising the Cap1 5'UTR, increases MCM expression levels in cells when introduced into those cells, *e.g.*, by at least 20%, *e.g.*, at least 20%, at least 25%, at least 35%, or at least 40%.

[0337] In some embodiments, at least one of the regions of linked nucleosides of C can comprise a sequence of linked nucleosides that can function as a 3' UTR. The sequence of linked nucleosides can be a natural or synthetic 3' UTR. As a non-limiting example, the chimeric polynucleotide can encode MCM and the sequence of linked nucleosides of C can encode the native 3' UTR of MCM or a non-heterologous 3' UTR such as, but not limited to a synthetic UTR.

[0338] In some embodiments, at least one of the regions of linked nucleosides of A comprises a sequence of linked nucleosides that functions as a 5' UTR and at least one of the regions of linked nucleosides of C comprises a sequence of linked nucleosides that functions as a 3' UTR. In some embodiments, the 5' UTR and the 3' UTR can be from the same or different species. In another embodiment, the 5' UTR and the 3' UTR can encode the native untranslated regions from different proteins from the same or different species.

[0339] Chimeric polynucleotides, including the parts or regions thereof, of the present disclosure can be classified as hemimers, gapmers, wingmers, or blockmers.

[0340] As used herein, a "hemimer" is a chimeric polynucleotide comprising a region or part that comprises half of one pattern, percent, position or population of a chemical modification(s) and half of a second pattern, percent, position or population of a chemical modification(s). Chimeric polynucleotides of the present disclosure can also comprise hemimer subregions. In some embodiments, a part or region is 50% of one and 50% of another.

[0341] In some embodiments, the entire chimeric polynucleotide can be 50% of one and 50% of the other. Any region or part of any chimeric polynucleotide of the disclosure can be a hemimer. Types of hemimers include pattern hemimers, population hemimers or position hemimers. By definition, hemimers are 50:50 percent hemimers.

[0342] As used herein, a "gapmer" is a chimeric polynucleotide having at least three parts or regions with a gap between the parts or regions. The "gap" can comprise a region of linked nucleosides or a single nucleoside that differs from the chimeric nature of the two parts or regions flanking it. The two parts or regions of a gapmer can be the same or different from each other.

[0343] As used herein, a "wingmer" is a chimeric polynucleotide having at least three parts or regions with a gap between the parts or regions. Unlike a gapmer, the two flanking parts or regions surrounding the gap in a wingmer are the same in degree or kind. Such similarity can be in the length of number of units of different modifications or in the number of modifications. The wings of a wingmer can be longer or shorter than the gap. The wing parts or regions can be 20, 30, 40, 50, 60 70, 80, 90 or 95% greater or shorter in length than the region that comprises the gap.

[0344] As used herein, a "blockmer" is a patterned polynucleotide where parts or regions are of equivalent size or number and type of modifications. Regions or subregions in a blockmer can be 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490 or 500, nucleosides long.

[0345] Chimeric polynucleotides, including the parts or regions thereof, of the present disclosure having a chemical modification pattern are referred to as "pattern chimeras." Pattern chimeras can also be referred to as blockmers. Pattern chimeras are those polynucleotides having a pattern of modifications within, across or among regions or parts.

[0346] Patterns of modifications within a part or region are those that start and stop within a defined region. Patterns of modifications across a part or region are those patterns that start in on part or region and end in another adjacent part or region. Patterns of modifications among parts or regions are those that begin and end in one part or region and are repeated in a different part or region, which is not necessarily adjacent to the first region or part.

[0347] The regions or subregions of pattern chimeras or blockmers can have simple alternating patterns such as ABAB[AB]_n where each "A" and each "B" represent different chemical modifications (at least one of the base, sugar or backbone linker), different types of chemical modifications (*e.g.*, naturally occurring and non-naturally occurring), different percentages of modifications or different populations of modifications. The pattern can repeat *n* number of times where *n*=3-300. Further, each A or B can represent from 1-2500 units (*e.g.*, nucleosides) in the pattern. Patterns can also be alternating multiples such as AABBAABB[AABB]_n (an alternating double multiple) or AAABBBAAABBB[AAABBB]_n (an alternating triple multiple) pattern. The pattern can repeat *n* number of times where *n*=3-300.

[0348] Different patterns can also be mixed together to form a second order pattern. For example, a single alternating pattern can be combined with a triple alternating pattern to form a second order alternating pattern A'B'. One example would be [ABABAB][AAABBBAAABBB] [ABABAB][AAABBBAAABBB] [ABABAB][AAABBBAAABBB], where [ABABAB] is A' and [AAABBBAAABBB] is B'. In like fashion, these patterns can be repeated *n* number of times, where *n*=3-300.

[0349] Patterns can include three or more different modifications to form an ABCABC[ABC]_n pattern. These three component patterns can also be multiples, such as AABBBCCAABBBCC[AABBBCC]_n and can be designed as combinations with other patterns such as ABCABCAABBBCCABCABCAABBBCC, and can be higher order patterns.

[0350] Regions or subregions of position, percent, and population modifications need not reflect an equal contribution from each modification type. They can form series such as "1-2-3-4," "1-2-4-8," where each integer represents the number of units of a particular modification type. Alternatively, they can be odd only, such as "1-3-3-1-3-1-5" or even only "2-4-2-4-6-4-8" or a mixture of both odd and even number of units such as "1-3-4-2-5-7-3-3-4".

[0351] Pattern chimeras can vary in their chemical modification by degree (such as those described above) or by kind (*e.g.*, different modifications).

[0352] Chimeric polynucleotides, including the parts or regions thereof, of the present disclosure having at least one region with two or more different chemical modifications of two or more nucleoside members of the same nucleoside type (A, C, G, T, or U) are referred to as "positionally modified" chimeras. Positionally modified chimeras are also referred to herein as "selective placement" chimeras or "selective placement polynucleotides". As the name implies, selective placement refers to the design of polynucleotides that, unlike polynucleotides in the art

where the modification to any A, C, G, T or U is the same by virtue of the method of synthesis, can have different modifications to the individual As, Cs, Gs, Ts or Us in a polynucleotide or region thereof. For example, in a positionally modified chimeric polynucleotide, there can be two or more different chemical modifications to any of the nucleoside types of As, Cs, Gs, Ts, or Us. There can also be combinations of two or more to any two or more of the same nucleoside type. For example, a positionally modified or selective placement chimeric polynucleotide can comprise 3 different modifications to the population of adenines in the molecule and also have 3 different modifications to the population of cytosines in the construct—all of which can have a unique, non-random, placement.

[0353] Chimeric polynucleotides, including the parts or regions thereof, of the present disclosure having a chemical modification percent are referred to as "percent chimeras." Percent chimeras can have regions or parts that comprise at least 1%, at least 2%, at least 5%, at least 8%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 99% positional, pattern or population of modifications. Alternatively, the percent chimera can be completely modified as to modification position, pattern, or population. The percent of modification of a percent chimera can be split between naturally occurring and non-naturally occurring modifications.

[0354] Chimeric polynucleotides, including the parts or regions thereof, of the present disclosure having a chemical modification population are referred to as "population chimeras." A population chimera can comprise a region or part where nucleosides (their base, sugar or backbone linkage, or combination thereof) have a select population of modifications. Such modifications can be selected from functional populations such as modifications that induce, alter or modulate a phenotypic outcome. For example, a functional population can be a population or selection of chemical modifications that increase the level of a cytokine. Other functional populations can individually or collectively function to decrease the level of one or more cytokines. Use of a selection of these like-function modifications in a chimeric polynucleotide would therefore constitute a "functional population chimera." As used herein, a "functional population chimera" can be one whose unique functional feature is defined by the population of modifications as described above or the term can apply to the overall function of the chimeric polynucleotide itself. For example, as a whole the chimeric polynucleotide can function in a different or superior way as compared to an unmodified or non-chimeric polynucleotide.

[0355] It should be noted that polynucleotides that have a uniform chemical modification of all of any of the same nucleoside type or a population of modifications produced by mere downward titration of the same starting modification in all of any of the same nucleoside type, or a measured percent of a chemical modification of all any of the same nucleoside type but with random incorporation, such as where all uridines are replaced by a uridine analog, *e.g.*, pseudouridine or 5-methoxyuridine, are not considered chimeric polynucleotides. Likewise, polynucleotides having a uniform chemical modification of two, three, or four of the same nucleoside type throughout the entire polynucleotide (such as all uridines and all cytosines, etc. are modified in the same way) are not considered chimeric polynucleotides. One example of a polynucleotide that is not chimeric is the canonical pseudouridine/5-methyl cytosine modified polynucleotide. These uniform polynucleotides are arrived at entirely via *in vitro* transcription (IVT) enzymatic synthesis; and due to the limitations of the synthesizing enzymes, they contain only one kind of modification at the occurrence of each of the same nucleoside type, *i.e.*, adenosine (A), thymidine (T), guanosine (G), cytidine (C) or uridine (U), found in the polynucleotide. Such polynucleotides can be characterized as IVT polynucleotides.

[0356] The chimeric polynucleotides of the present disclosure can be structurally modified or chemically modified. When the chimeric polynucleotides of the present disclosure are chemically and/or structurally modified, the polynucleotides can be referred to as "modified chimeric polynucleotides."

[0357] In some embodiments, the chimeric polynucleotides can encode two or more peptides or polypeptides of interest. Such peptides or polypeptides of interest include an enzyme and its substrate, a label and its binding molecule, a second messenger and its enzyme, or the components of multimeric proteins or complexes.

[0358] The regions or parts of the chimeric polynucleotides can be separated by a linker or spacer moiety. Such linkers or spaces can be nucleic acid based or non-nucleosidic.

[0359] In some embodiments, the chimeric polynucleotides can include a sequence encoding a self-cleaving peptide described herein, such as, but not limited to, a 2A peptide. The polynucleotide sequence of the 2A peptide in the chimeric polynucleotide can be modified or sequence-optimized by the methods described herein and/or are known in the art.

[0360] Notwithstanding the foregoing, the chimeric polynucleotides of the present disclosure can comprise a region or part that is not positionally modified or not chimeric as defined herein. For example, a region or part of a chimeric polynucleotide can be uniformly modified at one or

more A, T, C, G, or U, but the polynucleotides will not be uniformly modified throughout the entire region or part.

[0361] Regions or parts of chimeric polynucleotides can be, in some embodiments, from 15-10,000 nucleosides in length and, in some embodiments, a polynucleotide can have from 2-100 different regions or patterns of regions as described herein.

[0362] In some embodiments, chimeric polynucleotides encode one or more polypeptides of interest. In another embodiment, the chimeric polynucleotides are substantially non-coding. In another embodiment, the chimeric polynucleotides have both coding and non-coding regions and parts.

[0363] In some embodiments, regions or subregions of the polynucleotides can range from being absent to 500 nucleotides in length (*e.g.*, at least 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, or 500 nucleotides). Where the region is a polyA tail, the length can be determined in units of, or as a function of, polyA Binding Protein binding. In this embodiment, the polyA tail is long enough to bind at least 4 monomers of PolyA Binding Protein. PolyA Binding Protein monomers bind to stretches of approximately 38 nucleotides. As such, it has been observed that polyA tails of about 80 nucleotides to about 160 nucleotides are functional. The chimeric polynucleotides of the present disclosure that function as an mRNA need not comprise a polyA tail.

[0364] In some embodiments of the present disclosure, chimeric polynucleotides that function as an mRNA have a capping region. The capping region can comprise a single cap or a series of nucleotides forming the cap. In this embodiment the capping region can be from 1 to 10, *e.g.*, 2-9, 3-8, 4-7, 1-5, 5-10, or at least 2, or 10 or fewer nucleotides in length. In some embodiments, the cap is absent.

[0365] The present disclosure contemplates chimeric polynucleotides that are circular or cyclic. As the name implies circular polynucleotides are circular in nature meaning that the termini are joined in some fashion, whether by ligation, covalent bond, common association with the same protein or other molecule or complex or by hybridization.

[0366] Chimeric polynucleotides, formulations and compositions comprising chimeric polynucleotides, and methods of making, using and administering chimeric polynucleotides are also described in International Patent Application No. PCT/US2014/53907, the contents of which is incorporated by reference in its entirety.

[0367] In some embodiments, the chimeric polynucleotide encodes an MCM protein or a functional fragment thereof. In some embodiments, the chimeric polynucleotides of the disclosure comprise any one of the human MCM nucleic acid sequences selected from SEQ ID NOs: 1-207, 732-765, and 772. In some embodiments, the chimeric polynucleotide encodes a human MCM or functional fragment thereof comprising at least one amino acid mutation from the wild type sequence. In some embodiments, the chimeric polynucleotide encodes an MCM mutant comprising one or more of the point mutations V69, T499, H532, A598, and V671. In some embodiments, the expression of the encoded polypeptide is increased. In some embodiments, the chimeric polynucleotide increases MCM expression levels in cells when introduced into those cells, *e.g.*, by 20-50%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%.

Circular Polynucleotide Architecture

[0368] The disclosure also includes a circular polynucleotide comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0369] Polynucleotides that are circular are known as "circular polynucleotides" or "circP." As used herein, "circular polynucleotides" or "circP" means a single stranded circular polynucleotide which acts substantially like, and has the properties of, an RNA. The term "circular" is also meant to encompass any secondary or tertiary configuration of the circP.

[0370] The present disclosure contemplates polynucleotides encoding MCM that are circular or cyclic. As the name implies circular polynucleotides are circular in nature meaning that the termini are joined in some fashion, whether by ligation, covalent bond, common association with the same protein or other molecule or complex or by hybridization.

[0371] The circular polynucleotides or circPs that encode at least one peptide or polypeptide of interest are known as circular RNAs or circRNA. As used herein, "circular RNA" or "circRNA" means a circular polynucleotide that can encode at least one peptide or polypeptide of interest.

[0372] The circPs that comprise at least one sensor sequence and do not encode a peptide or polypeptide of interest are known as circular sponges or circSP. As used herein, "circular sponges," "circular polynucleotide sponges" or "circSP" means a circular polynucleotide that comprises at least one sensor sequence and does not encode a polypeptide of interest.

[0373] As used herein, "sensor sequence" means a receptor or pseudo-receptor for endogenous nucleic acid binding molecules. Non-limiting examples of sensor sequences include, microRNA binding sites, microRNA seed sequences, microRNA binding sites without the seed sequence, transcription factor binding sites and artificial binding sites engineered to act as pseudo-receptors and portions and fragments thereof.

[0374] The circPs that comprise at least one sensor sequence and encode at least one peptide or polypeptide of interest are known as circular RNA sponges or circRNA-SP. As used herein, "circular RNA sponges" or "circRNA-SP" means a circular polynucleotide that comprises at least one sensor sequence and at least one region encoding at least one peptide or polypeptide of interest.

[0375] As used herein, the term "circular construct" refers to a circular polynucleotide transcript that can act substantially similar to and have properties of a RNA molecule. In some embodiments, the circular construct acts as an mRNA. If the circular construct encodes one or more peptides or polypeptides of interest (*e.g.*, a circRNA or circRNA-SP) then the polynucleotide transcript retains sufficient structural and/or chemical features to allow the polypeptide of interest encoded therein to be translated. Circular constructs can be polynucleotides of the disclosure. When structurally or chemically modified, the construct can be referred to as a modified circP, modified circSP, modified circRNA or modified circRNA-SP.

[0376] Circular polynucleotides, formulations and compositions comprising circular polynucleotides, and methods of making, using and administering circular polynucleotides are also disclosed in International Patent Application No. PCT/US2014/53904 the contents of which is incorporated by reference in its entirety.

[0377] In some embodiments, the circular polynucleotide encodes an MCM protein or a functional fragment thereof. In some embodiments, the circular polynucleotides of the disclosure comprise any one of the human MCM nucleic acid selected from SEQ ID NOs: 1-207, 732-765, and 772. In some embodiments, the circular polynucleotide encodes a human MCM or functional fragment thereof comprising at least one amino acid mutation from the wild type sequence. In some embodiments, the circular polynucleotide encodes an MCM mutant comprising one or more of the point mutations V69, T499, H532, A598, and V671. In some embodiments, the expression of the encoded polypeptide is increased. In some embodiments, the circular polynucleotide increases MCM expression levels in cells when introduced into those cells, *e.g.*,

by 20-50%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%.

Multimers of Polynucleotides

[0378] The disclosure also includes multimers of polynucleotides comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0379] In some embodiments, multiple distinct chimeric polynucleotides and/or IVT polynucleotides can be linked together through the 3'-end using nucleotides that are modified at the 3'-terminus. Chemical conjugation can be used to control the stoichiometry of delivery into cells. For example, the glyoxylate cycle enzymes, isocitrate lyase and malate synthase, can be supplied into cells at a 1:1 ratio to alter cellular fatty acid metabolism. This ratio can be controlled by chemically linking chimeric polynucleotides and/or IVT polynucleotides using a 3'-azido terminated nucleotide on one polynucleotides species and a C5-ethynyl or alkynyl-containing nucleotide on the opposite polynucleotide species. The modified nucleotide is added post-transcriptionally using terminal transferase (New England Biolabs, Ipswich, MA) according to the manufacturer's protocol. After the addition of the 3'-modified nucleotide, the two polynucleotides species can be combined in an aqueous solution, in the presence or absence of copper, to form a new covalent linkage via a click chemistry mechanism as described in the literature.

[0380] In another example, more than two chimeric polynucleotides and/or IVT polynucleotides can be linked together using a functionalized linker molecule. For example, a functionalized saccharide molecule can be chemically modified to contain multiple chemical reactive groups (SH-, NH₂-, N₃, etc...) to react with the cognate moiety on a 3'-functionalized mRNA molecule (*i.e.*, a 3'-maleimide ester, 3'-NHS-ester, alkynyl). The number of reactive groups on the modified saccharide can be controlled in a stoichiometric fashion to directly control the stoichiometric ratio of conjugated chimeric polynucleotides and/or IVT polynucleotides.

[0381] In some embodiments, the chimeric polynucleotides and/or IVT polynucleotides can be linked together in a pattern. The pattern can be a simple alternating pattern such as CD[CD]_x where each "C" and each "D" represent a chimeric polynucleotide, IVT polynucleotide, different chimeric polynucleotides or different IVT polynucleotides. The pattern can repeat x number of

times, where $x = 1-300$. Patterns can also be alternating multiples such as CCDD[CCDD] $_x$ (an alternating double multiple) or CCCDDD[CCDDDD] $_x$ (an alternating triple multiple) pattern. The alternating double multiple or alternating triple multiple can repeat x number of times, where $x = 1-300$.

Conjugates and Combinations of Polynucleotides

[0382] The disclosure also includes conjugates and combinations of polynucleotides comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0383] In order to further enhance protein production, polynucleotides of the present disclosure can be designed to be conjugated to other polynucleotides, dyes, or other agents.

[0384] Conjugation can result in increased stability and/or half-life and can be particularly useful in targeting the polynucleotides to specific sites in the cell, tissue or organism.

[0385] In some embodiments, the polynucleotides can be administered with, conjugated to or further encode one or more of RNAi agents, siRNAs, shRNAs, miRNAs, miRNA binding sites, antisense RNAs, ribozymes, catalytic DNA, tRNA, RNAs that induce triple helix formation, aptamers or vectors, and the like.

Bifunctional Polynucleotides

[0386] The disclosure also includes bifunctional polynucleotides comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0387] In some embodiments of the disclosure are bifunctional polynucleotides (*e.g.*, bifunctional IVT polynucleotides, bifunctional chimeric polynucleotides or bifunctional circular polynucleotides). As the name implies, bifunctional polynucleotides are those having or capable of at least two functions. These molecules are also by convention be referred to as multi-functional.

[0388] The multiple functionalities of bifunctional polynucleotides can be encoded by the RNA (the function cannot manifest until the encoded product is translated) or can be a property of the polynucleotide itself. It can be structural or chemical. Bifunctional modified polynucleotides can comprise a function that is covalently or electrostatically associated with the

polynucleotides. Further, the two functions can be provided in the context of a complex of a chimeric polynucleotide and another molecule.

[0389] Bifunctional polynucleotides can encode peptides that are anti-proliferative. These peptides can be linear, cyclic, constrained or random coil. They can function as aptamers, signaling molecules, ligands or mimics or mimetics thereof. Anti-proliferative peptides can, as translated, be from 3 to 50 amino acids in length. They can be 5-40, 10-30, or approximately 15 amino acids long. They can be single chain, multichain or branched and can form complexes, aggregates or any multi-unit structure once translated.

Noncoding Polynucleotides

[0390] The disclosure also includes a noncoding polynucleotide comprising the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0391] The polynucleotides described herein can further comprise sequences that are partially or substantially not translatable, *e.g.*, having a noncoding region. As one non-limiting example, the noncoding region can be the first region of the IVT polynucleotide or the circular polynucleotide. Alternatively, the noncoding region can be a region other than the first region. As another non-limiting example, the noncoding region can be the A, B and/or C region of the chimeric polynucleotide.

[0392] Such molecules are generally not translated, but can exert an effect on protein production by one or more of binding to and sequestering one or more translational machinery components such as a ribosomal protein or a transfer RNA (tRNA), thereby effectively reducing protein expression in the cell or modulating one or more pathways or cascades in a cell that in turn alters protein levels. The polynucleotide can contain or encode one or more long noncoding RNA (lncRNA, or lincRNA) or portion thereof, a small nucleolar RNA (sno-RNA), micro RNA (miRNA), small interfering RNA (siRNA) or Piwi-interacting RNA (piRNA). Examples of such lncRNA molecules and RNAi constructs designed to target such lncRNA any of which can be encoded in the polynucleotides are disclosed in International Publication, WO2012/018881 A2, the contents of which are incorporated herein by reference in their entirety.

Cytotoxic Nucleosides

[0393] In some embodiments, the polynucleotides of the present disclosure (*i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide) can further incorporate one or more cytotoxic nucleosides.

Untranslated Regions (UTRs)

[0394] Untranslated regions (UTRs) are nucleic acid sections of a polynucleotide before a start codon (5'UTR) and after a stop codon (3'UTR) that are not translated. In some embodiments, a polynucleotide (e.g., a ribonucleic acid (RNA), e.g., a messenger RNA (mRNA)) of the disclosure comprising an open reading frame (ORF) encoding an MCM polypeptide further comprises UTR (e.g., a 5'UTR or functional fragment thereof, a 3'UTR or functional fragment thereof, or a combination thereof).

[0395] A UTR can be homologous or heterologous to the coding region in a polynucleotide. In some embodiments, the UTR is homologous to the ORF encoding the MCM polypeptide. In some embodiments, the UTR is heterologous to the ORF encoding the MCM polypeptide. In some embodiments, the polynucleotide comprises two or more 5'UTRs or functional fragments thereof, each of which have the same or different nucleotide sequences. In some embodiments, the polynucleotide comprises two or more 3'UTRs or functional fragments thereof, each of which have the same or different nucleotide sequences.

[0396] In some embodiments, the 5'UTR or functional fragment thereof, 3' UTR or functional fragment thereof, or any combination thereof is sequence optimized.

[0397] In some embodiments, the 5'UTR or functional fragment thereof, 3' UTR or functional fragment thereof, or any combination thereof comprises at least one chemically modified nucleobase, e.g., 5-methoxyuracil.

[0398] UTRs can have features that provide a regulatory role, e.g., increased or decreased stability, localization and/or translation efficiency. A polynucleotide comprising a UTR can be administered to a cell, tissue, or organism, and one or more regulatory features can be measured using routine methods. In some embodiments, a functional fragment of a 5'UTR or 3'UTR comprises one or more regulatory features of a full length 5' or 3' UTR, respectively.

[0399] Natural 5'UTRs bear features that play roles in translation initiation. They harbor signatures like Kozak sequences that are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus

CCR(A/G)CCAUGG, where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'. 5'UTRs also have been known to form secondary structures that are involved in elongation factor binding.

[0400] By engineering the features typically found in abundantly expressed genes of specific target organs, one can enhance the stability and protein production of a polynucleotide. For example, introduction of 5'UTR of liver-expressed mRNA, such as albumin, serum amyloid A, Apolipoprotein A/B/E, transferrin, alpha fetoprotein, erythropoietin, or Factor VIII, can enhance expression of polynucleotides in hepatic cell lines or liver. Likewise, use of 5'UTR from other tissue-specific mRNA to improve expression in that tissue is possible for muscle (e.g., MyoD, Myosin, Myoglobin, Myogenin, Herculin), for endothelial cells (e.g., Tie-1, CD36), for myeloid cells (e.g., C/EBP, AML1, G-CSF, GM-CSF, CD11b, MSR, Fr-1, i-NOS), for leukocytes (e.g., CD45, CD18), for adipose tissue (e.g., CD36, GLUT4, ACRP30, adiponectin) and for lung epithelial cells (e.g., SP-A/B/C/D).

[0401] In some embodiments, UTRs are selected from a family of transcripts whose proteins share a common function, structure, feature or property. For example, an encoded polypeptide can belong to a family of proteins (i.e., that share at least one function, structure, feature, localization, origin, or expression pattern), which are expressed in a particular cell, tissue or at some time during development. The UTRs from any of the genes or mRNA can be swapped for any other UTR of the same or different family of proteins to create a new polynucleotide.

[0402] In some embodiments, the 5'UTR and the 3'UTR can be heterologous. In some embodiments, the 5'UTR can be derived from a different species than the 3'UTR. In some embodiments, the 3'UTR can be derived from a different species than the 5'UTR.

[0403] Co-owned International Patent Application No. PCT/US2014/021522 (Publ. No. WO/2014/164253, incorporated herein by reference in its entirety) provides a listing of exemplary UTRs that can be utilized in the polynucleotide of the present disclosure as flanking regions to an ORF.

[0404] Exemplary UTRs of the application include, but are not limited to, one or more 5'UTR and/or 3'UTR derived from the nucleic acid sequence of: a globin, such as an α - or β -globin (e.g., a *Xenopus*, mouse, rabbit, or human globin); a strong Kozak translational initiation signal; a CYBA (e.g., human cytochrome b-245 α polypeptide); an albumin (e.g., human albumin7); a HSD17B4 (hydroxysteroid (17- β) dehydrogenase); a virus (e.g., a tobacco etch virus (TEV), a Venezuelan equine encephalitis virus (VEEV), a Dengue virus, a cytomegalovirus

(CMV) (e.g., CMV immediate early 1 (IE1)), a hepatitis virus (e.g., hepatitis B virus), a sindbis virus, or a PAV barley yellow dwarf virus); a heat shock protein (e.g., hsp70); a translation initiation factor (e.g., eIF4G); a glucose transporter (e.g., hGLUT1 (human glucose transporter 1)); an actin (e.g., human α or β actin); a GAPDH; a tubulin; a histone; a citric acid cycle enzyme; a topoisomerase (e.g., a 5'UTR of a TOP gene lacking the 5' TOP motif (the oligopyrimidine tract)); a ribosomal protein Large 32 (L32); a ribosomal protein (e.g., human or mouse ribosomal protein, such as, for example, rps9); an ATP synthase (e.g., ATP5A1 or the β subunit of mitochondrial H^+ -ATP synthase); a growth hormone (e.g., bovine (bGH) or human (hGH)); an elongation factor (e.g., elongation factor 1 α 1 (EEF1A1)); a manganese superoxide dismutase (MnSOD); a myocyte enhancer factor 2A (MEF2A); a β -F1-ATPase, a creatine kinase, a myoglobin, a granulocyte-colony stimulating factor (G-CSF); a collagen (e.g., collagen type I, alpha 2 (Col1A2), collagen type I, alpha 1 (Col1A1), collagen type VI, alpha 2 (Col6A2), collagen type VI, alpha 1 (Col6A1)); a ribophorin (e.g., ribophorin I (RPNI)); a low density lipoprotein receptor-related protein (e.g., LRP1); a cardiotrophin-like cytokine factor (e.g., Nnt1); calreticulin (Calr); a procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1 (Plod1); and a nucleobindin (e.g., Nucb1).

[0405] Other exemplary 5' and 3' UTRs include, but are not limited to, those described in Karikó et al., *Mol. Ther.* 2008 16(11):1833-1840; Karikó et al., *Mol. Ther.* 2012 20(5):948-953; Karikó et al., *Nucleic Acids Res.* 2011 39(21):e142; Strong et al., *Gene Therapy* 1997 4:624-627; Hansson et al., *J. Biol. Chem.* 2015 290(9):5661-5672; Yu et al., *Vaccine* 2007 25(10):1701-1711; Cafri et al., *Mol. Ther.* 2015 23(8):1391-1400; Andries et al., *Mol. Pharm.* 2012 9(8):2136-2145; Crowley et al., *Gene Ther.* 2015 Jun 30, doi:10.1038/gt.2015.68; Ramunas et al., *FASEB J.* 2015 29(5):1930-1939; Wang et al., *Curr. Gene Ther.* 2015 15(4):428-435; Holtkamp et al., *Blood* 2006 108(13):4009-4017; Kormann et al., *Nat. Biotechnol.* 2011 29(2):154-157; Poleganov et al., *Hum. Gen. Ther.* 2015 26(11):751-766; Warren et al., *Cell Stem Cell* 2010 7(5):618-630; Mandal and Rossi, *Nat. Protoc.* 2013 8(3):568-582; Holcik and Liebhaver, *PNAS* 1997 94(6):2410-2414; Ferizi et al., *Lab Chip.* 2015 15(17):3561-3571; Thess et al., *Mol. Ther.* 2015 23(9):1456-1464; Boros et al., *PLoS One* 2015 10(6):e0131141; Boros et al., *J. Photochem. Photobiol. B.* 2013 129:93-99; Andries et al., *J. Control. Release* 2015 217:337-344; Zinckgraf et al., *Vaccine* 2003 21(15):1640-9; Garneau et al., *J. Virol.* 2008 82(2):880-892; Holden and Harris, *Virology* 2004 329(1):119-133; Chiu et al., *J. Virol.* 2005 79(13):8303-8315; Wang et al., *EMBO J.* 1997 16(13):4107-4116; Al-Zoghaibi et al., *Gene* 2007

391(1-2):130-9; Vivinus et al., Eur. J. Biochem. 2001 268(7):1908-1917; Gan and Rhoads, J. Biol. Chem. 1996 271(2):623-626; Boado et al., J. Neurochem. 1996 67(4):1335-1343; Knirsch and Clerch, Biochem. Biophys. Res. Commun. 2000 272(1):164-168; Chung et al., Biochemistry 1998 37(46):16298-16306; Izquierdo and Cueva, Biochem. J. 2000 346 Pt 3:849-855; Dwyer et al., J. Neurochem. 1996 66(2):449-458; Black et al., Mol. Cell. Biol. 1997 17(5):2756-2763; Izquierdo and Cueva, Mol. Cell. Biol. 1997 17(9):5255-5268; US8278036; US8748089; US8835108; US9012219; US2010/0129877; US2011/0065103; US2011/0086904; US2012/0195936; US2014/020675; US2013/0195967; US2014/029490; US2014/0206753; WO2007/036366; WO2011/015347; WO2012/072096; WO2013/143555; WO2014/071963; WO2013/185067; WO2013/182623; WO2014/089486; WO2013/185069; WO2014/144196; WO2014/152659; 2014/152673; WO2014/152940; WO2014/152774; WO2014/153052; WO2014/152966, WO2014/152513; WO2015/101414; WO2015/101415; WO2015/062738, and WO2015/024667; the contents of each of which are incorporated herein by reference in their entirety.

[0406] In some embodiments, the 5'UTR is selected from the group consisting of a β -globin 5'UTR; a 5'UTR containing a strong Kozak translational initiation signal; a cytochrome b-245 α polypeptide (CYBA) 5'UTR; a hydroxysteroid (17- β) dehydrogenase (HSD17B4) 5'UTR; a Tobacco etch virus (TEV) 5'UTR; a Venezuelan equine encephalitis virus (VEEV) 5'UTR; a 5' proximal open reading frame of rubella virus (RV) RNA encoding nonstructural proteins; a Dengue virus (DEN) 5'UTR; a heat shock protein 70 (Hsp70) 5'UTR; a eIF4G 5'UTR; a GLUT1 5'UTR; functional fragments thereof and any combination thereof.

[0407] In some embodiments, the 3'UTR is selected from the group consisting of a β -globin 3'UTR; a CYBA 3'UTR; an albumin 3'UTR; a growth hormone (GH) 3'UTR; a VEEV 3'UTR; a hepatitis B virus (HBV) 3'UTR; α -globin 3'UTR; a DEN 3'UTR; a PAV barley yellow dwarf virus (BYDV-PAV) 3'UTR; an elongation factor 1 α 1 (EEF1A1) 3'UTR; a manganese superoxide dismutase (MnSOD) 3'UTR; a β subunit of mitochondrial H(+)-ATP synthase (β -mRNA) 3'UTR; a GLUT1 3'UTR; a MEF2A 3'UTR; a β -F1-ATPase 3'UTR; functional fragments thereof and combinations thereof.

[0408] Other exemplary UTRs include, but are not limited to, one or more of the UTRs, including any combination of UTRs, disclosed in WO2014/164253, the contents of which are incorporated herein by reference in their entirety. Shown in Table 21 of U.S. Provisional Application No. 61/775,509 and in Table 22 of U.S. Provisional Application No. 61/829,372, the

contents of each are incorporated herein by reference in their entirety, is a listing start and stop sites for 5'UTRs and 3'UTRs. In Table 21, each 5'UTR (5'-UTR-005 to 5'-UTR 68511) is identified by its start and stop site relative to its native or wild-type (homologous) transcript (ENST; the identifier used in the ENSEMBL database).

[0409] Wild-type UTRs derived from any gene or mRNA can be incorporated into the polynucleotides of the disclosure. In some embodiments, a UTR can be altered relative to a wild type or native UTR to produce a variant UTR, e.g., by changing the orientation or location of the UTR relative to the ORF; or by inclusion of additional nucleotides, deletion of nucleotides, swapping or transposition of nucleotides. In some embodiments, variants of 5' or 3' UTRs can be utilized, for example, mutants of wild type UTRs, or variants wherein one or more nucleotides are added to or removed from a terminus of the UTR.

[0410] Additionally, one or more synthetic UTRs can be used in combination with one or more non-synthetic UTRs. *See*, e.g., Mandal and Rossi, Nat. Protoc. 2013 8(3):568-82, the contents of which are incorporated herein by reference in their entirety, and sequences available at www.addgene.org/Derrick_Rossi/. UTRs or portions thereof can be placed in the same orientation as in the transcript from which they were selected or can be altered in orientation or location. Hence, a 5' and/or 3' UTR can be inverted, shortened, lengthened, or combined with one or more other 5' UTRs or 3' UTRs.

[0411] In some embodiments, the polynucleotide comprises multiple UTRs, e.g., a double, a triple or a quadruple 5'UTR or 3'UTR. For example, a double UTR comprises two copies of the same UTR either in series or substantially in series. For example, a double beta-globin 3'UTR can be used (see US2010/0129877, the contents of which are incorporated herein by reference in its entirety).

[0412] Tables 5 and 6 provide a listing of exemplary UTRs that can be utilized in the polynucleotides of the present disclosure. Shown in Table 5 is a listing of a 5'-untranslated region of the disclosure. Variants of 5' UTRs can be utilized wherein one or more nucleotides are added or removed to the termini, including A, T, C or G.

Table 5. 5'-Untranslated Regions

5' UTR Identifier	Name/Description	Sequence	SEQ ID NO.
5UTR-001	Upstream UTR	GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATAT AAGAGCCACC	215
5UTR-002	Upstream UTR	GGGAGATCAGAGAGAAAAGAAGAGTAAGAAGAAATAT AAGAGCCACC	216

5UTR-003	Upstream UTR	GGAATAAAAGTCTCAACACAACATATACAAAACAAAC GAATCTCAAGCAATCAAGCATTCTACTTCTATTGCAG CAATTTAAATCATTTCTTTTAAAGCAAAAGCAATTTT CTGAAAATTTTCACCATTACGAACGATAGCAAC	217
5UTR-004	Upstream UTR	GGGAGACAAGCUUGGCAUUCGGGUACUGUUGGUAAG CCACC	218
5UTR-005	Upstream UTR	GGGAGATCAGAGAGAAAAGAAGAGTAAGAAGAAATAT AAGAGCCACC	219
5UTR-006	Upstream UTR	GGAATAAAAGTCTCAACACAACATATACAAAACAAAC GAATCTCAAGCAATCAAGCATTCTACTTCTATTGCAG CAATTTAAATCATTTCTTTTAAAGCAAAAGCAATTTT CTGAAAATTTTCACCATTACGAACGATAGCAAC	220
5UTR-007	Upstream UTR	GGGAGACAAGCUUGGCAUUCGGGUACUGUUGGUAAG CCACC	221
5UTR-008	Upstream UTR	GGGAATTAACAGAGAAAAGAAGAGTAAGAAGAAATAT AAGAGCCACC	222
5UTR-009	Upstream UTR	GGGAAATTAGACAGAAAAGAAGAGTAAGAAGAAATAT AAGAGCCACC	223
5UTR-010	Upstream UTR	GGGAAATAAGAGAGTAAAGAACAGTAAGAAGAAATAT AAGAGCCACC	224
5UTR-011	Upstream UTR	GGGAAAAAAGAGAGAAAAGAAGACTAAGAAGAAATAT AAGAGCCACC	225
5UTR-012	Upstream UTR	GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGATATAT AAGAGCCACC	226
5UTR-013	Upstream UTR	GGGAAATAAGAGACAAAACAAGAGTAAGAAGAAATAT AAGAGCCACC	227
5UTR-014	Upstream UTR	GGGAAATTAGAGAGTAAAGAACAGTAAGTAGAATTAA AAGAGCCACC	228
5UTR-015	Upstream UTR	GGGAAATAAGAGAGAATAGAAGAGTAAGAAGAAATAT AAGAGCCACC	229
5UTR-016	Upstream UTR	GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAAATT AAGAGCCACC	230
5UTR-017	Upstream UTR	GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATTT AAGAGCCACC	231
5UTR-018	Upstream UTR	TCAAGCTTTTGGACCCTCGTACAGAAGCTAATACGAC TCACTATAGGGAAATAAGAGAGAAAAGAAGAGTAAGA AGAAATATAAGAGCCACC	266
142-3p UTR-001	Upstream UTR including miR142-3p	TGATAATAGTCCATAAAGTAGGAAACACTACAGCTGG AGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCC CCCCAGCCCCCTCCTCCCCTTCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	725
142-3p UTR-002	Upstream UTR including miR142-3p	TGATAATAGGCTGGAGCCTCGGTGGCTCCATAAAGTA GGAAACACTACACATGCTTCTTGCCCCTTGGGCCTCC CCCCAGCCCCCTCCTCCCCTTCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	726
142-3p UTR-003	Upstream UTR including miR142-3p	TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTG CCCCTTCCATAAAGTAGGAAACACTACATGGGCCTCC CCCCAGCCCCCTCCTCCCCTTCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	727
142-3p UTR-004	Upstream UTR including miR142-3p	TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTG CCCCTTGGGCCTCCCCCAGTCCATAAAGTAGGAAAC ACTACACCCCTCCTCCCCTTCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	728
142-3p UTR-005	Upstream UTR including miR142-3p	TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTG CCCCTTGGGCCTCCCCCAGCCCCCTCCTCCCCTTCTC CATAAAGTAGGAAACACTACACTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	729

142-3p UTR-006	Upstream UTR including miR142-3p	TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTG CCCCCTGGGCCTCCCCCAGCCCCCTCCTCCCCTTCCT GCACCCGTACCCCTCCATAAAGTAGGAAACACTACA GTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	730
142-3p UTR-007	Upstream UTR including miR142-3p	TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTG CCCCCTGGGCCTCCCCCAGCCCCCTCCTCCCCTTCCT GCACCCGTACCCCGTGGTCTTTGAATAAAGTTCCAT AAAGTAGGAAACACTACACTGAGTGGGCGGC	731

[0413] In a particular embodiment, the 5' UTR useful for the polynucleotides comprises SEQ ID NO: 266.

[0414] Shown in Table 6 is a listing of 3'-untranslated regions of the disclosure. Variants of 3' UTRs can be utilized wherein one or more nucleotides are added or removed to the termini, including A, T, C or G.

Table 6. 3'-Untranslated Regions

3' UTR Identifier	Name/ Description	Sequence	SEQ ID NO.
3UTR-001	Creatine Kinase	GCGCCTGCCACCTGCCACCGACTGCTGGAACCCAGCCAGTGG GAGGGCCTGGCCCACCAGAGTCCTGCTCCCTCACTCCTCGCCC CGCCCCCTGTCCAGAGTCCCACCTGGGGGCTCTCTCCACCCT TCTCAGAGTTCCAGTTTCAACCAGAGTTCCAACCAATGGGCTC CATCCTCTGGATTCTGGCCAATGAAATATCTCCCTGGCAGGGT CCTCTTCTTTTCCCAGAGCTCCACCCCAACCAGGAGCTCTAGT TAATGGAGAGCTCCCAGCACACTCGGAGCTTGTGCTTTGTCTC CACGCAAAGCGATAAATAAAAGCATTGGTGGCCTTTGGTCTTT GAATAAAGCCTGAGTAGGAAGTCTAGA	232
3UTR-002	Myoglobin	GCCCCTGCCGCTCCCACCCCCACCCATCTGGGCCCCGGGTTCA AGAGAGAGCGGGGTCTGATCTCGTGTAGCCATATAGAGTTTGC TTCTGAGTGTCTGCTTTGTTTAGTAGAGGTGGGCAGGAGGAGC TGAGGGGCTGGGGCTGGGGTGTGTAAGTTGGCTTTGCATGCCC AGCGATGCGCCTCCCTGTGGGATGTCATCACCCCTGGGAACCGG GAGTGGCCCTTGGCTCACTGTGTTCTGCATGGTTTGGATCTGA ATTAATTGTCCTTTCTTCTAAATCCCAACCGAACTTCTTCCAA CCTCCAAACTGGCTGTAACCCCAAATCCAAGCCATTAATACTACA CCTGACAGTAGCAATTGTCTGATTAATCACTGGCCCCTTGAAG ACAGCAGAATGTCCCTTTGCAATGAGGAGGAGATCTGGGCTGG GCGGGCCAGCTGGGGAAGCATTGACTATCTGGAACCTTGTGTG TGCCTCCTCAGGTATGGCAGTGACTCACCTGGTTTTAATAAAA CAACCTGCAACATCTCATGGTCTTTGAATAAAGCCTGAGTAGG AAGTCTAGA	233
3UTR-003	α -actin	ACACACTCCACCTCCAGCACGCGACTTCTCAGGACGACGAATC TTCTCAATGGGGGGGCGGCTGAGCTCCAGCCACCCCGCAGTCA CTTTCTTTGTAACAACTTCCGTTGCTGCCATCGTAAACTGACA CAGTGTTTATAACGTGTACATACATTAACCTATTACCTCATTT TGTTATTTTTTCGAAACAAAGCCCTGTGGAAGAAAATGGAAAC TTGAAGAAGCATTAAAGTCATTCTGTTAAGCTGCGTAAATGGT CTTTGAATAAAGCCTGAGTAGGAAGTCTAGA	234
3UTR-004	Albumin	CATCACATTTAAAGCATCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAAAGCTTATTCATCTGTTTTTCTTTTT CGTTGGTGTAAGCCAACACCCTGTCTAAAAAACATAAATTTTC	235

		TTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAA AATGGAAAGAATCTAATAGAGTGGTACAGCACTGTTATTTTTC AAAGATGTGTTGCTATCCTGAAAATTCTGTAGGTTCTGTGGAA GTTCCAGTGTTCTCTCTTATTCCACTTCGGTAGAGGATTTCTA GTTTCTTGTGGGCTAATTAAATAAATCATTAATACTCTTCTAA TGGTCTTTGAATAAAGCCTGAGTAGGAAGTCTAGA	
3UTR-005	α -globin	GCTGCCTTCTGCGGGGCTTGCCTTCTGGCCATGCCCTTCTTCT CTCCCTTGCACCTGTACCTCTTGGTCTTTGAATAAAGCCTGAG TAGGAAGGCGGGCCGCTCGAGCATGCATCTAGA	236
3UTR-006	G-CSF	GCCAAGCCCTCCCCATCCCATGTATTTATCTCTATTTAATATT TATGTCTATTTAAGCCTCATATTTAAAGACAGGGAAGAGCAGA ACGGAGCCCCAGGCCTCTGTGTCCTTCCCTGCATTTCTGAGTT TCATTCTCCTGCCTGTAGCAGTGAGAAAAAGCTCCTGTCCTCC CATCCCCTGGACTGGGAGGTAGATAGGTAAATACCAAGTATTT ATTACTATGACTGCTCCCCAGCCCTGGCTCTGCAATGGGCACT GGGATGAGCCGCTGTGAGCCCCCTGGTCTGAGGGTCCCCACCT GGGACCCTTGAGAGTATCAGGTCTCCACGTGGGAGACAAGAA ATCCCTGTTTAATATTTAAACAGCAGTGTTCCCCATCTGGGTC CTTGCACCCCTCACTCTGGCCTCAGCCGACTGCACAGCGGCCC CTGCATCCCCTTGGCTGTGAGGCCCTGGACAAGCAGAGGTGG CCAGAGCTGGGAGGCATGGCCCTGGGGTCCACGAATTTGCTG GGGAATCTCGTTTTTCTTCTTAAGACTTTTGGGACATGGTTTG ACTCCCGAACATCACCGACGCGTCTCCTGTTTTTCTGGGTGGC CTCGGGACACCTGCCCTGCCCCCACGAGGGTCAGGACTGTGAC TCTTTTTTAGGGCCAGGCAGGTGCCTGGACATTTGCCTTGCTGG ACGGGGACTGGGGATGTGGGAGGGAGCAGACAGGAGGAATCAT GTCAGGCCTGTGTGTGAAAGGAAGCTCCACTGTCACCCTCCAC CTCTTCACCCCCCACTCACCAGTGTCCTCCACTGTCACATT GTAAGTGAAGTTTCAAGGATAATAAAGTGTTTGCCTCCATGGTCT TTGAATAAAGCCTGAGTAGGAAGGCGGGCCGCTCGAGCATGCAT CTAGA	237
3UTR-007	Colla2; collagen, type I, alpha 2	ACTCAATCTAAATTAAAAAAGAAAGAAATTTGAAAAAATTTTC TCTTTGCCATTTCTTCTTCTTCTTTTTTTAACTGAAAGCTGAAT CCTTCCATTTCTTCTGCACATCTACTTGCTTAAATTGTGGGCA AAAGAGAAAAAGAAGGATTGATCAGAGCATTGTGCAATACAGT TTCATTAACTCCTTCCCCCGCTCCCCCAAAATTTGAATTTTT TTTTCAACACTCTTACACCTGTTATGGAAAATGTCAACCTTTG TAAGAAAACCAAAATAAAAAATTGAAAAATAAAAACCATAAACA TTTGCACCACTTGTGGCTTTTGAATATCTTCCACAGAGGGAAG TTTAAAACCCAAACTTCCAAAGGTTTAAACTACCTCAAAACAC TTTCCCATGAGTGTGATCCACATTGTTAGGTGCTGACCTAGAC AGAGATGAACTGAGGTCCTTGTTTTGTTTTGTTTATAAATAACA AGGTGCTAATTAATAGTATTTTCAAGATACTTGAAGAATGTTGAT GGTGCTAGAAGAATTTGAGAAGAAATACTCCTGTATTGAGTTG TATCGTGTGGTGTATTTTTTAAAAAATTTGATTTAGCATTCAT ATTTTCCATCTTATTCCCAATTAAAAGTATGCAGATTATTTGC CCAAATCTTCTTCAGATTCAGCATTTGTTCTTTGCCAGTCTCA TTTTCATCTTCTTCCATGGTTCCACAGAAGCTTTGTTTCTTGG GCAAGCAGAAAAATTAAATTGTACCTATTTTGTATATGTGAGA TGTTTAAATAAATTGTGAAAAAATGAAATAAAGCATGTTTGG TTTTCCAAAAGAACATAT	238
3UTR-008	Col6a2; collagen, type VI, alpha 2	CGCCGCGCGCCCGGGCCCCGCGAGTGGAGGGTCGTGAGCCCACCC CGTCCATGGTGCTAAGCGGGCCCCGGGTCCCACACGGCCAGCAC CGCTGCTCACTCGGACGACGCCCTGGGCCTGCACCTCTCCAGC TCCTCCCACGGGGTCCCCGTAGCCCCGGCCCCCGCCAGCCCC AGGTCTCCCCAGGCCCTCCGCAGGCTGCCCGGCCTCCCTCCCC CTGCAGCCATCCCAAGGCTCCTGACCTACCTGGCCCCCTGAGCT CTGGAGCAAGCCCTGACCCAATAAAGGCTTTGAACCCAT	239

3UTR-009	RPN1; ribophorin I	GGGGCTAGAGCCCTCTCCGCACAGCGTGGAGACGGGGCAAGGA GGGGGGTTATTAGGATTGGTGGTTTTGTTTTGCTTTGTTTTAAA GCCGTGGGAAAATGGCACAACCTTTACCTCTGTGGGAGATGCAA CACTGAGAGCCAAGGGGTGGGAGTTGGGATAATTTTTTATATAA AAGAAGTTTTTCCACTTTGAATTGCTAAAAGTGGCATTTTTTCC TATGTGCAGTCACTCCTCTCATTTCTAAAATAGGGACGTGGCC AGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGC CGAGGCAGGCGGCTCACGAGGTCAGGAGATCGAGACTATCCTG GCTAACACGGTAAAACCCCTGTCTCTACTAAAAGTACAAAAAAT TAGCTGGGCGTGGTGGTGGGCACCTGTAGTCCCAGCTACTCGG GAGGCTGAGGCAGGAGAAAGGCATGAATCCAAGAGGCAGAGCT TGCAGTGAGCTGAGATCACGCCATTGCACTCCAGCCTGGGCAA CAGTGTTAAGACTCTGTCTCAAATATAAATAAATAAATAAATA AATAAATAAATAAATAAATAAAGCGAGATGTTGCCCTCAAA	240
3UTR-010	LRP1; low density lipoprotein receptor- related protein 1	GGCCCTGCCCCGTCGGACTGCCCCAGAAAGCCTCCTGCCCC TGCCAGTGAAGTCCTTCAGTGAGCCCCCTCCCAGCCAGCCCTT CCCTGGCCCCGCCGGATGTATAAATGTAAAATGAAGGAATTA CATTTTATATGTGAGCGAGCAAGCCGGCAAGCGAGCACAGTAT TATTTCTCCATCCCCCTCCCTGCCTGCTCCTTGGCACCCCCATG CTGCCTTCAGGGAGACAGGCAGGGAGGGCTTGGGGCTGCACCT CCTACCCTCCCACCAGAACGCACCCCACTGGGAGAGCTGGTGG TGCAGCCTTCCCCTCCCTGTATAAGACACTTTGCCAAGGCTCT CCCCCTCTGCCCCATCCCTGCTTGCCCGCTCCCACAGCTTCCT GAGGGCTAATTCTGGGAAGGGAGAGTTCTTTGCTGCCCCCTGTC TGGAAGACGTGGCTCTGGGTGAGGTAGGCGGGAAAGGATGGAG TGTTTTAGTTCTTGGGGGAGGCCACCCCAAACCCCAAGCCCCAA CTCCAGGGGCACCTATGAGATGGCCATGCTCAACCCCCCTCCC AGACAGGCCCTCCCTGTCTCCAGGGCCCCCACCGAGGTTCCCA GGGCTGGAGACTTCCTCTGGTAAACATTCCCTCCAGCCTCCCCT CCCCTGGGGACGCCAAGGAGGTGGGCCACACCCAGGAAGGGAA AGCGGGCAGCCCCGTTTTTGGGGACGTGAACGTTTTAATAATTT TTGCTGAATTCCTTTACAATAAATAACACAGATAATTGTTATA AATAAAATTGT	241
3UTR-011	Nnt1; cardiotrophin-like cytokine factor 1	ATATTAAGGATCAAGCTGTTAGCTAATAATGCCACCTCTGCAG TTTTTGGGAACAGGCAAATAAAGTATCAGTATACATGGTGATGT ACATCTGTAGCAAAGCTCTTGGAGAAAATGAAGACTGAAGAAA GCAAAGCAAAAACCTGTATAGAGAGATTTTTTCAAAGCAGTAAT CCCTCAATTTTAAAAAAGGATTGAAAATTCTAAATGTCTTTCT GTGCATATTTTTTGTGTTAGGAATCAAAAGTATTTTATAAAAG GAGAAAGAACAGCCTCATTTTAGATGTAGTCCTGTTGGATTTT TTATGCCTCCTCAGTAACCAGAAATGTTTTAAAAAACTAAGTG TTTAGGATTTCAAGACAACATTATACATGGCTCTGAAATATCT GACACAATGTAAACATTGCAGGCACCTGCATTTTATGTTTTTT TTTTCAACAAATGTGACTAATTTGAAACTTTTATGAACTTCTG AGCTGTCCCCCTTGCAATTCAACCGCAGTTTGAATTAATCATAT CAAATCAGTTTTTAATTTTTTAAATTGTACTTCAGAGTCTATAT TTCAAGGGCACATTTTCTCACTACTATTTTAATACATTAAAGG ACTAAATAATCTTTCAGAGATGCTGGAAACAAATCATTTGCTT TATATGTTTCATTAGAATACCAATGAAACATACAACCTTGAAAA TTAGTAATAGTATTTTTTGAAGATCCCATTTCTAATTGGAGATC TCTTTAATTTTCGATCAACTTATAATGTGTAGTACTATATTAAG TGCACCTGAGTGGAATTCAACATTTGACTAATAAAATGAGTTC ATCATGTTGGCAAGTGATGTGGCAATTATCTCTGGTGACAAAA GAGTAAATCAAATATTTCTGCCTGTTACAAATATCAAGGAAG ACCTGCTACTATGAAATAGATGACATTAATCTGTCTTCACTGT TTATAATACGGATGGATTTTTTTTTCAAATCAGTGTGTGTTTTG AGGTCTTATGTAATTGATGACATTTGAGAGAAATGGTGGCTTT TTTTAGCTACCTCTTTGTTTCATTTAAGCACCAGTAAAGATCAT	242

		GTCTTTTTATAGAAGTGTAGATTTTCTTTGTGACTTTGCTATC GTGCCTAAAGCTCTAAATATAGGTGAATGTGTGATGAATACTC AGATTATTTGTCTCTCTATATAATTAGTTTGGTACTAAGTTTC TCAAAAATTATTAACACATGAAAGACAATCTCTAAACCAGAA AAAGAAGTAGTACAAATTTTGTACTGTAATGCTCGCGTTTAG TGAGTTTAAACACACAGTATCTTTTGGTTTATAATCAGTTT CTATTTTGCTGTGCCTGAGATTAAGATCTGTGTATGTGTGTGT GTGTGTGTGTGCGTTTGTGTGTTAAAGCAGAAAAGACTTTTTT AAAAGTTTAAAGTGATAAATGCAATTTGTTAATTGATCTTAGA TCACTAGTAACTCAGGGCTGAATTATACCATGTATATTCTAT TAGAAGAAAGTAAACACCATCTTTATTCCTGCCCTTTTTCTTC TCTCAAAGTAGTTGTAGTTATATCTAGAAAGAAGCAATTTTGA TTTCTTGAAAAGGTAGTTCCTGCACTCAGTTTAAACTAAAAAT AATCATACTTGGATTTTATTTATTTTGTGCATAGTAAAAATTT TAATTTATATATATTTTATTTAGTATTATCTTATTCTTTGCT ATTTGCCAATCCTTTGTGCATCAATTGTGTTAAATGAATTGAAA ATTCATGCCCTGTTCAATTTTATTTTACTTTATTGGTTAGGATA TTTAAAGGATTTTGTATATATAATTTCTTAAATTAATATTCC AAAAGGTTAGTGGACTTAGATTATAAATTATGGCAAAAATCTA AAAACAACAAAAATGATTTTATACATTTCTATTTCAATTATTCC TCTTTTCCAATAAGTCATACAATTGGTAGATATGACTTATTT TATTTTGTATTATTTCACTATATCTTTATGATATTTAAGTATA AATAATTAAAAAATTTATTGTACCTTATAGTCTGTCACCAAA AAAAAAATTATCTGTAGGTAGTGAAATGCTAATGTTGATTT GTCTTTAAGGGCTTGTTAACTATCCTTTATTTTCTCATTTGTC TTAAATTAGGAGTTTGTGTTTAAATTACTCATCTAAGCAAAAA ATGTATATAAATCCCATTACTGGGTATATACCCAAAGGATTAT AAATCATGCTGCTATAAAGACACATGCACACGTATGTTTATTG CAGCACTATTCACAATAGCAAAGACTTGGAACCAACCCAAATG TCCATCAATGATAGACTTGATTAAGAAAATGTGCACATATACA CCATGGAATACTATGCAGCCATAAAAAAGGATGAGTTCATGTC CTTTGTAGGGACATGGATAAAGCTGGAAACCATCAFTCTGAGC AAATATTGCAAGGACAGAAAACCAAACACTGCATGTTCTCAC TCATAGGTGGGAATTGAACAATGAGAACACTTGGACACAAGGT GGGGAACACCACACACCAGGGCCTGTCATGGGGTGGGGGGAGT GGGGAGGGATAGCATTAGGAGATATACCTAATGTAAATGATGA GTTAATGGGTGCAGCACACCAACATGGCACATGTATACATATG TAGCAAACCTGCACGTTGTGCACATGTACCCTAGAACTTAAAG TATAATTAAAAAAGAAAACAGAAAGCTATTTATAAAGAA GTTATTTGCTGAAATAAATGTGATCTTTCCCATTAATAAATA AAGAAATTTGGGGTAAAAAACACAATATATTGTATTCTTGA AAAATTCTAAGAGAGTGGATGTGAAGTGTTCTCACCACAAAAG TGATAACTAATTGAGGTAATGCACATATTAATTAGAAAGATTT TGTCATTCCACAATGTATATATACTTAAAAATATGTTATACAC AATAAATACATACATTAATAAATAAGTAAATGTA	
3UTR-012	Col6a1; collagen, type VI, alpha 1	CCCACCCTGCACGCCGGCACCAAACCCTGTCCTCCCACCCTC CCCACTCATCACTAAACAGAGTAAATGTGATGCGAATTTTCC CGACCAACCTGATTGCTAGATTTTTTTTTTAAGGAAAAGCTTGG AAAGCCAGGACACAACGCTGCTGCCTGCTTTGTGCAGGGTCCT CCGGGGCTCAGCCCTGAGTTGGCATCACCTGCGCAGGGCCCTC TGGGGCTCAGCCCTGAGCTAGTGTACCTGCACAGGGCCCTCT GAGGCTCAGCCCTGAGCTGGCGTCACCTGTGCAGGGCCCTCTG GGGCTCAGCCCTGAGCTGGCCTCACCTGGGTTCACCCACCCCGG GCTCTCCTGCCCTGCCCTCCTGCCCGCCCTCCCTCCTGCCTGC GCAGCTCCTTCCCTAGGCACCTCTGTGCTGCATCCCACCAGCC TGAGCAAGACGCCCTCTCGGGGCTGTGCCGCACTAGCCTCCC TCTCCTCTGTCCCATAGCTGGTTTTTCCCACCAATCCTCACC TAACAGTACTTTACAATTAACTCAAAGCAAGCTCTTCTCCT	243

		CAGCTTGGGGCAGCCATTGGCCTCTGTCTCGTTTTGGGAAACC AAGGTCAGGAGGCCGTTGCAGACATAAATCTCGGCGACTCGGC CCCGTCTCCTGAGGGTCCTGCTGGTGACCGGCCTGGACCTTGG CCCTACAGCCCTGGAGGCCGCTGCTGACCAGCACTGACCCCGA CCTCAGAGAGTACTCGCAGGGGCGCTGGCTGCACTCAAGACCC TCGAGATTAACGGTGCTAACCCCGTCTGCTCCTCCCTCCCGCA GAGACTGGGGCCTGGACTGGACATGAGAGCCCCCTTGGTGCCAC AGAGGGCTGTGTCTTACTAGAAACAACGCAAACCTCTCCTTCC TCAGAATAGTGATGTGTTGACGTTTTATCAAAGGCCCCCTTT CTATGTTTCATGTTAGTTTTGCTCCTTCTGTGTTTTTTCTGAA CCATATCCATGTTGCTGACTTTTCCAAATAAAGGTTTTCACTC CTCTC	
3UTR-013	Calr; calreticulin	AGAGGCCTGCCTCCAGGGCTGGACTGAGGCCTGAGCGCTCCTG CCGCAGAGCTGGCCGCGCCAAATAATGTCTCTGTGAGACTCGA GAACTTTCATTTTTTTCCAGGCTGGTTCGGATTTGGGGTGGAT TTTGGTTTTGTTCCTCCCTCCACTCTCCCCACCCCCCTCCCC GCCCTTTTTTTTTTTTTTTTTTTTAACTGGTATTTTATCTTTGA TTCTCCTTCAGCCCTCACCCCTGGTTCTCATCTTTCTTGATCA ACATCTTTTCTTGCTCTGTCCCCTTCTCTCATCTCTTAGCTC CCCTCCAACCTGGGGGGCAGTGGTGTGGAGAAGCCACAGGCCT GAGATTTTCATCTGCTCTCCTTCCTGGAGCCCAGAGGAGGGCAG CAGAAGGGGGTGGTGTCTCCAACCCCCCAGCACTGAGGAAGAA CGGGGCTCTTCTCATTTACCCCTCCCTTTCTCCCCTGCCCCC AGGACTGGGCCACTTCTGGGTGGGGCAGTGGGTCCCAGATTGG CTCACACTGAGAATGTAAGAACTACAAACAAAATTTCTATTAA ATTAAATTTTGTGTCTCC	244
3UTR-014	Colla1; collagen, type I, alpha 1	CTCCCTCCATCCCAACCTGGCTCCCTCCCACCCAACCAACTTT CCCCCAACCCGGAAACAGACAAGCAACCCAAACTGAACCCCC TCAAAAGCCAAAAAATGGGAGACAATTTACATGGACTTTGGA AAATATTTTTTTTCTTTGCATTCTCTCAAACCTTAGTTTTTT ATCTTTGACCAACCGAACATGACCAAAAACCAAAAGTGCAATC AACCTTACCAAAAAAAAAAAAAAAAAAAGAATAAATAAATAAC TTTTTAAAAAAGGAAGCTTGGTCCACTTGCTTGAAGACCCATG CGGGGGTAAGTCCCTTTCTGCCCCTTGGGCTTATGAAACCCCA ATGCTGCCCTTTCTGCTCCTTTCTCCACACCCCCCTTGGGGCC TCCCTCCACTCCTTCCCAAATCTGTCTCCCCAGAAGACACAG GAAACAATGTATTGTCTGCCCAGCAATCAAAGGCAATGCTCAA ACACCCAAGTGGCCCCCACCCTCAGCCCGCTCCTGCCCGCCCA GCACCCCCAGGCCCTGGGGGACCTGGGGTTCTCAGACTGCCAA AGAAGCCTTGCCATCTGGCGCTCCCATGGCTCTTGCAACATCT CCCCTTCGTTTTTTGAGGGGGTTCATGCCGGGGGAGCCACCAGCC CCTCACTGGGTTCGGAGGAGAGTCAGGAAGGGCCACGACAAAG CAGAAACATCGGATTTGGGGAACGCGTGTCAATCCCTTGTGCC GCAGGGCTGGGCGGGAGAGACTGTTCTGTTCTTGTGTAAGTGT TGTTGCTGAAAGACTACCTCGTTCTTGTCTTGATGTGTCACCG GGGCAACTGCCTGGGGGCGGGGATGGGGGCAGGGTGGAAGCGG CTCCCCATTTTATAACCAAGGTGCTACATCTATGTGATGGGTG GGGTGGGGAGGGAATCACTGGTGCTATAGAAATTGAGATGCCC CCCCAGGCCAGCAAATGTTCCCTTTTTGTTCAAAGTCTATTTTT ATTCCTTGATATTTTTCTTTTTTTTTTTTTTTTTTTTGTGGATG GGGACTTGTGAATTTTTCTAAAGGTGCTATTTAACATGGGAGG AGAGCGTGTGCGGCTCCAGCCCAGCCCGCTGCTCACTTTCCAC CCTCTCTCCACCTGCCTCTGGCTTCTCAGGCCTCTGCTCTCCG ACCTCTCTCCTCTGAAACCCTCCTCCACAGCTGCAGCCCATCC TCCCGGCTCCCTCCTAGTCTGTCCCTGCGTCCTCTGTCCCCGGG TTTCAGAGACAACCTTCCCAAAGCACAAAGCAGTTTTTCCCCCT AGGGGTGGGAGGAAGCAAAGACTCTGTACCTATTTTGTATGT GTATAATAATTTGAGATGTTTTTAATTATTTTGATTGCTGGAA	245

		TAAAGCATGTGGAAATGACCCAAACATAATCCGCAGTGGCCTC CTAATTTCTTCTTTGGAGTTGGGGGAGGGGTAGACATGGGGA AGGGGCTTTGGGGTGATGGGCTTGCCTTCCATTCTGCCCCTT CCCTCCCCACTATTCTCTTAGATCCCTCCATAACCCCACTC CCCTTTCTCTCACCCCTTCTTATACCGCAAACCTTTCTACTTCC TCTTTCATTTTCTATTCTTGCAATTTCTTGCACCTTTTCCAA ATCCTCTTCTCCCCTGCAATACCATAACAGGCAATCCACGTGCA CAACACACACACACACTCTTACATCTGGGGTTGTCCAAACCT CATACCCACTCCCCTTCAAGCCCATCCACTCTCCACCCCTGG ATGCCCTGCACCTGGTGGCGGTGGGATGCTCATGGATACTGGG AGGGTGAGGGGAGTGGAACCCGTGAGGAGGACCTGGGGGCCTC TCCTTGAAGTACATGAAGGGTCATCTGGCCTCTGCTCCCTTC TCACCCACGCTGACCTCCTGCCGAAGGAGCAACGCAACAGGAG AGGGGTCTGCTGAGCCTGGCGAGGGTCTGGGAGGGACCAGGAG GAAGGCGTGCTCCCTGCTCGCTGTCTGGCCCTGGGGGAGTGA GGGAGACAGACACCTGGGAGAGCTGTGGGGAAGGCACTCGCAC CGTGCTCTTGGGAAGGAAGGAGACCTGGCCCTGCTCACCACGG ACTGGGTGCCTCGACCTCCTGAATCCCCAGAACACAACCCCC TGGGCTGGGGTGGTCTGGGGAACCATCGTGCCCCCGCCTCCCG CCTACTCCTTTTAAAGCTT	
3UTR-015	Plod1; procollagen- lysine, 2- oxoglutarate 5- dioxygenase 1	TTGGCCAGGCCTGACCCTCTTGGACCTTTCTTCTTTGCCGACA ACCACTGCCAGCAGCCTCTGGGACCTCGGGGTCCCAGGGAAC CCAGTCCAGCCTCCTGGCTGTTGACTTCCCATTGCTCTTGGAG CCACCAATCAAAGAGATTCAAAGAGATTCTGCAGGCCAGAGG CGGAACACACCTTTATGGCTGGGGCTCTCCGTGGTGTCTGGA CCCAGCCCCCTGGAGACACCATTCACTTTTACTGCTTTGTAGTG ACTCGTGCTCTCCAACCTGTCTTCTGAAAAACCAAGGCCCC TTCCCCACCTCTTCCATGGGGTGAGACTTGAGCAGAACAGGG GCTTCCCCAAGTTGCCAGAAAGACTGTCTGGGTGAGAAGCCA TGGCCAGAGCTTCTCCCAGGCACAGGTGTTGCACCAGGGACTT CTGCTTCAAGTTTGGGGTAAAGACACCTGGATCAGACTCCAA GGGCTGCCCTGAGTCTGGGACTTCTGCCTCCATGGCTGGTCAT GAGAGCAAACCGTAGTCCCCCTGGAGACAGCGACTCCAGAGAAC CTCTTGGGAGACAGAAGAGGCATCTGTGCACAGCTCGATCTTC TACTTGCCTGTGGGGAGGGGAGTGACAGGTCCACACACCACAC TGGGTACCCCTGTCCTGGATGCCTCTGAAGAGAGGGACAGACC GTCAGAACTGGAGAGTTTCTATTAAAGGTCATTTAAACCA	246
3UTR-016	Nucb1; nucleobindi n 1	TCCTCCGGGACCCCAGCCCTCAGGATTCTGATGCTCCAAGGC GACTGATGGGCGCTGGATGAAGTGGCACAGTCAGCTTCCCTGG GGGCTGGTGTCTGTTGGGCTCCTGGGGCGGGGGCACGGCCTG GCATTTACGCATTGCTGCCACCCAGGTCCACCTGTCTCCAC TTTACAGCCTCCAAGTCTGTGGCTCTTCCCTTCTGTCTCCG AGGGGCTTGCCTTCTCTCGTGTCCAGTGAGGTGCTCAGTGATC GGCTTAACTTAGAGAAGCCCCGCCCCCTCCCCTTCTCCGTCTGT CCCAAGAGGGTCTGCTCTGAGCCTGCGTTCTAGGTGGCTCGG CCTCAGCTGCCTGGGTGTGGCCGCCCTAGCATCCTGTATGCC CACAGCTACTGGAATCCCCGCTGCTGCTCCGGGCCAAGCTTCT GGTTGATTAAATGAGGGCATGGGGTGGTCCCTCAAGACCTTCCC CTACCTTTTGTGGAACCAAGTATGCCTCAAAGACAGTGTCCCC TCCACAGCTGGGTGCCAGGGGCAGGGGATCCTCAGTATAGCCG GTGAACCTGATAACAGGAGCCTGGGCCTCCCTGAACCCCTGG CTTCCAGCCATCTCATCGCCAGCCTCCTCCTGGACCTCTTGGC CCCCAGCCCCCTTCCCCACACAGCCCCAGAAGGGTCCCAGAGCT GACCCCACTCCAGGACCTAGGCCAGCCCCCTCAGCCTCATCTG GAGCCCTGAAGACCAGTCCCACCCACCTTTCTGGCCTCATCT GACACTGCTCCGCATCCTGCTGTGTGTCTGTTCATGTTCCG GTTCCATCCAAATACACTTTCTGGAACAAA	247
3UTR-017	α -globin	GCTGGAGCCTCGGTGGCCATGCTTCTTGGCCCTTGGGCCTCCC	248

		CCCAGCCCCCTCCTCCCCCTTCCTGCACCCGTACCCCCGTGGTCT TTGAATAAAGTCTGAGTGGGCGGC	
3UTR-018		TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCCTT GGGCCTCCCCCAGCCCCCTCCTCCCCCTTCCTGCACCCGTACCC CCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	267
3UTR-019		TGATAATAGTCCATAAAGTAGGAAACACTACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCCTTGGGCCTCCCCCAGCCCCCTC CTCCCCCTTCCTGCACCCGTACCCCCCGCATTATTACTCACGGT ACGAGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	773
3UTR-020		TGATAATAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCCTT GGGCCTCCCCCAGCCCCCTCCTCCCCCTTCCTGCACCCGTACCC CCCGCATTATTACTCACGGTACGAGTGGTCTTTGAATAAAGTC TGAGTGGGCGGC	774

[0415] In certain embodiments, the 3' UTR useful for the polynucleotides comprises SEQ ID NO: 267.

[0416] In certain embodiments, the 5'UTR and/or 3'UTR sequence of the disclosure comprises a nucleotide sequence at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identical to a sequence selected from the group consisting of 5'UTR sequences comprising any of SEQ ID NOs: 215-231 and/or 3'UTR sequences comprising any of SEQ ID NOs: 232-248, and any combination thereof. The polynucleotides of the disclosure can comprise combinations of features. For example, the ORF can be flanked by a 5'UTR that comprises a strong Kozak translational initiation signal and/or a 3'UTR comprising an oligo(dT) sequence for templated addition of a poly-A tail. A 5'UTR can comprise a first polynucleotide fragment and a second polynucleotide fragment from the same and/or different UTRs (see, e.g., US2010/0293625, herein incorporated by reference in its entirety).

[0417] It is also within the scope of the present disclosure to have patterned UTRs. As used herein "patterned UTRs" include a repeating or alternating pattern, such as ABABAB or AABBAABBAABB or ABCABCABC or variants thereof repeated once, twice, or more than 3 times. In these patterns, each letter, A, B, or C represent a different UTR nucleic acid sequence.

[0418] Other non-UTR sequences can be used as regions or subregions within the polynucleotides of the disclosure. For example, introns or portions of intron sequences can be incorporated into the polynucleotides of the disclosure. Incorporation of intronic sequences can increase protein production as well as polynucleotide expression levels. In some embodiments, the polynucleotide of the disclosure comprises an internal ribosome entry site (IRES) instead of or in addition to a UTR (see, e.g., Yakubov et al., Biochem. Biophys. Res. Commun. 2010

394(1):189-193, the contents of which are incorporated herein by reference in their entirety). In some embodiments, the polynucleotide of the disclosure comprises 5' and/or 3' sequence associated with the 5' and/or 3' ends of rubella virus (RV) genomic RNA, respectively, or deletion derivatives thereof, including the 5' proximal open reading frame of RV RNA encoding nonstructural proteins (e.g., see Pogue et al., J. Virol. 67(12):7106-7117, the contents of which are incorporated herein by reference in their entirety). Viral capsid sequences can also be used as a translational enhancer, e.g., the 5' portion of a capsid sequence, (e.g., semliki forest virus and sindbis virus capsid RNAs as described in Sjöberg et al., Biotechnology (NY) 1994 12(11):1127-1131, and Frolov and Schlesinger J. Virol. 1996 70(2):1182-1190, the contents of each of which are incorporated herein by reference in their entirety). In some embodiments, the polynucleotide comprises an IRES instead of a 5'UTR sequence. In some embodiments, the polynucleotide comprises an ORF and a viral capsid sequence. In some embodiments, the polynucleotide comprises a synthetic 5'UTR in combination with a non-synthetic 3'UTR.

[0419] In some embodiments, the UTR can also include at least one translation enhancer polynucleotide, translation enhancer element, or translational enhancer elements (collectively, "TEE," which refers to nucleic acid sequences that increase the amount of polypeptide or protein produced from a polynucleotide. As a non-limiting example, the TEE can include those described in US2009/0226470, incorporated herein by reference in its entirety, and others known in the art. As a non-limiting example, the TEE can be located between the transcription promoter and the start codon. In some embodiments, the 5'UTR comprises a TEE.

[0420] In one aspect, a TEE is a conserved element in a UTR that can promote translational activity of a nucleic acid such as, but not limited to, cap-dependent or cap-independent translation. The conservation of these sequences has been shown across 14 species including humans. See, e.g., Panek et al., "An evolutionary conserved pattern of 18S rRNA sequence complementarity to mRNA 5'UTRs and its implications for eukaryotic gene translation regulation," Nucleic Acids Research 2013, doi:10.1093/nar/gkt548, incorporated herein by reference in its entirety.

[0421] In one non-limiting example, the TEE comprises the TEE sequence in the 5'-leader of the Gtx homeodomain protein. See Chappell et al., PNAS 2004 101:9590-9594, incorporated herein by reference in its entirety.

[0422] In another non-limiting example, the TEE comprises a TEE having one or more of the sequences of SEQ ID NOs: 1-35 in US2009/0226470, US2013/0177581, and WO2009/075886;

SEQ ID NOs: 1-5 and 7-645 in WO2012/009644; and SEQ ID NO: 1 WO1999/024595, US6310197, and US6849405; the contents of each of which are incorporated herein by reference in their entirety.

[0423] In some embodiments, the TEE is an internal ribosome entry site (IRES), HCV-IRES, or an IRES element such as, but not limited to, those described in: US7468275, US2007/0048776, US2011/0124100, WO2007/025008, and WO2001/055369; the contents of each of which are incorporated herein by reference in their entirety. The IRES elements can include, but are not limited to, the Gtx sequences (e.g., Gtx9-nt, Gtx8-nt, Gtx7-nt) as described by Chappell et al., PNAS 2004 101:9590-9594, Zhou et al., PNAS 2005 102:6273-6278, US2007/0048776, US2011/0124100, and WO2007/025008; the contents of each of which are incorporated herein by reference in their entirety.

[0424] "Translational enhancer polynucleotide" or "translation enhancer polynucleotide sequence" refer to a polynucleotide that includes one or more of the TEE provided herein and/or known in the art (see, e.g., US6310197, US6849405, US7456273, US7183395, US2009/0226470, US2007/0048776, US2011/0124100, US2009/0093049, US2013/0177581, WO2009/075886, WO2007/025008, WO2012/009644, WO2001/055371, WO1999/024595, EP2610341A1, and EP2610340A1; the contents of each of which are incorporated herein by reference in their entirety), or their variants, homologs, or functional derivatives. In some embodiments, the polynucleotide of the disclosure comprises one or multiple copies of a TEE. The TEE in a translational enhancer polynucleotide can be organized in one or more sequence segments. A sequence segment can harbor one or more of the TEEs provided herein, with each TEE being present in one or more copies. When multiple sequence segments are present in a translational enhancer polynucleotide, they can be homogenous or heterogeneous. Thus, the multiple sequence segments in a translational enhancer polynucleotide can harbor identical or different types of the TEE provided herein, identical or different number of copies of each of the TEE, and/or identical or different organization of the TEE within each sequence segment. In one embodiment, the polynucleotide of the disclosure comprises a translational enhancer polynucleotide sequence.

[0425] In some embodiments, a 5'UTR and/or 3'UTR of a polynucleotide of the disclosure comprises at least one TEE or portion thereof that is disclosed in: WO1999/024595, WO2012/009644, WO2009/075886, WO2007/025008, WO1999/024595, WO2001/055371, EP2610341A1, EP2610340A1, US6310197, US6849405, US7456273, US7183395,

US2009/0226470, US2011/0124100, US2007/0048776, US2009/0093049, or US2013/0177581, the contents of each are incorporated herein by reference in their entirety.

[0426] In some embodiments, a 5'UTR and/or 3'UTR of a polynucleotide of the disclosure comprises a TEE that is at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to a TEE disclosed in: US2009/0226470, US2007/0048776, US2013/0177581, US2011/0124100, WO1999/024595, WO2012/009644, WO2009/075886, WO2007/025008, EP2610341A1, EP2610340A1, US6310197, US6849405, US7456273, US7183395, Chappell et al., PNAS 2004 101:9590-9594, Zhou et al., PNAS 2005 102:6273-6278, and Supplemental Table 1 and in Supplemental Table 2 of Wellensiek et al., "Genome-wide profiling of human cap-independent translation-enhancing elements," Nature Methods 2013, DOI:10.1038/NMETH.2522; the contents of each of which are incorporated herein by reference in their entirety.

[0427] In some embodiments, a 5'UTR and/or 3'UTR of a polynucleotide of the disclosure comprises a TEE which is selected from a 5-30 nucleotide fragment, a 5-25 nucleotide fragment, a 5-20 nucleotide fragment, a 5-15 nucleotide fragment, or a 5-10 nucleotide fragment (including a fragment of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides) of a TEE sequence disclosed in: US2009/0226470, US2007/0048776, US2013/0177581, US2011/0124100, WO1999/024595, WO2012/009644, WO2009/075886, WO2007/025008, EP2610341A1, EP2610340A1, US6310197, US6849405, US7456273, US7183395, Chappell et al., PNAS 2004 101:9590-9594, Zhou et al., PNAS 2005 102:6273-6278, and Supplemental Table 1 and in Supplemental Table 2 of Wellensiek et al., "Genome-wide profiling of human cap-independent translation-enhancing elements," Nature Methods 2013, DOI:10.1038/NMETH.2522.

[0428] In some embodiments, a 5'UTR and/or 3'UTR of a polynucleotide of the disclosure comprises a TEE which is a transcription regulatory element described in any of US7456273, US7183395, US2009/0093049, and WO2001/055371, the contents of each of which are incorporated herein by reference in their entirety. The transcription regulatory elements can be identified by methods known in the art, such as, but not limited to, the methods described in US7456273, US7183395, US2009/0093049, and WO2001/055371.

[0429] In some embodiments, a 5'UTR and/or 3'UTR comprising at least one TEE described herein can be incorporated in a monocistronic sequence such as, but not limited to, a vector system or a nucleic acid vector. As non-limiting examples, the vector systems and nucleic acid vectors can include those described in US7456273, US7183395, US2007/0048776, US2009/0093049, US2011/0124100, WO2007/025008, and WO2001/055371.

[0430] In some embodiments, a 5'UTR and/or 3'UTR of a polynucleotide of the disclosure comprises a TEE or portion thereof described herein. In some embodiments, the TEEs in the 3'UTR can be the same and/or different from the TEE located in the 5'UTR.

[0431] In some embodiments, a 5'UTR and/or 3'UTR of a polynucleotide of the disclosure can include at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18 at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55 or more than 60 TEE sequences. In one embodiment, the 5'UTR of a polynucleotide of the disclosure can include 1-60, 1-55, 1-50, 1-45, 1-40, 1-35, 1-30, 1-25, 1-20, 1-15, 1-10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 TEE sequences. The TEE sequences in the 5'UTR of the polynucleotide of the disclosure can be the same or different TEE sequences. A combination of different TEE sequences in the 5'UTR of the polynucleotide of the disclosure can include combinations in which more than one copy of any of the different TEE sequences are incorporated. The TEE sequences can be in a pattern such as ABABAB or AABBAABBAABB or ABCABCABC or variants thereof repeated one, two, three, or more than three times. In these patterns, each letter, A, B, or C represent a different TEE nucleotide sequence.

[0432] In some embodiments, the TEE can be identified by the methods described in US2007/0048776, US2011/0124100, WO2007/025008, WO2012/009644, the contents of each of which are incorporated herein by reference in their entirety.

[0433] In some embodiments, the 5'UTR and/or 3'UTR comprises a spacer to separate two TEE sequences. As a non-limiting example, the spacer can be a 15 nucleotide spacer and/or other spacers known in the art. As another non-limiting example, the 5'UTR and/or 3'UTR comprises a TEE sequence-spacer module repeated at least once, at least twice, at least 3 times, at least 4 times, at least 5 times, at least 6 times, at least 7 times, at least 8 times, at least 9 times, at least 10 times, or more than 10 times in the 5'UTR and/or 3'UTR, respectively. In some

embodiments, the 5'UTR and/or 3'UTR comprises a TEE sequence-spacer module repeated 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 times.

[0434] In some embodiments, the spacer separating two TEE sequences can include other sequences known in the art that can regulate the translation of the polynucleotide of the disclosure, e.g., miR sequences described herein (e.g., miR binding sites and miR seeds). As a non-limiting example, each spacer used to separate two TEE sequences can include a different miR sequence or component of a miR sequence (e.g., miR seed sequence).

[0435] In some embodiments, a polynucleotide of the disclosure comprises a miR and/or TEE sequence. In some embodiments, the incorporation of a miR sequence and/or a TEE sequence into a polynucleotide of the disclosure can change the shape of the stem loop region, which can increase and/or decrease translation. See e.g., Kedde et al., Nature Cell Biology 2010 12(10):1014-20, herein incorporated by reference in its entirety).

Sensor Sequences and MicroRNA (miRNA) Binding Sites

[0436] Polynucleotides of the disclosure can include regulatory elements, for example, microRNA (miRNA) binding sites, transcription factor binding sites, structured mRNA sequences and/or motifs, artificial binding sites engineered to act as pseudo-receptors for endogenous nucleic acid binding molecules, and combinations thereof. In some embodiments, polynucleotides including such regulatory elements are referred to as including “sensor sequences”. Non-limiting examples of sensor sequences are described in U.S. Publication 2014/0200261, the contents of which are incorporated herein by reference in their entirety.

[0437] In some embodiments, a polynucleotide (e.g., a ribonucleic acid (RNA), e.g., a messenger RNA (mRNA)) of the disclosure comprises an open reading frame (ORF) encoding a polypeptide of interest and further comprises one or more miRNA binding site(s). Inclusion or incorporation of miRNA binding site(s) provides for regulation of polynucleotides of the disclosure, and in turn, of the polypeptides encoded therefrom, based on tissue-specific and/or cell-type specific expression of naturally-occurring miRNAs.

[0438] A miRNA, e.g., a natural-occurring miRNA, is a 19-25 nucleotide long noncoding RNA that binds to a polynucleotide and down-regulates gene expression either by reducing stability or by inhibiting translation of the polynucleotide. A miRNA sequence comprises a “seed” region, i.e., a sequence in the region of positions 2-8 of the mature miRNA. A miRNA

seed can comprise positions 2-8 or 2-7 of the mature miRNA. In some embodiments, a miRNA seed can comprise 7 nucleotides (e.g., nucleotides 2-8 of the mature miRNA), wherein the seed-complementary site in the corresponding miRNA binding site is flanked by an adenosine (A) opposed to miRNA position 1. In some embodiments, a miRNA seed can comprise 6 nucleotides (e.g., nucleotides 2-7 of the mature miRNA), wherein the seed-complementary site in the corresponding miRNA binding site is flanked by an adenosine (A) opposed to miRNA position 1. See, for example, Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP; Mol Cell. 2007 Jul 6;27(1):91-105. miRNA profiling of the target cells or tissues can be conducted to determine the presence or absence of miRNA in the cells or tissues. In some embodiments, a polynucleotide (e.g., a ribonucleic acid (RNA), e.g., a messenger RNA (mRNA)) of the disclosure comprises one or more microRNA binding sites, microRNA target sequences, microRNA complementary sequences, or microRNA seed complementary sequences. Such sequences can correspond to, e.g., have complementarity to, any known microRNA such as those taught in US Publication US2005/0261218 and US Publication US2005/0059005, the contents of each of which are incorporated herein by reference in their entirety.

[0439] As used herein, the term “microRNA (miRNA or miR) binding site” refers to a sequence within a polynucleotide, e.g., within a DNA or within an RNA transcript, including in the 5'UTR and/or 3'UTR, that has sufficient complementarity to all or a region of a miRNA to interact with, associate with or bind to the miRNA. In some embodiments, a polynucleotide of the disclosure comprising an ORF encoding a polypeptide of interest and further comprises one or more miRNA binding site(s). In exemplary embodiments, a 5'UTR and/or 3'UTR of the polynucleotide (e.g., a ribonucleic acid (RNA), e.g., a messenger RNA (mRNA)) comprises the one or more miRNA binding site(s).

[0440] A miRNA binding site having sufficient complementarity to a miRNA refers to a degree of complementarity sufficient to facilitate miRNA-mediated regulation of a polynucleotide, e.g., miRNA-mediated translational repression or degradation of the polynucleotide. In exemplary aspects of the disclosure, a miRNA binding site having sufficient complementarity to the miRNA refers to a degree of complementarity sufficient to facilitate miRNA-mediated degradation of the polynucleotide, e.g., miRNA-guided RNA-induced silencing complex (RISC)-mediated cleavage of mRNA. The miRNA binding site can have complementarity to, for example, a 19-25 nucleotide miRNA sequence, to a 19-23 nucleotide miRNA sequence, or to a 22 nucleotide miRNA sequence. A miRNA binding site can be

complementary to only a portion of a miRNA, e.g., to a portion less than 1, 2, 3, or 4 nucleotides of the full length of a naturally-occurring miRNA sequence. Full or complete complementarity (e.g., full complementarity or complete complementarity over all or a significant portion of the length of a naturally-occurring miRNA) is preferred when the desired regulation is mRNA degradation.

[0441] In some embodiments, a miRNA binding site includes a sequence that has complementarity (e.g., partial or complete complementarity) with an miRNA seed sequence. In some embodiments, the miRNA binding site includes a sequence that has complete complementarity with a miRNA seed sequence. In some embodiments, a miRNA binding site includes a sequence that has complementarity (e.g., partial or complete complementarity) with an miRNA sequence. In some embodiments, the miRNA binding site includes a sequence that has complete complementarity with a miRNA sequence. In some embodiments, a miRNA binding site has complete complementarity with a miRNA sequence but for 1, 2, or 3 nucleotide substitutions, terminal additions, and/or truncations.

[0442] In some embodiments, the miRNA binding site is the same length as the corresponding miRNA. In other embodiments, the miRNA binding site is one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve nucleotide(s) shorter than the corresponding miRNA at the 5' terminus, the 3' terminus, or both. In still other embodiments, the microRNA binding site is two nucleotides shorter than the corresponding microRNA at the 5' terminus, the 3' terminus, or both. The miRNA binding sites that are shorter than the corresponding miRNAs are still capable of degrading the mRNA incorporating one or more of the miRNA binding sites or preventing the mRNA from translation.

[0443] In some embodiments, the miRNA binding site binds the corresponding mature miRNA that is part of an active RISC containing Dicer. In another embodiment, binding of the miRNA binding site to the corresponding miRNA in RISC degrades the mRNA containing the miRNA binding site or prevents the mRNA from being translated. In some embodiments, the miRNA binding site has sufficient complementarity to miRNA so that a RISC complex comprising the miRNA cleaves the polynucleotide comprising the miRNA binding site. In other embodiments, the miRNA binding site has imperfect complementarity so that a RISC complex comprising the miRNA induces instability in the polynucleotide comprising the miRNA binding site. In another embodiment, the miRNA binding site has imperfect complementarity so that a

RISC complex comprising the miRNA represses transcription of the polynucleotide comprising the miRNA binding site.

[0444] In some embodiments, the miRNA binding site has one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve mismatch(es) from the corresponding miRNA.

[0445] In some embodiments, the miRNA binding site has at least about ten, at least about eleven, at least about twelve, at least about thirteen, at least about fourteen, at least about fifteen, at least about sixteen, at least about seventeen, at least about eighteen, at least about nineteen, at least about twenty, or at least about twenty-one contiguous nucleotides complementary to at least about ten, at least about eleven, at least about twelve, at least about thirteen, at least about fourteen, at least about fifteen, at least about sixteen, at least about seventeen, at least about eighteen, at least about nineteen, at least about twenty, or at least about twenty-one, respectively, contiguous nucleotides of the corresponding miRNA, .

[0446] By engineering one or more miRNA binding sites into a polynucleotide of the disclosure, the polynucleotide can be targeted for degradation or reduced translation, provided the miRNA in question is available. This can reduce off-target effects upon delivery of the polynucleotide. For example, if a polynucleotide of the disclosure is not intended to be delivered to a tissue or cell but ends up in said tissue or cell, then a miRNA abundant in the tissue or cell can inhibit the expression of the gene of interest if one or multiple binding sites of the miRNA are engineered into the 5'UTR and/or 3'UTR of the polynucleotide.

[0447] Conversely, miRNA binding sites can be removed from polynucleotide sequences in which they naturally occur in order to increase protein expression in specific tissues. For example, a binding site for a specific miRNA can be removed from a polynucleotide to improve protein expression in tissues or cells containing the miRNA.

[0448] In one embodiment, a polynucleotide of the disclosure can include at least one miRNA-binding site in the 5'UTR and/or 3'UTR in order to regulate cytotoxic or cytoprotective mRNA therapeutics to specific cells such as, but not limited to, normal and/or cancerous cells. In another embodiment, a polynucleotide of the disclosure can include two, three, four, five, six, seven, eight, nine, ten, or more miRNA-binding sites in the 5'-UTR and/or 3'-UTR in order to regulate cytotoxic or cytoprotective mRNA therapeutics to specific cells such as, but not limited to, normal and/or cancerous cells.

[0449] Regulation of expression in multiple tissues can be accomplished through introduction or removal of one or more miRNA binding sites, e.g., one or more distinct miRNA

binding sites. The decision whether to remove or insert a miRNA binding site can be made based on miRNA expression patterns and/or their profilings in tissues and/or cells in development and/or disease. Identification of miRNAs, miRNA binding sites, and their expression patterns and role in biology have been reported (e.g., Bonauer et al., *Curr Drug Targets* 2010 11:943-949; Anand and Cheresch *Curr Opin Hematol* 2011 18:171-176; Contreras and Rao *Leukemia* 2012 26:404-413 (2011 Dec 20. doi: 10.1038/leu.2011.356); Bartel *Cell* 2009 136:215-233; Landgraf et al, *Cell*, 2007 129:1401-1414; Gentner and Naldini, *Tissue Antigens*. 2012 80:393-403 and all references therein; each of which is incorporated herein by reference in its entirety).

[0450] miRNAs and miRNA binding sites can correspond to any known sequence, including non-limiting examples described in U.S. Publication Nos. 2014/0200261, 2005/0261218, and 2005/0059005, each of which are incorporated herein by reference in their entirety.

[0451] Examples of tissues where miRNA are known to regulate mRNA, and thereby protein expression, include, but are not limited to, liver (miR-122), muscle (miR-133, miR-206, miR-208), endothelial cells (miR-17-92, miR-126), myeloid cells (miR-142-3p, miR-142-5p, miR-16, miR-21, miR-223, miR-24, miR-27), adipose tissue (let-7, miR-30c), heart (miR-1d, miR-149), kidney (miR-192, miR-194, miR-204), and lung epithelial cells (let-7, miR-133, miR-126).

[0452] Specifically, miRNAs are known to be differentially expressed in immune cells (also called hematopoietic cells), such as antigen presenting cells (APCs) (e.g., dendritic cells and macrophages), macrophages, monocytes, B lymphocytes, T lymphocytes, granulocytes, natural killer cells, etc. Immune cell specific miRNAs are involved in immunogenicity, autoimmunity, the immune-response to infection, inflammation, as well as unwanted immune response after gene therapy and tissue/organ transplantation. Immune cells specific miRNAs also regulate many aspects of development, proliferation, differentiation and apoptosis of hematopoietic cells (immune cells). For example, miR-142 and miR-146 are exclusively expressed in immune cells, particularly abundant in myeloid dendritic cells. It has been demonstrated that the immune response to a polynucleotide can be shut-off by adding miR-142 binding sites to the 3'-UTR of the polynucleotide, enabling more stable gene transfer in tissues and cells. miR-142 efficiently degrades exogenous polynucleotides in antigen presenting cells and suppresses cytotoxic elimination of transduced cells (e.g., Annoni A et al., *blood*, 2009, 114, 5152-5161; Brown BD, et al., *Nat med.* 2006, 12(5), 585-591; Brown BD, et al., *blood*, 2007, 110(13): 4144-4152, each of which is incorporated herein by reference in its entirety).

[0453] An antigen-mediated immune response can refer to an immune response triggered by foreign antigens, which, when entering an organism, are processed by the antigen presenting cells and displayed on the surface of the antigen presenting cells. T cells can recognize the presented antigen and induce a cytotoxic elimination of cells that express the antigen.

[0454] Introducing a miR-142 binding site into the 5'UTR and/or 3'UTR of a polynucleotide of the disclosure can selectively repress gene expression in antigen presenting cells through miR-142 mediated degradation, limiting antigen presentation in antigen presenting cells (e.g., dendritic cells) and thereby preventing antigen-mediated immune response after the delivery of the polynucleotide. The polynucleotide is then stably expressed in target tissues or cells without triggering cytotoxic elimination.

[0455] In one embodiment, binding sites for miRNAs that are known to be expressed in immune cells, in particular, antigen presenting cells, can be engineered into a polynucleotide of the disclosure to suppress the expression of the polynucleotide in antigen presenting cells through miRNA mediated RNA degradation, subduing the antigen-mediated immune response. Expression of the polynucleotide is maintained in non-immune cells where the immune cell specific miRNAs are not expressed. For example, in some embodiments, to prevent an immunogenic reaction against a liver specific protein, any miR-122 binding site can be removed and a miR-142 (and/or miR-146) binding site can be engineered into the 5'UTR and/or 3'UTR of a polynucleotide of the disclosure.

[0456] To further drive the selective degradation and suppression in APCs and macrophage, a polynucleotide of the disclosure can include a further negative regulatory element in the 5'UTR and/or 3'UTR, either alone or in combination with miR-142 and/or miR-146 binding sites. As a non-limiting example, the further negative regulatory element is a Constitutive Decay Element (CDE).

[0457] Immune cell specific miRNAs include, but are not limited to, hsa-let-7a-2-3p, hsa-let-7a-3p, hsa-7a-5p, hsa-let-7c, hsa-let-7e-3p, hsa-let-7e-5p, hsa-let-7g-3p, hsa-let-7g-5p, hsa-let-7i-3p, hsa-let-7i-5p, miR-10a-3p, miR-10a-5p, miR-1184, hsa-let-7f-1--3p, hsa-let-7f-2--5p, hsa-let-7f-5p, miR-125b-1-3p, miR-125b-2-3p, miR-125b-5p, miR-1279, miR-130a-3p, miR-130a-5p, miR-132-3p, miR-132-5p, miR-142-3p, miR-142-5p, miR-143-3p, miR-143-5p, miR-146a-3p, miR-146a-5p, miR-146b-3p, miR-146b-5p, miR-147a, miR-147b, miR-148a-5p, miR-148a-3p, miR-150-3p, miR-150-5p, miR-151b, miR-155-3p, miR-155-5p, miR-15a-3p, miR-15a-5p, miR-15b-5p, miR-15b-3p, miR-16-1-3p, miR-16-2-3p, miR-16-5p, miR-17-5p, miR-181a-3p,

miR-181a-5p, miR-181a-2-3p, miR-182-3p, miR-182-5p, miR-197-3p, miR-197-5p, miR-21-5p, miR-21-3p, miR-214-3p, miR-214-5p, miR-223-3p, miR-223-5p, miR-221-3p, miR-221-5p, miR-23b-3p, miR-23b-5p, miR-24-1-5p, miR-24-2-5p, miR-24-3p, miR-26a-1-3p, miR-26a-2-3p, miR-26a-5p, miR-26b-3p, miR-26b-5p, miR-27a-3p, miR-27a-5p, miR-27b-3p, miR-27b-5p, miR-28-3p, miR-28-5p, miR-2909, miR-29a-3p, miR-29a-5p, miR-29b-1-5p, miR-29b-2-5p, miR-29c-3p, miR-29c-5p, miR-30e-3p, miR-30e-5p, miR-331-5p, miR-339-3p, miR-339-5p, miR-345-3p, miR-345-5p, miR-346, miR-34a-3p, miR-34a-5p, , miR-363-3p, miR-363-5p, miR-372, miR-377-3p, miR-377-5p, miR-493-3p, miR-493-5p, miR-542, miR-548b-5p, miR-548c-5p, miR-548i, miR-548j, miR-548n, miR-574-3p, miR-598, miR-718, miR-935, miR-99a-3p, miR-99a-5p, miR-99b-3p, and miR-99b-5p. Furthermore, novel miRNAs can be identified in immune cell through micro-array hybridization and microtome analysis (e.g., Jima DD et al, Blood, 2010, 116:e118-e127; Vaz C et al., BMC Genomics, 2010, 11,288, the content of each of which is incorporated herein by reference in its entirety.)

[0458] miRNAs that are known to be expressed in the liver include, but are not limited to, miR-107, miR-122-3p, miR-122-5p, miR-1228-3p, miR-1228-5p, miR-1249, miR-129-5p, miR-1303, miR-151a-3p, miR-151a-5p, miR-152, miR-194-3p, miR-194-5p, miR-199a-3p, miR-199a-5p, miR-199b-3p, miR-199b-5p, miR-296-5p, miR-557, miR-581, miR-939-3p, and miR-939-5p. MiRNA binding sites from any liver specific miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the liver. Liver specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g., APC) miRNA binding sites in a polynucleotide of the disclosure.

[0459] miRNAs that are known to be expressed in the lung include, but are not limited to, let-7a-2-3p, let-7a-3p, let-7a-5p, miR-126-3p, miR-126-5p, miR-127-3p, miR-127-5p, miR-130a-3p, miR-130a-5p, miR-130b-3p, miR-130b-5p, miR-133a, miR-133b, miR-134, miR-18a-3p, miR-18a-5p, miR-18b-3p, miR-18b-5p, miR-24-1-5p, miR-24-2-5p, miR-24-3p, miR-296-3p, miR-296-5p, miR-32-3p, miR-337-3p, miR-337-5p, miR-381-3p, and miR-381-5p. miRNA binding sites from any lung specific miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the lung. Lung specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g., APC) miRNA binding sites in a polynucleotide of the disclosure.

[0460] miRNAs that are known to be expressed in the heart include, but are not limited to, miR-1, miR-133a, miR-133b, miR-149-3p, miR-149-5p, miR-186-3p, miR-186-5p, miR-208a,

miR-208b, miR-210, miR-296-3p, miR-320, miR-451a, miR-451b, miR-499a-3p, miR-499a-5p, miR-499b-3p, miR-499b-5p, miR-744-3p, miR-744-5p, miR-92b-3p, and miR-92b-5p. mMIRNA binding sites from any heart specific microRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the heart. Heart specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g., APC) miRNA binding sites in a polynucleotide of the disclosure.

[0461] miRNAs that are known to be expressed in the nervous system include, but are not limited to, miR-124-5p, miR-125a-3p, miR-125a-5p, miR-125b-1-3p, miR-125b-2-3p, miR-125b-5p, miR-1271-3p, miR-1271-5p, miR-128, miR-132-5p, miR-135a-3p, miR-135a-5p, miR-135b-3p, miR-135b-5p, miR-137, miR-139-5p, miR-139-3p, miR-149-3p, miR-149-5p, miR-153, miR-181c-3p, miR-181c-5p, miR-183-3p, miR-183-5p, miR-190a, miR-190b, miR-212-3p, miR-212-5p, miR-219-1-3p, miR-219-2-3p, miR-23a-3p, miR-23a-5p, miR-30a-5p, miR-30b-3p, miR-30b-5p, miR-30c-1-3p, miR-30c-2-3p, miR-30c-5p, miR-30d-3p, miR-30d-5p, miR-329, miR-342-3p, miR-3665, miR-3666, miR-380-3p, miR-380-5p, miR-383, miR-410, miR-425-3p, miR-425-5p, miR-454-3p, miR-454-5p, miR-483, miR-510, miR-516a-3p, miR-548b-5p, miR-548c-5p, miR-571, miR-7-1-3p, miR-7-2-3p, miR-7-5p, miR-802, miR-922, miR-9-3p, and miR-9-5p. miRNAs enriched in the nervous system further include those specifically expressed in neurons, including, but not limited to, miR-132-3p, miR-132-5p, miR-148b-3p, miR-148b-5p, miR-151a-3p, miR-151a-5p, miR-212-3p, miR-212-5p, miR-320b, miR-320e, miR-323a-3p, miR-323a-5p, miR-324-5p, miR-325, miR-326, miR-328, miR-922 and those specifically expressed in glial cells, including, but not limited to, miR-1250, miR-219-1-3p, miR-219-2-3p, miR-219-5p, miR-23a-3p, miR-23a-5p, miR-3065-3p, miR-3065-5p, miR-30e-3p, miR-30e-5p, miR-32-5p, miR-338-5p, and miR-657. miRNA binding sites from any CNS specific miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the nervous system. Nervous system specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g., APC) miRNA binding sites in a polynucleotide of the disclosure.

[0462] miRNAs that are known to be expressed in the pancreas include, but are not limited to, miR-105-3p, miR-105-5p, miR-184, miR-195-3p, miR-195-5p, miR-196a-3p, miR-196a-5p, miR-214-3p, miR-214-5p, miR-216a-3p, miR-216a-5p, miR-30a-3p, miR-33a-3p, miR-33a-5p, miR-375, miR-7-1-3p, miR-7-2-3p, miR-493-3p, miR-493-5p, and miR-944. MiRNA binding sites from any pancreas specific miRNA can be introduced to or removed from a polynucleotide

of the disclosure to regulate expression of the polynucleotide in the pancreas. Pancreas specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g. APC) miRNA binding sites in a polynucleotide of the disclosure.

[0463] miRNAs that are known to be expressed in the kidney include, but are not limited to, miR-122-3p, miR-145-5p, miR-17-5p, miR-192-3p, miR-192-5p, miR-194-3p, miR-194-5p, miR-20a-3p, miR-20a-5p, miR-204-3p, miR-204-5p, miR-210, miR-216a-3p, miR-216a-5p, miR-296-3p, miR-30a-3p, miR-30a-5p, miR-30b-3p, miR-30b-5p, miR-30c-1-3p, miR-30c-2-3p, miR-30c-5p, miR-324-3p, miR-335-3p, miR-335-5p, miR-363-3p, miR-363-5p, and miR-562. miRNA binding sites from any kidney specific miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the kidney. Kidney specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g., APC) miRNA binding sites in a polynucleotide of the disclosure.

[0464] miRNAs that are known to be expressed in the muscle include, but are not limited to, let-7g-3p, let-7g-5p, miR-1, miR-1286, miR-133a, miR-133b, miR-140-3p, miR-143-3p, miR-143-5p, miR-145-3p, miR-145-5p, miR-188-3p, miR-188-5p, miR-206, miR-208a, miR-208b, miR-25-3p, and miR-25-5p. MiRNA binding sites from any muscle specific miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the muscle. Muscle specific miRNA binding sites can be engineered alone or further in combination with immune cell (e.g., APC) miRNA binding sites in a polynucleotide of the disclosure.

[0465] miRNAs are also differentially expressed in different types of cells, such as, but not limited to, endothelial cells, epithelial cells, and adipocytes.

[0466] miRNAs that are known to be expressed in endothelial cells include, but are not limited to, let-7b-3p, let-7b-5p, miR-100-3p, miR-100-5p, miR-101-3p, miR-101-5p, miR-126-3p, miR-126-5p, miR-1236-3p, miR-1236-5p, miR-130a-3p, miR-130a-5p, miR-17-5p, miR-17-3p, miR-18a-3p, miR-18a-5p, miR-19a-3p, miR-19a-5p, miR-19b-1-5p, miR-19b-2-5p, miR-19b-3p, miR-20a-3p, miR-20a-5p, miR-217, miR-210, miR-21-3p, miR-21-5p, miR-221-3p, miR-221-5p, miR-222-3p, miR-222-5p, miR-23a-3p, miR-23a-5p, miR-296-5p, miR-361-3p, miR-361-5p, miR-421, miR-424-3p, miR-424-5p, miR-513a-5p, miR-92a-1-5p, miR-92a-2-5p, miR-92a-3p, miR-92b-3p, and miR-92b-5p. Many novel miRNAs are discovered in endothelial cells from deep-sequencing analysis (e.g., Voellenkle C et al., RNA, 2012, 18, 472-484, herein incorporated by reference in its entirety). miRNA binding sites from any endothelial cell specific

miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the endothelial cells.

[0467] miRNAs that are known to be expressed in epithelial cells include, but are not limited to, let-7b-3p, let-7b-5p, miR-1246, miR-200a-3p, miR-200a-5p, miR-200b-3p, miR-200b-5p, miR-200c-3p, miR-200c-5p, miR-338-3p, miR-429, miR-451a, miR-451b, miR-494, miR-802 and miR-34a, miR-34b-5p, miR-34c-5p, miR-449a, miR-449b-3p, miR-449b-5p specific in respiratory ciliated epithelial cells, let-7 family, miR-133a, miR-133b, miR-126 specific in lung epithelial cells, miR-382-3p, miR-382-5p specific in renal epithelial cells, and miR-762 specific in corneal epithelial cells. miRNA binding sites from any epithelial cell specific miRNA can be introduced to or removed from a polynucleotide of the disclosure to regulate expression of the polynucleotide in the epithelial cells.

[0468] In addition, a large group of miRNAs are enriched in embryonic stem cells, controlling stem cell self-renewal as well as the development and/or differentiation of various cell lineages, such as neural cells, cardiac, hematopoietic cells, skin cells, osteogenic cells and muscle cells (e.g., Kuppusamy KT et al., Curr. Mol Med, 2013, 13(5), 757-764; Vidigal JA and Ventura A, Semin Cancer Biol. 2012, 22(5-6), 428-436; Goff LA et al., PLoS One, 2009, 4:e7192; Morin RD et al., Genome Res, 2008, 18, 610-621; Yoo JK et al., Stem Cells Dev. 2012, 21(11), 2049-2057, each of which is herein incorporated by reference in its entirety). MiRNAs abundant in embryonic stem cells include, but are not limited to, let-7a-2-3p, let-a-3p, let-7a-5p, let7d-3p, let-7d-5p, miR-103a-2-3p, miR-103a-5p, miR-106b-3p, miR-106b-5p, miR-1246, miR-1275, miR-138-1-3p, miR-138-2-3p, miR-138-5p, miR-154-3p, miR-154-5p, miR-200c-3p, miR-200c-5p, miR-290, miR-301a-3p, miR-301a-5p, miR-302a-3p, miR-302a-5p, miR-302b-3p, miR-302b-5p, miR-302c-3p, miR-302c-5p, miR-302d-3p, miR-302d-5p, miR-302e, miR-367-3p, miR-367-5p, miR-369-3p, miR-369-5p, miR-370, miR-371, miR-373, miR-380-5p, miR-423-3p, miR-423-5p, miR-486-5p, miR-520c-3p, miR-548e, miR-548f, miR-548g-3p, miR-548g-5p, miR-548i, miR-548k, miR-548l, miR-548m, miR-548n, miR-548o-3p, miR-548o-5p, miR-548p, miR-664a-3p, miR-664a-5p, miR-664b-3p, miR-664b-5p, miR-766-3p, miR-766-5p, miR-885-3p, miR-885-5p, miR-93-3p, miR-93-5p, miR-941, miR-96-3p, miR-96-5p, miR-99b-3p and miR-99b-5p. Many predicted novel miRNAs are discovered by deep sequencing in human embryonic stem cells (e.g., Morin RD et al., Genome Res, 2008, 18, 610-621; Goff LA et al., PLoS One, 2009, 4:e7192; Bar M et al., Stem cells, 2008, 26, 2496-2505, the content of each of which is incorporated herein by reference in its entirety).

[0469] In one embodiment, the binding sites of embryonic stem cell specific miRNAs can be included in or removed from the 3'UTR of a polynucleotide of the disclosure to modulate the development and/or differentiation of embryonic stem cells, to inhibit the senescence of stem cells in a degenerative condition (e.g. degenerative diseases), or to stimulate the senescence and apoptosis of stem cells in a disease condition (e.g. cancer stem cells).

[0470] Many miRNA expression studies are conducted to profile the differential expression of miRNAs in various cancer cells/tissues and other diseases. Some miRNAs are abnormally over-expressed in certain cancer cells and others are under-expressed. For example, miRNAs are differentially expressed in cancer cells (WO2008/154098, US2013/0059015, US2013/0042333, WO2011/157294); cancer stem cells (US2012/0053224); pancreatic cancers and diseases (US2009/0131348, US2011/0171646, US2010/0286232, US8389210); asthma and inflammation (US8415096); prostate cancer (US2013/0053264); hepatocellular carcinoma (WO2012/151212, US2012/0329672, WO2008/054828, US8252538); lung cancer cells (WO2011/076143, WO2013/033640, WO2009/070653, US2010/0323357); cutaneous T cell lymphoma (WO2013/011378); colorectal cancer cells (WO2011/0281756, WO2011/076142); cancer positive lymph nodes (WO2009/100430, US2009/0263803); nasopharyngeal carcinoma (EP2112235); chronic obstructive pulmonary disease (US2012/0264626, US2013/0053263); thyroid cancer (WO2013/066678); ovarian cancer cells (US2012/0309645, WO2011/095623); breast cancer cells (WO2008/154098, WO2007/081740, US2012/0214699), leukemia and lymphoma (WO2008/073915, US2009/0092974, US2012/0316081, US2012/0283310, WO2010/018563, the content of each of which is incorporated herein by reference in its entirety.)

[0471] As a non-limiting example, miRNA binding sites for miRNAs that are over-expressed in certain cancer and/or tumor cells can be removed from the 3'UTR of a polynucleotide of the disclosure, restoring the expression suppressed by the over-expressed miRNAs in cancer cells, thus ameliorating the corresponsive biological function, for instance, transcription stimulation and/or repression, cell cycle arrest, apoptosis and cell death. Normal cells and tissues, wherein miRNAs expression is not up-regulated, will remain unaffected.

[0472] miRNA can also regulate complex biological processes such as angiogenesis (e.g., miR-132) (Anand and Cheresi Curr Opin Hematol 2011 18:171-176). In the polynucleotides of the disclosure, miRNA binding sites that are involved in such processes can be removed or introduced, in order to tailor the expression of the polynucleotides to biologically relevant cell

types or relevant biological processes. In this context, the polynucleotides of the disclosure are defined as auxotrophic polynucleotides.

[0473] In some embodiments, a polynucleotide of the disclosure comprises a miRNA binding site, wherein the miRNA binding site comprises one or more nucleotide sequences selected from **Table 7**, including one or more copies of any one or more of the miRNA binding site sequences. In some embodiments, a polynucleotide of the disclosure further comprises at least one, two, three, four, five, six, seven, eight, nine, ten, or more of the same or different miRNA binding sites selected from **Table 7**, including any combination thereof. In some embodiments, the miRNA binding site binds to miR-142 or is complementary to miR-142. In some embodiments, the miR-142 comprises SEQ ID NO: 720. In some embodiments, the miRNA binding site binds to miR-142-3p or miR-142-5p. In some embodiments, the miR-142-3p binding site comprises SEQ ID NO: 721. In some embodiments, the miR-142-5p binding site comprises SEQ ID NO: 723. In some embodiments, the miRNA binding site comprises a nucleotide sequence at least 80%, at least 85%, at least 90%, at least 95%, or 100% identical to SEQ ID NO: 722 or SEQ ID NO: 724.

Table 7. miR-142 and miR-142 binding sites

SEQ ID NO.	Description	Sequence
720	miR-142	GACAGUGCAGUCACCCAUAAAGUAGAAAGCA CUACUAAACAGCACUGGAGGGUGUAGUGUUUC CUACUUUAUGGAUGAGUGUACUGUG
721	miR-142-3p	UGUAGUGUUUCCUACUUUAUGGA
722	miR-142-3p binding site	UCCAUAAAGUAGGAAACACUACA
723	miR-142-5p	CAUAAAGUAGAAAGCACUACU
724	miR-142-5p binding site	AGUAGUGCUUUCUACUUUAUG

[0474] In some embodiments, a miRNA binding site is inserted in the polynucleotide of the disclosure in any position of the polynucleotide (e.g., the 5'UTR and/or 3'UTR). In some embodiments, the 5'UTR comprises a miRNA binding site. In some embodiments, the 3'UTR comprises a miRNA binding site. In some embodiments, the 5'UTR and the 3'UTR comprise a miRNA binding site. The insertion site in the polynucleotide can be anywhere in the polynucleotide as long as the insertion of the miRNA binding site in the polynucleotide does not

interfere with the translation of a functional polypeptide in the absence of the corresponding miRNA; and in the presence of the miRNA, the insertion of the miRNA binding site in the polynucleotide and the binding of the miRNA binding site to the corresponding miRNA are capable of degrading the polynucleotide or preventing the translation of the polynucleotide.

[0475] In some embodiments, a miRNA binding site is inserted in at least about 30 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the disclosure comprising the ORF. In some embodiments, a miRNA binding site is inserted in at least about 10 nucleotides, at least about 15 nucleotides, at least about 20 nucleotides, at least about 25 nucleotides, at least about 30 nucleotides, at least about 35 nucleotides, at least about 40 nucleotides, at least about 45 nucleotides, at least about 50 nucleotides, at least about 55 nucleotides, at least about 60 nucleotides, at least about 65 nucleotides, at least about 70 nucleotides, at least about 75 nucleotides, at least about 80 nucleotides, at least about 85 nucleotides, at least about 90 nucleotides, at least about 95 nucleotides, or at least about 100 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the disclosure. In some embodiments, a miRNA binding site is inserted in about 10 nucleotides to about 100 nucleotides, about 20 nucleotides to about 90 nucleotides, about 30 nucleotides to about 80 nucleotides, about 40 nucleotides to about 70 nucleotides, about 50 nucleotides to about 60 nucleotides, about 45 nucleotides to about 65 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the disclosure.

[0476] miRNA gene regulation can be influenced by the sequence surrounding the miRNA such as, but not limited to, the species of the surrounding sequence, the type of sequence (e.g., heterologous, homologous, exogenous, endogenous, or artificial), regulatory elements in the surrounding sequence and/or structural elements in the surrounding sequence. The miRNA can be influenced by the 5'UTR and/or 3'UTR. As a non-limiting example, a non-human 3'UTR can increase the regulatory effect of the miRNA sequence on the expression of a polypeptide of interest compared to a human 3'UTR of the same sequence type.

[0477] In one embodiment, other regulatory elements and/or structural elements of the 5'UTR can influence miRNA mediated gene regulation. One example of a regulatory element and/or structural element is a structured IRES (Internal Ribosome Entry Site) in the 5'UTR, which is necessary for the binding of translational elongation factors to initiate protein translation. EIF4A2 binding to this secondarily structured element in the 5'-UTR is necessary for miRNA mediated gene expression (Meijer HA et al., Science, 2013, 340, 82-85, herein

incorporated by reference in its entirety). The polynucleotides of the disclosure can further include this structured 5'UTR in order to enhance microRNA mediated gene regulation.

[0478] At least one miRNA binding site can be engineered into the 3'UTR of a polynucleotide of the disclosure. In this context, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or more miRNA binding sites can be engineered into a 3'UTR of a polynucleotide of the disclosure. For example, 1 to 10, 1 to 9, 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 2, or 1 miRNA binding sites can be engineered into the 3'UTR of a polynucleotide of the disclosure. In one embodiment, miRNA binding sites incorporated into a polynucleotide of the disclosure can be the same or can be different miRNA sites. A combination of different miRNA binding sites incorporated into a polynucleotide of the disclosure can include combinations in which more than one copy of any of the different miRNA sites are incorporated. In another embodiment, miRNA binding sites incorporated into a polynucleotide of the disclosure can target the same or different tissues in the body. As a non-limiting example, through the introduction of tissue-, cell-type-, or disease-specific miRNA binding sites in the 3'-UTR of a polynucleotide of the disclosure, the degree of expression in specific cell types (e.g., hepatocytes, myeloid cells, endothelial cells, cancer cells, etc.) can be reduced.

[0479] In one embodiment, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR, about halfway between the 5' terminus and 3' terminus of the 3'UTR and/or near the 3' terminus of the 3'UTR in a polynucleotide of the disclosure. As a non-limiting example, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR and about halfway between the 5' terminus and 3' terminus of the 3'UTR. As another non-limiting example, a miRNA binding site can be engineered near the 3' terminus of the 3'UTR and about halfway between the 5' terminus and 3' terminus of the 3'UTR. As yet another non-limiting example, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR and near the 3' terminus of the 3'UTR.

[0480] In another embodiment, a 3'UTR can comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 miRNA binding sites. The miRNA binding sites can be complementary to a miRNA, miRNA seed sequence, and/or miRNA sequences flanking the seed sequence.

[0481] In one embodiment, a polynucleotide of the disclosure can be engineered to include more than one miRNA site expressed in different tissues or different cell types of a subject. As a non-limiting example, a polynucleotide of the disclosure can be engineered to include miR-192

and miR-122 to regulate expression of the polynucleotide in the liver and kidneys of a subject. In another embodiment, a polynucleotide of the disclosure can be engineered to include more than one miRNA site for the same tissue.

[0482] In some embodiments, the therapeutic window and or differential expression associated with the polypeptide encoded by a polynucleotide of the disclosure can be altered with a miRNA binding site. For example, a polynucleotide encoding a polypeptide that provides a death signal can be designed to be more highly expressed in cancer cells by virtue of the miRNA signature of those cells. Where a cancer cell expresses a lower level of a particular miRNA, the polynucleotide encoding the binding site for that miRNA (or miRNAs) would be more highly expressed. Hence, the polypeptide that provides a death signal triggers or induces cell death in the cancer cell. Neighboring noncancer cells, harboring a higher expression of the same miRNA would be less affected by the encoded death signal as the polynucleotide would be expressed at a lower level due to the effects of the miRNA binding to the binding site or “sensor” encoded in the 3'UTR. Conversely, cell survival or cytoprotective signals can be delivered to tissues containing cancer and non-cancerous cells where a miRNA has a higher expression in the cancer cells—the result being a lower survival signal to the cancer cell and a larger survival signal to the normal cell. Multiple polynucleotides can be designed and administered having different signals based on the use of miRNA binding sites as described herein.

[0483] In some embodiments, the expression of a polynucleotide of the disclosure can be controlled by incorporating at least one sensor sequence in the polynucleotide and formulating the polynucleotide for administration. As a non-limiting example, a polynucleotide of the disclosure can be targeted to a tissue or cell by incorporating a miRNA binding site and formulating the polynucleotide in a lipid nanoparticle comprising a cationic lipid, including any of the lipids described herein.

[0484] A polynucleotide of the disclosure can be engineered for more targeted expression in specific tissues, cell types, or biological conditions based on the expression patterns of miRNAs in the different tissues, cell types, or biological conditions. Through introduction of tissue-specific miRNA binding sites, a polynucleotide of the disclosure can be designed for optimal protein expression in a tissue or cell, or in the context of a biological condition.

[0485] In some embodiments, a polynucleotide of the disclosure can be designed to incorporate miRNA binding sites that either have 100% identity to known miRNA seed sequences or have less than 100% identity to miRNA seed sequences. In some embodiments, a

polynucleotide of the disclosure can be designed to incorporate miRNA binding sites that have at least: 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to known miRNA seed sequences. The miRNA seed sequence can be partially mutated to decrease miRNA binding affinity and as such result in reduced downmodulation of the polynucleotide. In essence, the degree of match or mis-match between the miRNA binding site and the miRNA seed can act as a rheostat to more finely tune the ability of the miRNA to modulate protein expression. In addition, mutation in the non-seed region of a miRNA binding site can also impact the ability of a miRNA to modulate protein expression.

[0486] In one embodiment, a miRNA sequence can be incorporated into the loop of a stem loop.

[0487] In another embodiment, a miRNA seed sequence can be incorporated in the loop of a stem loop and a miRNA binding site can be incorporated into the 5' or 3' stem of the stem loop.

[0488] In one embodiment, a translation enhancer element (TEE) can be incorporated on the 5' end of the stem of a stem loop and a miRNA seed can be incorporated into the stem of the stem loop. In another embodiment, a TEE can be incorporated on the 5' end of the stem of a stem loop, a miRNA seed can be incorporated into the stem of the stem loop and a miRNA binding site can be incorporated into the 3' end of the stem or the sequence after the stem loop. The miRNA seed and the miRNA binding site can be for the same and/or different miRNA sequences.

[0489] In one embodiment, the incorporation of a miRNA sequence and/or a TEE sequence changes the shape of the stem loop region which can increase and/or decrease translation. (see e.g., Kedde et al., "A Pumilio-induced RNA structure switch in p27-3'UTR controls miR-221 and miR-22 accessibility." *Nature Cell Biology*. 2010, incorporated herein by reference in its entirety).

[0490] In one embodiment, the 5'-UTR of a polynucleotide of the disclosure can comprise at least one miRNA sequence. The miRNA sequence can be, but is not limited to, a 19 or 22 nucleotide sequence and/or a miRNA sequence without the seed.

[0491] In one embodiment the miRNA sequence in the 5'UTR can be used to stabilize a polynucleotide of the disclosure described herein.

[0492] In another embodiment, a miRNA sequence in the 5'UTR of a polynucleotide of the disclosure can be used to decrease the accessibility of the site of translation initiation such as, but not limited to a start codon. See, e.g., Matsuda et al., *PLoS One*. 2010 11(5):e15057;

incorporated herein by reference in its entirety, which used antisense locked nucleic acid (LNA) oligonucleotides and exon-junction complexes (EJC) around a start codon (-4 to +37 where the A of the AUG codons is +1) in order to decrease the accessibility to the first start codon (AUG). Matsuda showed that altering the sequence around the start codon with an LNA or EJC affected the efficiency, length and structural stability of a polynucleotide. A polynucleotide of the disclosure can comprise a miRNA sequence, instead of the LNA or EJC sequence described by Matsuda et al, near the site of translation initiation in order to decrease the accessibility to the site of translation initiation. The site of translation initiation can be prior to, after or within the miRNA sequence. As a non-limiting example, the site of translation initiation can be located within a miRNA sequence such as a seed sequence or binding site. As another non-limiting example, the site of translation initiation can be located within a miR-122 sequence such as the seed sequence or the mir-122 binding site.

[0493] In some embodiments, a polynucleotide of the disclosure can include at least one miRNA in order to dampen the antigen presentation by antigen presenting cells. The miRNA can be the complete miRNA sequence, the miRNA seed sequence, the miRNA sequence without the seed, or a combination thereof. As a non-limiting example, a miRNA incorporated into a polynucleotide of the disclosure can be specific to the hematopoietic system. As another non-limiting example, a miRNA incorporated into a polynucleotide of the disclosure to dampen antigen presentation is miR-142-3p.

[0494] In some embodiments, a polynucleotide of the disclosure can include at least one miRNA in order to dampen expression of the encoded polypeptide in a tissue or cell of interest. As a non-limiting example, a polynucleotide of the disclosure can include at least one miR-122 binding site in order to dampen expression of an encoded polypeptide of interest in the liver. As another non-limiting example a polynucleotide of the disclosure can include at least one miR-142-3p binding site, miR-142-3p seed sequence, miR-142-3p binding site without the seed, miR-142-5p binding site, miR-142-5p seed sequence, miR-142-5p binding site without the seed, miR-146 binding site, miR-146 seed sequence and/or miR-146 binding site without the seed sequence.

[0495] In some embodiments, a polynucleotide of the disclosure can comprise at least one miRNA binding site in the 3'UTR in order to selectively degrade mRNA therapeutics in the immune cells to subdue unwanted immunogenic reactions caused by therapeutic delivery. As a non-limiting example, the miRNA binding site can make a polynucleotide of the disclosure more

unstable in antigen presenting cells. Non-limiting examples of these miRNAs include mir-142-5p, mir-142-3p, mir-146a-5p, and mir-146-3p.

[0496] In one embodiment, a polynucleotide of the disclosure comprises at least one miRNA sequence in a region of the polynucleotide that can interact with a RNA binding protein.

[0497] In some embodiments, the polynucleotide of the disclosure (e.g., a RNA, e.g., a mRNA) comprising (i) a sequence-optimized nucleotide sequence (e.g., an ORF) encoding a MCM polypeptide (e.g., the wild-type sequence, functional fragment, or variant thereof) and (ii) a miRNA binding site (e.g., a miRNA binding site that binds to miR-142).

[0498] In some embodiments, the polynucleotide of the disclosure comprises a uracil-modified sequence encoding a MCM polypeptide disclosed herein and a miRNA binding site disclosed herein, e.g., a miRNA binding site that binds to miR-142. In some embodiments, the uracil-modified sequence encoding a MCM polypeptide comprises at least one chemically modified nucleobase, e.g., 5-methoxyuracil. In some embodiments, at least 95% of a type of nucleobase (e.g., uracil) in a uracil-modified sequence encoding a MCM polypeptide of the disclosure are modified nucleobases. In some embodiments, at least 95% of uracil in a uracil-modified sequence encoding a MCM polypeptide is 5-methoxyuridine. In some embodiments, the polynucleotide comprising a nucleotide sequence encoding a MCM polypeptide disclosed herein and a miRNA binding site is formulated with a delivery agent, e.g., a compound having the Formula (I), e.g., any of Compounds 1-147.

3' UTR and the AU Rich Elements

[0499] The disclosure also includes a polynucleotide that comprises both one or more 3' untranslated regions as well as the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0500] Natural or wild type 3' UTRs are known to have stretches of Adenosines and Uridines embedded in them. These AU rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU rich elements (AREs) can be separated into three classes (Chen et al, 1995): Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. C-Myc and MyoD contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A)(U/A) nonamers. Molecules containing this type of AREs include GM-CSF and TNF- α . Class III AREs are less well defined. These U rich regions do not contain an AUUUA motif. c-Jun and Myogenin are

two well-studied examples of this class. Most proteins binding to the AREs are known to destabilize the messenger, whereas members of the ELAV family, most notably HuR, have been documented to increase the stability of mRNA. HuR binds to AREs of all the three classes. Engineering the HuR specific binding sites into the 3' UTR of nucleic acid molecules will lead to HuR binding and thus, stabilization of the message *in vivo*.

[0501] Introduction, removal or modification of 3' UTR AU rich elements (AREs) can be used to modulate the stability of polynucleotides of the disclosure. When engineering specific polynucleotides, one or more copies of an ARE can be introduced to make polynucleotides of the disclosure less stable and thereby curtail translation and decrease production of the resultant protein. Likewise, AREs can be identified and removed or mutated to increase the intracellular stability and thus increase translation and production of the resultant protein. Transfection experiments can be conducted in relevant cell lines, using polynucleotides of the disclosure and protein production can be assayed at various time points post-transfection. For example, cells can be transfected with different ARE-engineering molecules and by using an ELISA kit to the relevant protein and assaying protein produced at 6 hour, 12 hour, 24 hour, 48 hour, and 7 days post-transfection.

Regions having a 5' Cap

[0502] The disclosure also includes a polynucleotide that comprises both a 5' Cap and the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0503] The 5' cap structure of a natural mRNA is involved in nuclear export, increasing mRNA stability and binds the mRNA Cap Binding Protein (CBP), which is responsible for mRNA stability in the cell and translation competency through the association of CBP with poly(A) binding protein to form the mature cyclic mRNA species. The cap further assists the removal of 5' proximal introns during mRNA splicing.

[0504] Endogenous mRNA molecules can be 5'-end capped generating a 5'-ppp-5'-triphosphate linkage between a terminal guanosine cap residue and the 5'-terminal transcribed sense nucleotide of the mRNA molecule. This 5'-guanylate cap can then be methylated to generate an N7-methyl-guanylate residue. The ribose sugars of the terminal and/or anteterminal transcribed nucleotides of the 5' end of the mRNA can optionally also be 2'-O-methylated. 5'-

decapping through hydrolysis and cleavage of the guanylate cap structure can target a nucleic acid molecule, such as an mRNA molecule, for degradation.

[0505] In some embodiments, polynucleotides can be designed to incorporate a cap moiety. Modifications to the polynucleotides of the present disclosure can generate a non-hydrolyzable cap structure preventing decapping and thus increasing mRNA half-life. Because cap structure hydrolysis requires cleavage of 5'-ppp-5' phosphodiester linkages, modified nucleotides can be used during the capping reaction. For example, a Vaccinia Capping Enzyme from New England Biolabs (Ipswich, MA) can be used with α -thio-guanosine nucleotides according to the manufacturer's instructions to create a phosphorothioate linkage in the 5'-ppp-5' cap. Additional modified guanosine nucleotides can be used such as α -methyl-phosphonate and seleno-phosphate nucleotides.

[0506] Additional modifications include, but are not limited to, 2'-O-methylation of the ribose sugars of 5'-terminal and/or 5'-antiterminal nucleotides of the polynucleotide (as mentioned above) on the 2'-hydroxyl group of the sugar ring. Multiple distinct 5'-cap structures can be used to generate the 5'-cap of a nucleic acid molecule, such as a polynucleotide that functions as an mRNA molecule.

[0507] Cap analogs, which herein are also referred to as synthetic cap analogs, chemical caps, chemical cap analogs, or structural or functional cap analogs, differ from natural (*i.e.*, endogenous, wild-type or physiological) 5'-caps in their chemical structure, while retaining cap function. Cap analogs can be chemically (*i.e.*, non-enzymatically) or enzymatically synthesized and/or linked to the polynucleotides of the disclosure.

[0508] For example, the Anti-Reverse Cap Analog (ARCA) cap contains two guanines linked by a 5'-5'-triphosphate group, wherein one guanine contains an N7 methyl group as well as a 3'-O-methyl group (*i.e.*, N7,3'-O-dimethyl-guanosine-5'-triphosphate-5'-guanosine (m⁷G-3'mppp-G; which can equivalently be designated 3' O-Me-m⁷G(5')ppp(5')G). The 3'-O atom of the other, unmodified, guanine becomes linked to the 5'-terminal nucleotide of the capped polynucleotide. The N7- and 3'-O-methylated guanine provides the terminal moiety of the capped polynucleotide.

[0509] Another exemplary cap is mCAP, which is similar to ARCA but has a 2'-O-methyl group on guanosine (*i.e.*, N7,2'-O-dimethyl-guanosine-5'-triphosphate-5'-guanosine, m⁷Gm-ppp-G).

[0510] In some embodiments, the cap is a dinucleotide cap analog. As a non-limiting example, the dinucleotide cap analog can be modified at different phosphate positions with a boranophosphate group or a phosphoroselenoate group such as the dinucleotide cap analogs described in US Patent No. US 8,519,110, the contents of which are herein incorporated by reference in its entirety.

[0511] In another embodiment, the cap is a cap analog is a N7-(4-chlorophenoxyethyl) substituted dinucleotide form of a cap analog known in the art and/or described herein. Non-limiting examples of a N7-(4-chlorophenoxyethyl) substituted dinucleotide form of a cap analog include a N7-(4-chlorophenoxyethyl)-G(5')ppp(5')G and a N7-(4-chlorophenoxyethyl)-m^{3'}-^oG(5')ppp(5')G cap analog (*See, e.g.*, the various cap analogs and the methods of synthesizing cap analogs described in Kore et al. *Bioorganic & Medicinal Chemistry* 2013 21:4570-4574; the contents of which are herein incorporated by reference in its entirety). In another embodiment, a cap analog of the present disclosure is a 4-chloro/bromophenoxyethyl analog.

[0512] While cap analogs allow for the concomitant capping of a polynucleotide or a region thereof, in an *in vitro* transcription reaction, up to 20% of transcripts can remain uncapped. This, as well as the structural differences of a cap analog from an endogenous 5'-cap structures of nucleic acids produced by the endogenous, cellular transcription machinery, can lead to reduced translational competency and reduced cellular stability.

[0513] Polynucleotides of the disclosure can also be capped post-manufacture (whether IVT or chemical synthesis), using enzymes, in order to generate more authentic 5'-cap structures. As used herein, the phrase "more authentic" refers to a feature that closely mirrors or mimics, either structurally or functionally, an endogenous or wild type feature. That is, a "more authentic" feature is better representative of an endogenous, wild-type, natural or physiological cellular function and/or structure as compared to synthetic features or analogs, etc., of the prior art, or which outperforms the corresponding endogenous, wild-type, natural or physiological feature in one or more respects. Non-limiting examples of more authentic 5'cap structures of the present disclosure are those that, among other things, have enhanced binding of cap binding proteins, increased half-life, reduced susceptibility to 5' endonucleases and/or reduced 5'decapping, as compared to synthetic 5'cap structures known in the art (or to a wild-type, natural or physiological 5'cap structure). For example, recombinant Vaccinia Virus Capping Enzyme and recombinant 2'-O-methyltransferase enzyme can create a canonical 5'-5'-triphosphate linkage between the 5'-terminal nucleotide of a polynucleotide and a guanine cap nucleotide wherein the

cap guanine contains an N7 methylation and the 5'-terminal nucleotide of the mRNA contains a 2'-O-methyl. Such a structure is termed the Cap1 structure. This cap results in a higher translational-competency and cellular stability and a reduced activation of cellular pro-inflammatory cytokines, as compared, *e.g.*, to other 5'cap analog structures known in the art. Cap structures include, but are not limited to, 7mG(5')ppp(5')N,pN2p (cap 0), 7mG(5')ppp(5')NlmpNp (cap 1), and 7mG(5')-ppp(5')NlmpN2mp (cap 2).

[0514] As a non-limiting example, capping chimeric polynucleotides post-manufacture can be more efficient as nearly 100% of the chimeric polynucleotides can be capped. This is in contrast to ~80% when a cap analog is linked to a chimeric polynucleotide in the course of an *in vitro* transcription reaction.

[0515] According to the present disclosure, 5' terminal caps can include endogenous caps or cap analogs. According to the present disclosure, a 5' terminal cap can comprise a guanine analog. Useful guanine analogs include, but are not limited to, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

Poly-A tails

[0516] The disclosure also includes a polynucleotide that comprises both a poly-A tail and the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0517] During RNA processing, a long chain of adenine nucleotides (poly-A tail) can be added to a polynucleotide such as an mRNA molecule in order to increase stability. Immediately after transcription, the 3' end of the transcript can be cleaved to free a 3' hydroxyl. Then poly-A polymerase adds a chain of adenine nucleotides to the RNA. The process, called polyadenylation, adds a poly-A tail that can be between, for example, approximately 80 to approximately 250 residues long, including approximately 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240 or 250 residues long.

[0518] PolyA tails can also be added after the construct is exported from the nucleus.

[0519] According to the present disclosure, terminal groups on the poly A tail can be incorporated for stabilization. Polynucleotides of the present disclosure can include des-3' hydroxyl tails. They can also include structural moieties or 2'-Omethyl modifications as taught

by Junjie Li, et al. (Current Biology, Vol. 15, 1501–1507, August 23, 2005, the contents of which are incorporated herein by reference in its entirety).

[0520] The polynucleotides of the present disclosure can be designed to encode transcripts with alternative polyA tail structures including histone mRNA. According to Norbury, "Terminal uridylation has also been detected on human replication-dependent histone mRNAs. The turnover of these mRNAs is thought to be important for the prevention of potentially toxic histone accumulation following the completion or inhibition of chromosomal DNA replication. These mRNAs are distinguished by their lack of a 3' poly(A) tail, the function of which is instead assumed by a stable stem-loop structure and its cognate stem-loop binding protein (SLBP); the latter carries out the same functions as those of PABP on polyadenylated mRNAs" (Norbury, "Cytoplasmic RNA: a case of the tail wagging the dog," Nature Reviews Molecular Cell Biology; AOP, published online 29 August 2013; doi:10.1038/nrm3645) the contents of which are incorporated herein by reference in its entirety.

[0521] Unique poly-A tail lengths provide certain advantages to the polynucleotides of the present disclosure.

[0522] Generally, the length of a poly-A tail, when present, is greater than 30 nucleotides in length. In another embodiment, the poly-A tail is greater than 35 nucleotides in length (*e.g.*, at least or greater than about 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, and 3,000 nucleotides). In some embodiments, the polynucleotide or region thereof includes from about 30 to about 3,000 nucleotides (*e.g.*, from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 750, from 30 to 1,000, from 30 to 1,500, from 30 to 2,000, from 30 to 2,500, from 50 to 100, from 50 to 250, from 50 to 500, from 50 to 750, from 50 to 1,000, from 50 to 1,500, from 50 to 2,000, from 50 to 2,500, from 50 to 3,000, from 100 to 500, from 100 to 750, from 100 to 1,000, from 100 to 1,500, from 100 to 2,000, from 100 to 2,500, from 100 to 3,000, from 500 to 750, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 2,500, from 500 to 3,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 2,500, from 1,000 to 3,000, from 1,500 to 2,000, from 1,500 to 2,500, from 1,500 to 3,000, from 2,000 to 3,000, from 2,000 to 2,500, and from 2,500 to 3,000).

[0523] In some embodiments, the poly-A tail is designed relative to the length of the overall polynucleotide or the length of a particular region of the polynucleotide. This design can be

based on the length of a coding region, the length of a particular feature or region or based on the length of the ultimate product expressed from the polynucleotides.

[0524] In this context, the poly-A tail can be 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100% greater in length than the polynucleotide or feature thereof. The poly-A tail can also be designed as a fraction of the polynucleotides to which it belongs. In this context, the poly-A tail can be 10, 20, 30, 40, 50, 60, 70, 80, or 90% or more of the total length of the construct, a construct region or the total length of the construct minus the poly-A tail. Further, engineered binding sites and conjugation of polynucleotides for Poly-A binding protein can enhance expression.

[0525] Additionally, multiple distinct polynucleotides can be linked together via the PABP (Poly-A binding protein) through the 3'-end using modified nucleotides at the 3'-terminus of the poly-A tail. Transfection experiments can be conducted in relevant cell lines at and protein production can be assayed by ELISA at 12hr, 24hr, 48hr, 72hr and day 7 post-transfection.

[0526] In some embodiments, the polynucleotides of the present disclosure are designed to include a polyA-G Quartet region. The G-quartet is a cyclic hydrogen bonded array of four guanine nucleotides that can be formed by G-rich sequences in both DNA and RNA. In this embodiment, the G-quartet is incorporated at the end of the poly-A tail. The resultant polynucleotide is assayed for stability, protein production and other parameters including half-life at various time points. It has been discovered that the polyA-G quartet results in protein production from an mRNA equivalent to at least 75% of that seen using a poly-A tail of 120 nucleotides alone.

Start codon region

[0527] The disclosure also includes a polynucleotide that comprises both a start codon region and the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide. In some embodiments, the polynucleotides of the present disclosure can have regions that are analogous to or function like a start codon region.

[0528] In some embodiments, the translation of a polynucleotide can initiate on a codon that is not the start codon AUG. Translation of the polynucleotide can initiate on an alternative start codon such as, but not limited to, ACG, AGG, AAG, CTG/CUG, GTG/GUG, ATA/AUA, ATT/AUU, TTG/UUG (see Touriol et al. *Biology of the Cell* 95 (2003) 169-178 and Matsuda and Mauro *PLoS ONE*, 2010 5:11; the contents of each of which are herein incorporated by reference in its entirety). As a non-limiting example, the translation of a polynucleotide begins

on the alternative start codon ACG. As another non-limiting example, polynucleotide translation begins on the alternative start codon CTG or CUG. As yet another non-limiting example, the translation of a polynucleotide begins on the alternative start codon GTG or GUG.

[0529] Nucleotides flanking a codon that initiates translation such as, but not limited to, a start codon or an alternative start codon, are known to affect the translation efficiency, the length and/or the structure of the polynucleotide. (*See, e.g.,* Matsuda and Mauro PLoS ONE, 2010 5:11; the contents of which are herein incorporated by reference in its entirety). Masking any of the nucleotides flanking a codon that initiates translation can be used to alter the position of translation initiation, translation efficiency, length and/or structure of a polynucleotide.

[0530] In some embodiments, a masking agent can be used near the start codon or alternative start codon in order to mask or hide the codon to reduce the probability of translation initiation at the masked start codon or alternative start codon. Non-limiting examples of masking agents include antisense locked nucleic acids (LNA) polynucleotides and exon-junction complexes (EJCs) (*See, e.g.,* Matsuda and Mauro describing masking agents LNA polynucleotides and EJCs (PLoS ONE, 2010 5:11); the contents of which are herein incorporated by reference in its entirety).

[0531] In another embodiment, a masking agent can be used to mask a start codon of a polynucleotide in order to increase the likelihood that translation will initiate on an alternative start codon.

[0532] In some embodiments, a masking agent can be used to mask a first start codon or alternative start codon in order to increase the chance that translation will initiate on a start codon or alternative start codon downstream to the masked start codon or alternative start codon.

[0533] In some embodiments, a start codon or alternative start codon can be located within a perfect complement for a miR binding site. The perfect complement of a miR binding site can help control the translation, length and/or structure of the polynucleotide similar to a masking agent. As a non-limiting example, the start codon or alternative start codon can be located in the middle of a perfect complement for a miR-122 binding site. The start codon or alternative start codon can be located after the first nucleotide, second nucleotide, third nucleotide, fourth nucleotide, fifth nucleotide, sixth nucleotide, seventh nucleotide, eighth nucleotide, ninth nucleotide, tenth nucleotide, eleventh nucleotide, twelfth nucleotide, thirteenth nucleotide, fourteenth nucleotide, fifteenth nucleotide, sixteenth nucleotide, seventeenth nucleotide, eighteenth nucleotide, nineteenth nucleotide, twentieth nucleotide or twenty-first nucleotide.

[0534] In another embodiment, the start codon of a polynucleotide can be removed from the polynucleotide sequence in order to have the translation of the polynucleotide begin on a codon that is not the start codon. Translation of the polynucleotide can begin on the codon following the removed start codon or on a downstream start codon or an alternative start codon. In a non-limiting example, the start codon ATG or AUG is removed as the first 3 nucleotides of the polynucleotide sequence in order to have translation initiate on a downstream start codon or alternative start codon. The polynucleotide sequence where the start codon was removed can further comprise at least one masking agent for the downstream start codon and/or alternative start codons in order to control or attempt to control the initiation of translation, the length of the polynucleotide and/or the structure of the polynucleotide.

Stop Codon Region

[0535] The disclosure also includes a polynucleotide that comprises both a stop codon region and the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide. In some embodiments, the polynucleotides of the present disclosure can include at least two stop codons before the 3' untranslated region (UTR). The stop codon can be selected from TGA, TAA and TAG. In some embodiments, the polynucleotides of the present disclosure include the stop codon TGA and one additional stop codon. In a further embodiment the additional stop codon can be TAA. In another embodiment, the polynucleotides of the present disclosure include three stop codons.

Insertions and Substitutions

[0536] The disclosure also includes a polynucleotide that comprises insertions and/or substitutions in the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide.

[0537] In some embodiments, the 5'UTR of the polynucleotide can be replaced by the insertion of at least one region and/or string of nucleosides of the same base. The region and/or string of nucleotides can include, but is not limited to, at least 3, at least 4, at least 5, at least 6, at least 7 or at least 8 nucleotides and the nucleotides can be natural and/or unnatural. As a non-limiting example, the group of nucleotides can include 5-8 adenine, cytosine, thymine, a string of any of the other nucleotides disclosed herein and/or combinations thereof.

[0538] In some embodiments, the 5'UTR of the polynucleotide can be replaced by the insertion of at least two regions and/or strings of nucleotides of two different bases such as, but not limited to, adenine, cytosine, thymine, any of the other nucleotides disclosed herein and/or combinations thereof. For example, the 5'UTR can be replaced by inserting 5-8 adenine bases followed by the insertion of 5-8 cytosine bases. In another example, the 5'UTR can be replaced by inserting 5-8 cytosine bases followed by the insertion of 5-8 adenine bases.

[0539] In some embodiments, the polynucleotide can include at least one substitution and/or insertion downstream of the transcription start site that can be recognized by an RNA polymerase. As a non-limiting example, at least one substitution and/or insertion can occur downstream of the transcription start site by substituting at least one nucleic acid in the region just downstream of the transcription start site (such as, but not limited to, +1 to +6). Changes to region of nucleotides just downstream of the transcription start site can affect initiation rates, increase apparent nucleotide triphosphate (NTP) reaction constant values, and increase the dissociation of short transcripts from the transcription complex curing initial transcription (Briebe et al, Biochemistry (2002) 41: 5144-5149; herein incorporated by reference in its entirety). The modification, substitution and/or insertion of at least one nucleoside can cause a silent mutation of the sequence or can cause a mutation in the amino acid sequence.

[0540] In some embodiments, the polynucleotide can include the substitution of at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12 or at least 13 guanine bases downstream of the transcription start site.

[0541] In some embodiments, the polynucleotide can include the substitution of at least 1, at least 2, at least 3, at least 4, at least 5 or at least 6 guanine bases in the region just downstream of the transcription start site. As a non-limiting example, if the nucleotides in the region are GGGAGA, the guanine bases can be substituted by at least 1, at least 2, at least 3 or at least 4 adenine nucleotides. In another non-limiting example, if the nucleotides in the region are GGGAGA the guanine bases can be substituted by at least 1, at least 2, at least 3 or at least 4 cytosine bases. In another non-limiting example, if the nucleotides in the region are GGGAGA the guanine bases can be substituted by at least 1, at least 2, at least 3 or at least 4 thymine, and/or any of the nucleotides described herein.

[0542] In some embodiments, the polynucleotide can include at least one substitution and/or insertion upstream of the start codon. For the purpose of clarity, one of skill in the art would appreciate that the start codon is the first codon of the protein coding region whereas the

transcription start site is the site where transcription begins. The polynucleotide can include, but is not limited to, at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7 or at least 8 substitutions and/or insertions of nucleotide bases. The nucleotide bases can be inserted or substituted at 1, at least 1, at least 2, at least 3, at least 4 or at least 5 locations upstream of the start codon. The nucleotides inserted and/or substituted can be the same base (*e.g.*, all A or all C or all T or all G), two different bases (*e.g.*, A and C, A and T, or C and T), three different bases (*e.g.*, A, C and T or A, C and T) or at least four different bases. As a non-limiting example, the guanine base upstream of the coding region in the polynucleotide can be substituted with adenine, cytosine, thymine, or any of the nucleotides described herein. In another non-limiting example the substitution of guanine bases in the polynucleotide can be designed so as to leave one guanine base in the region downstream of the transcription start site and before the start codon (see Esvelt *et al.* Nature (2011) 472(7344):499-503; the contents of which is herein incorporated by reference in its entirety). As a non-limiting example, at least 5 nucleotides can be inserted at 1 location downstream of the transcription start site but upstream of the start codon and the at least 5 nucleotides can be the same base type.

IV. Methods of Making Polynucleotides

[0543] The present disclosure also provides methods for making a polynucleotide disclosed herein or a complement thereof. In some aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein, and encoding an MCM polypeptide or a functional fragment thereof, can be constructed using *in vitro* transcription. In other aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein, and encoding an MCM polypeptide or a functional fragment thereof, can be constructed by chemical synthesis using an oligonucleotide synthesizer. In other aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein, and encoding an MCM polypeptide or a functional fragment thereof is made by using a host cell. In certain aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein, and encoding an MCM polypeptide or a functional fragment thereof is made by one or more combination of the IVT, chemical synthesis, host cell expression, or any other methods known in the art.

[0544] Naturally occurring nucleosides, non-naturally occurring nucleosides, or combinations thereof, can totally or partially naturally replace occurring nucleosides present in the candidate nucleotide sequence and can be incorporated into a sequence-optimized nucleotide

sequence (e.g., an mRNA) encoding an MCM polypeptide. The resultant mRNAs can then be examined for their ability to produce protein and/or produce a therapeutic outcome.

In Vitro Transcription-enzymatic synthesis

[0545] A polynucleotide disclosed herein can be transcribed using an *in vitro* transcription (IVT) system. The system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and a polymerase. The NTPs can be selected from, but are not limited to, those described herein including natural and unnatural (modified) NTPs. The polymerase can be selected from, but is not limited to, T7 RNA polymerase, T3 RNA polymerase and mutant polymerases such as, but not limited to, polymerases able to incorporate modified nucleic acids. See U.S. Publ. No. US20130259923, which is herein incorporated by reference in its entirety.

[0546] The IVT system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and a polymerase. The NTPs can be selected from, but are not limited to, those described herein including natural and unnatural (modified) NTPs. The polymerase can be selected from, but is not limited to, T7 RNA polymerase, T3 RNA polymerase and mutant polymerases such as, but not limited to, polymerases able to incorporate polynucleotides disclosed herein.

[0547] Any number of RNA polymerases or variants can be used in the synthesis of the polynucleotides of the present disclosure.

[0548] RNA polymerases can be modified by inserting or deleting amino acids of the RNA polymerase sequence. As a non-limiting example, the RNA polymerase can be modified to exhibit an increased ability to incorporate a 2'-modified nucleotide triphosphate compared to an unmodified RNA polymerase (see International Publication WO2008078180 and U.S. Patent 8,101,385; herein incorporated by reference in their entireties).

[0549] Variants can be obtained by evolving an RNA polymerase, optimizing the RNA polymerase amino acid and/or nucleic acid sequence and/or by using other methods known in the art. As a non-limiting example, T7 RNA polymerase variants can be evolved using the continuous directed evolution system set out by Esvelt et al. (Nature (2011) 472(7344):499-503; herein incorporated by reference in its entirety) where clones of T7 RNA polymerase can encode at least one mutation such as, but not limited to, lysine at position 93 substituted for threonine (K93T), I4M, A7T, E63V, V64D, A65E, D66Y, T76N, C125R, S128R, A136T, N165S, G175R,

H176L, Y178H, F182L, L196F, G198V, D208Y, E222K, S228A, Q239R, T243N, G259D, M267I, G280C, H300R, D351A, A354S, E356D, L360P, A383V, Y385C, D388Y, S397R, M401T, N410S, K450R, P451T, G452V, E484A, H523L, H524N, G542V, E565K, K577E, K577M, N601S, S684Y, L699I, K713E, N748D, Q754R, E775K, A827V, D851N or L864F. As another non-limiting example, T7 RNA polymerase variants can encode at least mutation as described in U.S. Pub. Nos. 20100120024 and 20070117112; herein incorporated by reference in their entireties. Variants of RNA polymerase can also include, but are not limited to, substitutional variants, conservative amino acid substitution, insertional variants, deletional variants and/or covalent derivatives.

[0550] In one aspect, the polynucleotide can be designed to be recognized by the wild type or variant RNA polymerases. In doing so, the polynucleotide can be modified to contain sites or regions of sequence changes from the wild type or parent chimeric polynucleotide.

[0551] Polynucleotide or nucleic acid synthesis reactions can be carried out by enzymatic methods utilizing polymerases. Polymerases catalyze the creation of phosphodiester bonds between nucleotides in a polynucleotide or nucleic acid chain. Currently known DNA polymerases can be divided into different families based on amino acid sequence comparison and crystal structure analysis. DNA polymerase I (pol I) or A polymerase family, including the Klenow fragments of *E. Coli*, *Bacillus* DNA polymerase I, *Thermus aquaticus* (Taq) DNA polymerases, and the T7 RNA and DNA polymerases, is among the best studied of these families. Another large family is DNA polymerase α (pol α) or B polymerase family, including all eukaryotic replicating DNA polymerases and polymerases from phages T4 and RB69. Although they employ similar catalytic mechanism, these families of polymerases differ in substrate specificity, substrate analog-incorporating efficiency, degree and rate for primer extension, mode of DNA synthesis, exonuclease activity, and sensitivity against inhibitors.

[0552] DNA polymerases are also selected based on the optimum reaction conditions they require, such as reaction temperature, pH, and template and primer concentrations. Sometimes a combination of more than one DNA polymerases is employed to achieve the desired DNA fragment size and synthesis efficiency. For example, Cheng et al. increase pH, add glycerol and dimethyl sulfoxide, decrease denaturation times, increase extension times, and utilize a secondary thermostable DNA polymerase that possesses a 3' to 5' exonuclease activity to effectively amplify long targets from cloned inserts and human genomic DNA. (Cheng et al., PNAS, Vol. 91, 5695-5699 (1994), the contents of which are incorporated herein by reference in

their entirety). RNA polymerases from bacteriophage T3, T7, and SP6 have been widely used to prepare RNAs for biochemical and biophysical studies. RNA polymerases, capping enzymes, and poly-A polymerases are disclosed in the co-pending International Publication No. WO2014028429, the contents of which are incorporated herein by reference in their entirety.

[0553] In one aspect, the RNA polymerase which can be used in the synthesis of the polynucleotides described herein is a Syn5 RNA polymerase. (see Zhu et al. Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety). The Syn5 RNA polymerase was recently characterized from marine cyanophage Syn5 by Zhu et al. where they also identified the promoter sequence (see Zhu et al. Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety). Zhu et al. found that Syn5 RNA polymerase catalyzed RNA synthesis over a wider range of temperatures and salinity as compared to T7 RNA polymerase. Additionally, the requirement for the initiating nucleotide at the promoter was found to be less stringent for Syn5 RNA polymerase as compared to the T7 RNA polymerase making Syn5 RNA polymerase promising for RNA synthesis.

[0554] In one aspect, a Syn5 RNA polymerase can be used in the synthesis of the polynucleotides described herein. As a non-limiting example, a Syn5 RNA polymerase can be used in the synthesis of the polynucleotide requiring a precise 3'-termini.

[0555] In one aspect, a Syn5 promoter can be used in the synthesis of the polynucleotides. As a non-limiting example, the Syn5 promoter can be 5'-ATTGGGCACCCGTAAGGG-3' as described by Zhu et al. (Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety).

[0556] In one aspect, a Syn5 RNA polymerase can be used in the synthesis of polynucleotides comprising at least one chemical modification described herein and/or known in the art. (see e.g., the incorporation of pseudo-UTP and 5Me-CTP described in Zhu et al. Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety).

[0557] In one aspect, the polynucleotides described herein can be synthesized using a Syn5 RNA polymerase which has been purified using modified and improved purification procedure described by Zhu et al. (Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety).

[0558] Various tools in genetic engineering are based on the enzymatic amplification of a target gene which acts as a template. For the study of sequences of individual genes or specific regions of interest and other research needs, it is necessary to generate multiple copies of a target

gene from a small sample of polynucleotides or nucleic acids. Such methods can be applied in the manufacture of the polynucleotides of the disclosure.

[0559] Polymerase chain reaction (PCR) has wide applications in rapid amplification of a target gene, as well as genome mapping and sequencing. The key components for synthesizing DNA comprise target DNA molecules as a template, primers complementary to the ends of target DNA strands, deoxynucleoside triphosphates (dNTPs) as building blocks, and a DNA polymerase. As PCR progresses through denaturation, annealing and extension steps, the newly produced DNA molecules can act as a template for the next circle of replication, achieving exponentially amplification of the target DNA. PCR requires a cycle of heating and cooling for denaturation and annealing. Variations of the basic PCR include asymmetric PCR [Innis et al., PNAS, vol. 85, 9436-9440 (1988)], inverse PCR [Ochman et al., Genetics, vol. 120(3), 621-623, (1988)], reverse transcription PCR (RT-PCR) (Freeman et al., BioTechniques, vol. 26(1), 112-22, 124-5 (1999), the contents of which are incorporated herein by reference in their entirety and so on). In RT-PCR, a single stranded RNA is the desired target and is converted to a double stranded DNA first by reverse transcriptase.

[0560] A variety of isothermal in vitro nucleic acid amplification techniques have been developed as alternatives or complements of PCR. For example, strand displacement amplification (SDA) is based on the ability of a restriction enzyme to form a nick. (Walker et al., PNAS, vol. 89, 392-396 (1992), the contents of which are incorporated herein by reference in their entirety)). A restriction enzyme recognition sequence is inserted into an annealed primer sequence. Primers are extended by a DNA polymerase and dNTPs to form a duplex. Only one strand of the duplex is cleaved by the restriction enzyme. Each single strand chain is then available as a template for subsequent synthesis. SDA does not require the complicated temperature control cycle of PCR.

[0561] Nucleic acid sequence-based amplification (NASBA), also called transcription mediated amplification (TMA), is also an isothermal amplification method that utilizes a combination of DNA polymerase, reverse transcriptase, RNase H, and T7 RNA polymerase. [Compton, Nature, vol. 350, 91-92 (1991)] the contents of which are incorporated herein by reference in their entirety. A target RNA is used as a template and a reverse transcriptase synthesizes its complementary DNA strand. RNase H hydrolyzes the RNA template, making space for a DNA polymerase to synthesize a DNA strand complementary to the first DNA strand which is complementary to the RNA target, forming a DNA duplex. T7 RNA polymerase

continuously generates complementary RNA strands of this DNA duplex. These RNA strands act as templates for new cycles of DNA synthesis, resulting in amplification of the target gene.

[0562] Rolling-circle amplification (RCA) amplifies a single stranded circular polynucleotide and involves numerous rounds of isothermal enzymatic synthesis where Φ 29 DNA polymerase extends a primer by continuously progressing around the polynucleotide circle to replicate its sequence over and over again. Therefore, a linear copy of the circular template is achieved. A primer can then be annealed to this linear copy and its complementary chain can be synthesized. [See Lizardi et al., *Nature Genetics*, vol. 19, 225-232 (1998)] the contents of which are incorporated herein by reference in their entirety. A single stranded circular DNA can also serve as a template for RNA synthesis in the presence of an RNA polymerase. (Daubendiek et al., *JACS*, vol. 117, 7818-7819 (1995), the contents of which are incorporated herein by reference in their entirety). An inverse rapid amplification of cDNA ends (RACE) RCA is described by Polidoros et al. A messenger RNA (mRNA) is reverse transcribed into cDNA, followed by RNase H treatment to separate the cDNA. The cDNA is then circularized by CircLigase into a circular DNA. The amplification of the resulting circular DNA is achieved with RCA. (Polidoros et al., *BioTechniques*, vol. 41, 35-42 (2006), the contents of which are incorporated herein by reference in their entirety).

[0563] Any of the foregoing methods can be utilized in the manufacture of one or more regions of the polynucleotides of the present disclosure.

[0564] Assembling polynucleotides or nucleic acids by a ligase is also widely used. DNA or RNA ligases promote intermolecular ligation of the 5' and 3' ends of polynucleotide chains through the formation of a phosphodiester bond. Ligase chain reaction (LCR) is a promising diagnosing technique based on the principle that two adjacent polynucleotide probes hybridize to one strand of a target gene and couple to each other by a ligase. If a target gene is not present, or if there is a mismatch at the target gene, such as a single-nucleotide polymorphism (SNP), the probes cannot ligase. (Wiedmann et al., *PCR Methods and Application*, vol.3 (4), s51-s64 (1994), the contents of which are incorporated herein by reference in their entirety). LCR can be combined with various amplification techniques to increase sensitivity of detection or to increase the amount of products if it is used in synthesizing polynucleotides and nucleic acids.

[0565] Several library preparation kits for nucleic acids are now commercially available. They include enzymes and buffers to convert a small amount of nucleic acid samples into an indexed library for downstream applications. For example, DNA fragments can be placed in a

NEBNext® ULTRATM DNA Library Prep Kit by NEWENGLAND BIOLABS® for end preparation, ligation, size selection, clean-up, PCR amplification and final clean-up.

[0566] Continued development is going on to improvement the amplification techniques. For example, US Pat. 8,367,328 to Asada et al. the contents of which are incorporated herein by reference in their entirety, teaches utilizing a reaction enhancer to increase the efficiency of DNA synthesis reactions by DNA polymerases. The reaction enhancer comprises an acidic substance or cationic complexes of an acidic substance. US Pat. 7,384,739 to Kitabayashi et al. the contents of which are incorporated herein by reference in their entirety, teaches a carboxylate ion-supplying substance that promotes enzymatic DNA synthesis, wherein the carboxylate ion-supplying substance is selected from oxalic acid, malonic acid, esters of oxalic acid, esters of malonic acid, salts of malonic acid, and esters of maleic acid. US Pat. 7,378,262 to Sobek et al. the contents of which are incorporated herein by reference in their entirety, discloses an enzyme composition to increase fidelity of DNA amplifications. The composition comprises one enzyme with 3' exonuclease activity but no polymerase activity and another enzyme that is a polymerase. Both of the enzymes are thermostable and are reversibly modified to be inactive at lower temperatures.

[0567] US Pat. No. 7,550,264 to Getts et al. teaches multiple round of synthesis of sense RNA molecules are performed by attaching oligodeoxynucleotides tails onto the 3' end of cDNA molecules and initiating RNA transcription using RNA polymerase, the contents of which are incorporated herein by reference in their entirety. US Pat. Publication No. 2013/0183718 to Rohayem teaches RNA synthesis by RNA-dependent RNA polymerases (RdRp) displaying an RNA polymerase activity on single-stranded DNA templates, the contents of which are incorporated herein by reference in their entirety. Oligonucleotides with non-standard nucleotides can be synthesized with enzymatic polymerization by contacting a template comprising non-standard nucleotides with a mixture of nucleotides that are complementary to the nucleotides of the template as disclosed in US Pat. No. 6,617,106 to Benner, the contents of which are incorporated herein by reference in their entirety.

Chemical synthesis

[0568] Standard methods can be applied to synthesize an isolated polynucleotide sequence encoding an isolated polypeptide of interest. For example, a single DNA or RNA oligomer containing a codon-optimized nucleotide sequence coding for the particular isolated polypeptide

can be synthesized. In other aspects, several small oligonucleotides coding for portions of the desired polypeptide can be synthesized and then ligated. In some aspects, the individual oligonucleotides typically contain 5' or 3' overhangs for complementary assembly.

[0569] A polynucleotide disclosed herein (e.g., mRNA) can be chemically synthesized using chemical synthesis methods and potential nucleobase substitutions known in the art. See, for example, International Publication Nos. WO2014093924, WO2013052523; WO2013039857, WO2012135805, WO2013151671; U.S. Publ. No. US20130115272; or U.S. Pat. Nos. US8999380, US8710200, all of which are herein incorporated by reference in their entireties.

V. Purification and Quantitation of Polynucleotides

Purification

[0570] Purification of the polynucleotides described herein (*i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide) can include, but is not limited to, polynucleotide clean-up, quality assurance and quality control. Clean-up can be performed by methods known in the arts such as, but not limited to, AGENCOURT® beads (Beckman Coulter Genomics, Danvers, MA), poly-T beads, LNATM oligo-T capture probes (EXIQON® Inc., Vedbaek, Denmark) or HPLC based purification methods such as, but not limited to, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC). The term "purified" when used in relation to a polynucleotide such as a "purified polynucleotide" refers to one that is separated from at least one contaminant. As used herein, a "contaminant" is any substance that makes another unfit, impure or inferior. Thus, a purified polynucleotide (e.g., DNA and RNA) is present in a form or setting different from that in which it is found in nature, or a form or setting different from that which existed prior to subjecting it to a treatment or purification method.

[0571] In some embodiments, purification of a polynucleotide of the disclosure removes impurities that can reduce or remove an unwanted immune response, *e.g.*, reducing cytokine activity.

[0572] In some embodiments, the polynucleotide of the disclosure is purified prior to administration using column chromatography (*e.g.*, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), or (LCMS)). In some embodiments, a column chromatography (*e.g.*, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and

hydrophobic interaction HPLC (HIC-HPLC), or (LCMS)) purified polynucleotide, that encodes an MCM polypeptide disclosed herein increases expression of MCM compared to polynucleotides encoding MCM purified by a different purification method. In some embodiments, a column chromatography (e.g., strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), or (LCMS)) purified polynucleotide encodes an MCM polypeptide or functional fragment thereof comprising one or more of the point mutations V69, T499, H532, A598, and V671. In some embodiments, the RP-HPLC purified polynucleotide increases MCM expression (e.g., by 20-50%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%).

[0573] A quality assurance and/or quality control check can be conducted using methods such as, but not limited to, gel electrophoresis, UV absorbance, or analytical HPLC.

[0574] In another embodiment, the polynucleotides can be sequenced by methods including, but not limited to reverse-transcriptase-PCR.

Quantification

[0575] In some embodiments, the polynucleotides of the present disclosure (*i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide) can be quantified in exosomes or when derived from one or more bodily fluid. As used herein "bodily fluids" include peripheral blood, serum, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, and umbilical cord blood. Alternatively, exosomes can be retrieved from an organ selected from the group consisting of lung, heart, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colon, breast, prostate, brain, esophagus, liver, and placenta.

[0576] In the exosome quantification method, a sample of not more than 2mL is obtained from the subject and the exosomes isolated by size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof. In the analysis,

the level or concentration of a polynucleotide can be an expression level, presence, absence, truncation or alteration of the administered construct. It is advantageous to correlate the level with one or more clinical phenotypes or with an assay for a human disease biomarker. The assay can be performed using construct specific probes, cytometry, qRT-PCR, real-time PCR, PCR, flow cytometry, electrophoresis, mass spectrometry, or combinations thereof while the exosomes can be isolated using immunohistochemical methods such as enzyme linked immunosorbent assay (ELISA) methods. Exosomes can also be isolated by size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof.

[0577] These methods afford the investigator the ability to monitor, in real time, the level of polynucleotides remaining or delivered. This is possible because the polynucleotides of the present disclosure differ from the endogenous forms due to the structural or chemical modifications.

[0578] In some embodiments, the polynucleotide can be quantified using methods such as, but not limited to, ultraviolet visible spectroscopy (UV/Vis). A non-limiting example of a UV/Vis spectrometer is a NANODROP® spectrometer (ThermoFisher, Waltham, MA). The quantified polynucleotide can be analyzed in order to determine if the polynucleotide can be of proper size, check that no degradation of the polynucleotide has occurred. Degradation of the polynucleotide can be checked by methods such as, but not limited to, agarose gel electrophoresis, HPLC based purification methods such as, but not limited to, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), liquid chromatography-mass spectrometry (LCMS), capillary electrophoresis (CE) and capillary gel electrophoresis (CGE).

VI. Pharmaceutical Compositions

[0579] The disclosure includes pharmaceutical compositions that comprise the polynucleotide described herein, *i.e.*, a polynucleotide comprising an ORF encoding an MCM polypeptide. In some embodiments, the formulation can contain polynucleotide encoding wild type MCM or MCM comprising a nucleotide sequence having significant sequence similarity to a polynucleotide selected from the group of SEQ ID NOs: 1-207, 732-765, and 772, wherein the ORF encodes an MCM polypeptide.

[0580] Pharmaceutical compositions can optionally comprise one or more additional active substances, *e.g.*, therapeutically and/or prophylactically active substances. Pharmaceutical compositions of the present disclosure can be sterile and/or pyrogen-free. General considerations in the formulation and/or manufacture of pharmaceutical agents can be found, for example, in *Remington: The Science and Practice of Pharmacy* 21st ed., Lippincott Williams & Wilkins, 2005 (incorporated herein by reference in its entirety). For the purposes of the present disclosure, the phrase "active ingredient" generally refers to polynucleotides to be delivered as described herein.

[0581] Formulations of the pharmaceutical compositions described herein can be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of associating the active ingredient with an excipient and/or one or more other accessory ingredients.

[0582] A pharmaceutical composition in accordance with the present disclosure can be prepared, packaged, and/or sold in bulk, as a single unit dose, and/or as a plurality of single unit doses. As used herein, a "unit dose" refers to a discrete amount of the pharmaceutical composition comprising a predetermined amount of the active ingredient. The amount of the active ingredient is generally equal to the dosage of the active ingredient that would be administered to a subject and/or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

[0583] Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the present disclosure can vary, depending upon the identity, size, and/or condition of the subject being treated and further depending upon the route by which the composition is to be administered. For example, the composition can comprise between 0.1% and 99% (w/w) of the active ingredient. By way of example, the composition can comprise between 0.1% and 100%, *e.g.*, between 0.5 and 50%, between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

[0584] In some embodiments, the formulations described herein can contain at least one polynucleotide. As a non-limiting example, the formulations can contain 1, 2, 3, 4 or 5 polynucleotides.

[0585] In some embodiments, the formulations described herein can comprise more than one type of polynucleotide. In some embodiments, the formulation can comprise a polynucleotide in linear and circular form. In another embodiment, the formulation can comprise a circular

polynucleotide and an IVT polynucleotide. In yet another embodiment, the formulation can comprise an IVT polynucleotide, a chimeric polynucleotide and a circular polynucleotide.

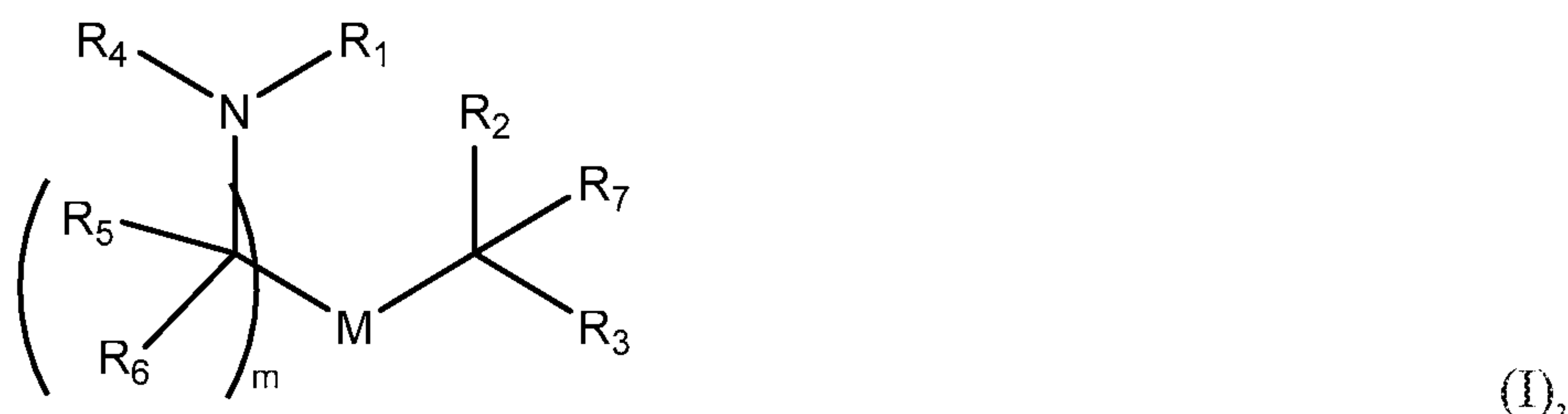
[0586] In some embodiments, compositions are administered to humans, human patients or subjects. Although the descriptions of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions that are suitable for administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to any other animal, *e.g.*, to non-human animals, *e.g.* non-human mammals. Modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various animals is well understood, and the ordinarily skilled veterinary pharmacologist can design and/or perform such modification with merely ordinary, if any, experimentation. Subjects to which administration of the pharmaceutical compositions is contemplated include, but are not limited to, humans and/or other primates; mammals, including commercially relevant mammals such as cattle, pigs, horses, sheep, cats, dogs, mice, and/or rats; and/or birds, including commercially relevant birds such as poultry, chickens, ducks, geese, and/or turkeys.

Delivery Agents

a. Lipid Compound

[0587] The present disclosure provides pharmaceutical compositions with advantageous properties. In particular, the present application provides pharmaceutical compositions comprising:

- (a) a polynucleotide comprising an ORF encoding an MCM polypeptide; and
- (b) a lipid compound having the formula (I)



wherein

R₁ is selected from the group consisting of C₅₋₂₀ alkyl, C₅₋₂₀ alkenyl, -R*YR'', -YR'', and -R''M'R';

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, $-R^*YR''$, $-YR''$, and $-R^*OR''$, or R_2 and R_3 , together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a carbocycle, heterocycle, $-OR$, $-O(CH_2)_nN(R)_2$, $-C(O)OR$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-N(R)_2$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, and $-C(R)N(R)_2C(O)OR$, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, $-R^*YR''$, $-YR''$, and H;

each R'' is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R^* is independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or stereoisomers thereof, wherein alkyl and alkenyl groups can be linear or branched.

[0588] In some embodiments, a subset of compounds of Formula (I) includes those in which when R_4 is $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, or $-CQ(R)_2$, then (i) Q is not $-N(R)_2$ when n is 1, 2, 3, 4 or 5, or (ii) Q is not 5, 6, or 7-membered heterocycloalkyl when n is 1 or 2.

[0589] In another embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-20} alkyl, C_{5-20} alkenyl, $-R^*YR''$, $-YR''$, and $-R''M'R'$;

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, $-R^*YR''$, $-YR''$, and $-R^*OR''$, or R_2 and R_3 , together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a C_{3-6} carbocycle, a 5- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S, $-OR$, $-O(CH_2)_nN(R)_2$, $-C(O)OR$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-CRN(R)_2C(O)OR$, and a 5- to 14-membered heterocycloalkyl having one or more heteroatoms selected from N, O, and S which is substituted with one or more substituents selected from oxo ($=O$), OH, amino, and C_{1-3} alkyl, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, $-R^*YR''$, $-YR''$, and H;

each R'' is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R^* is independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and
 m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,
 or salts or stereoisomers thereof.

[0590] In yet another embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-20} alkyl, C_{5-20} alkenyl, $-R^*YR''$, $-YR''$, and $-R''M'R'$;

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, $-R^*YR''$, $-YR''$, and $-R^*OR''$, or R_2 and R_3 , together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a C_{3-6} carbocycle, a 5- to 14-membered heterocycle having one or more heteroatoms selected from N, O, and S, $-OR$, $-O(CH_2)_nN(R)_2$, $-C(O)OR$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-CRN(R)_2C(O)OR$, and each n is independently selected from 1, 2, 3, 4, and 5; and when Q is a 5- to 14-membered heterocycle and (i) R_4 is $-(CH_2)_nQ$ in which n is 1 or 2, or (ii) R_4 is $-(CH_2)_nCHQR$ in which n is 1, or (iii) R_4 is $-CHQR$, and $-CQ(R)_2$, then Q is either a 5- to 14-membered heteroaryl or 8- to 14-membered heterocycloalkyl;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, $-R^*YR''$, $-YR''$, and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or stereoisomers thereof.

[0591] In still another embodiments, another subset of compounds of Formula (I) includes those in which

R₁ is selected from the group consisting of C₅₋₂₀ alkyl, C₅₋₂₀ alkenyl, -R*YR'', -YR'', and -R''M'R';

R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, C₂₋₁₄ alkenyl, -R*YR'', -YR'', and -R*OR'', or R₂ and R₃, together with the atom to which they are attached, form a heterocycle or carbocycle;

R₄ is selected from the group consisting of a C₃₋₆ carbocycle, -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, -CQ(R)₂, and unsubstituted C₁₋₆ alkyl, where Q is selected from a C₃₋₆ carbocycle, a 5- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S, -OR, -O(CH₂)_nN(R)₂, -C(O)OR, -OC(O)R, -CX₃, -CX₂H, -CXH₂, -CN, -C(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)C(O)N(R)₂, -N(R)C(S)N(R)₂, -CRN(R)₂C(O)OR, and each n is independently selected from 1, 2, 3, 4, and 5;

each R₅ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R₆ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, an aryl group, and a heteroaryl group;

R₇ is selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or stereoisomers thereof.

In yet another embodiments, another subset of compounds of Formula (I) includes those in which

R₁ is selected from the group consisting of C₅₋₂₀ alkyl, C₅₋₂₀ alkenyl, -R*YR'', -YR'', and -R''M'R';

R₂ and R₃ are independently selected from the group consisting of H, C₂₋₁₄ alkyl, C₂₋₁₄ alkenyl, -R*YR'', -YR'', and -R*OR'', or R₂ and R₃, together with the atom to which they are attached, form a heterocycle or carbocycle;

R₄ is -(CH₂)_nQ or -(CH₂)_nCHQR, where Q is -N(R)₂, and n is selected from 3, 4, and 5;

each R₅ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R₆ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, an aryl group, and a heteroaryl group;

R₇ is selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₁₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or stereoisomers thereof.

[0592] In still another embodiments, another subset of compounds of Formula (I) includes those in which

R₁ is selected from the group consisting of C₅₋₂₀ alkyl, C₅₋₂₀ alkenyl, -R*YR'', -YR'', and -R''M'R';

R₂ and R₃ are independently selected from the group consisting of C₁₋₁₄ alkyl, C₂₋₁₄ alkenyl, -R*YR'', -YR'', and -R*OR'', or R₂ and R₃, together with the atom to which they are attached, form a heterocycle or carbocycle;

R₄ is selected from the group consisting of -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, and -CQ(R)₂, where Q is -N(R)₂, and n is selected from 1, 2, 3, 4, and 5;

each R₅ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R₆ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, an aryl group, and a heteroaryl group;

R₇ is selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₁₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and
 m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,
 or salts or stereoisomers thereof.

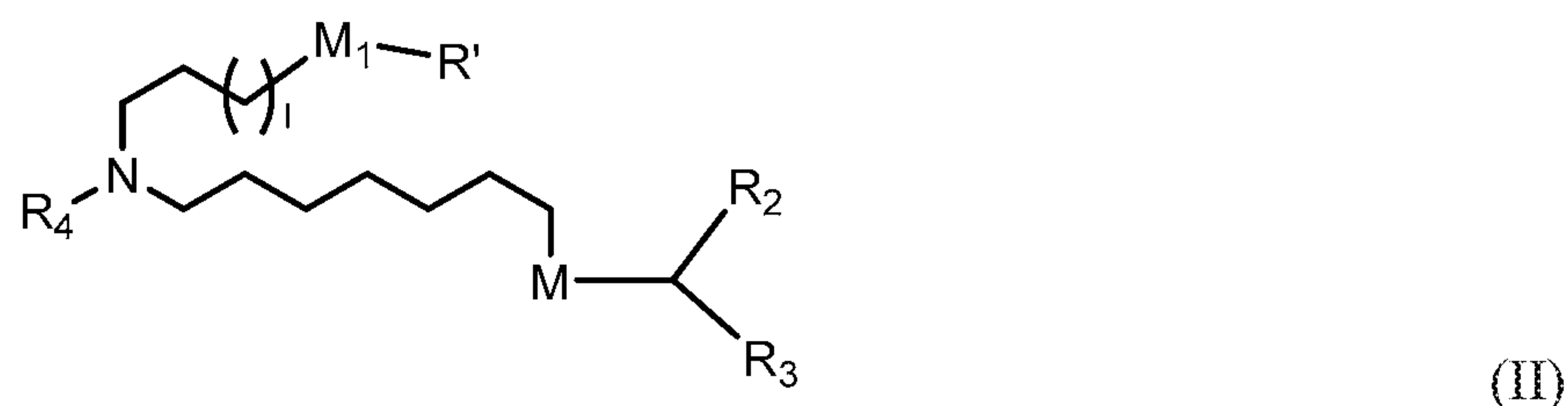
[0593] In certain embodiments, a subset of compounds of Formula (I) includes those of Formula (IA):



or a salt or stereoisomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; m is selected from 5, 6, 7, 8, and 9; M₁ is a bond or M'; R₄ is unsubstituted C₁₋₃ alkyl, or -(CH₂)_nQ, in which Q is OH, -NHC(S)N(R)₂, or -NHC(O)N(R)₂; M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -P(O)(OR')O-, an aryl group, and a heteroaryl group; and

R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl.

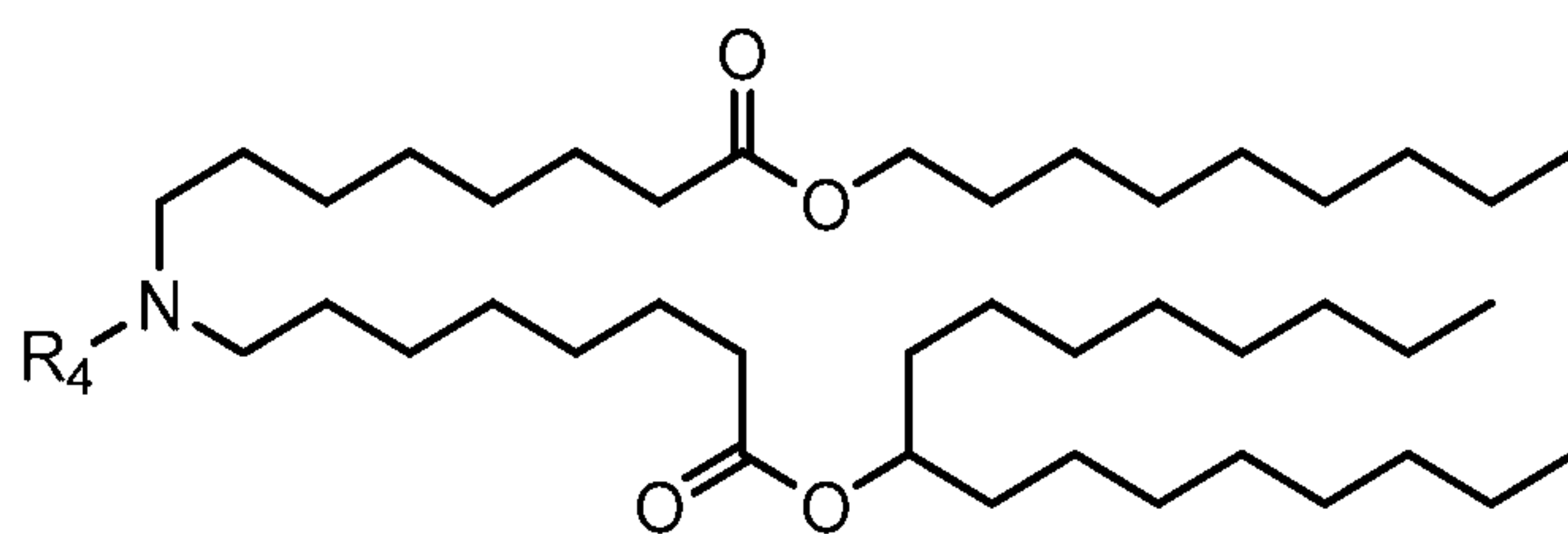
[0594] In certain embodiments, a subset of compounds of Formula (I) includes those of Formula (II):



or a salt or stereoisomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; M₁ is a bond or M'; R₄ is unsubstituted C₁₋₃ alkyl, or -(CH₂)_nQ, in which n is 2, 3, or 4, and Q is OH, -NHC(S)N(R)₂, or -NHC(O)N(R)₂; M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -P(O)(OR')O-, an aryl group, and a heteroaryl group; and

R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl.

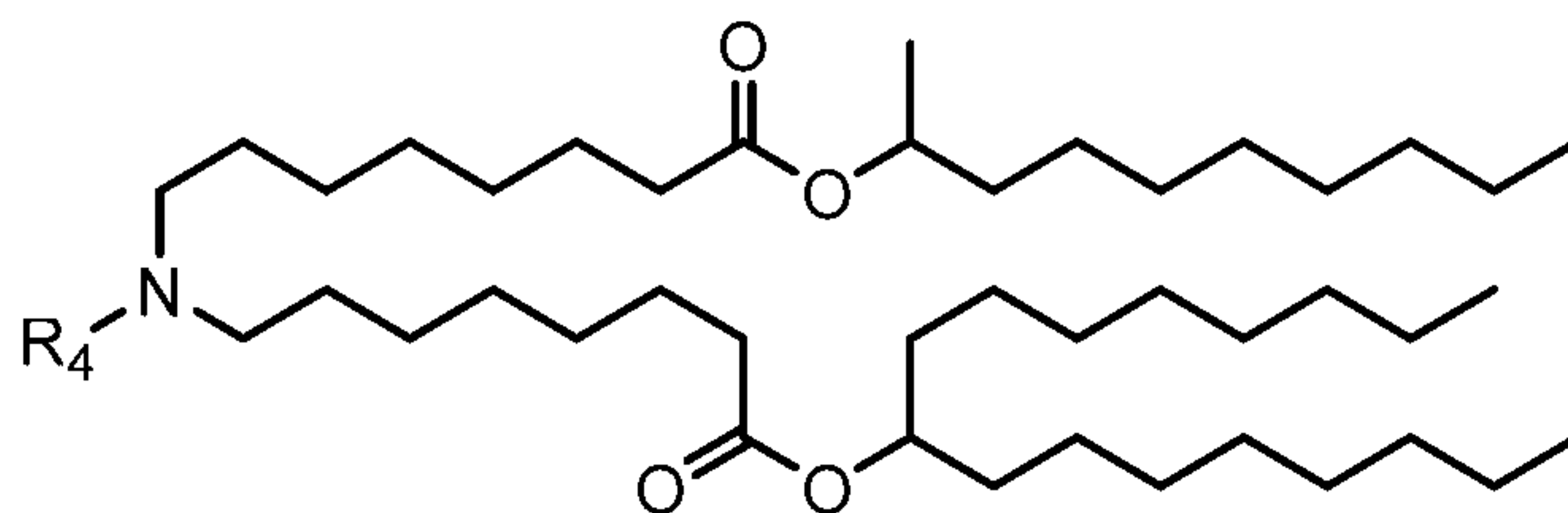
[0595] In some embodiments, the compound of formula (I) is of the formula (IIa),



(IIa),

or a salt thereof, wherein R_4 is as described above.

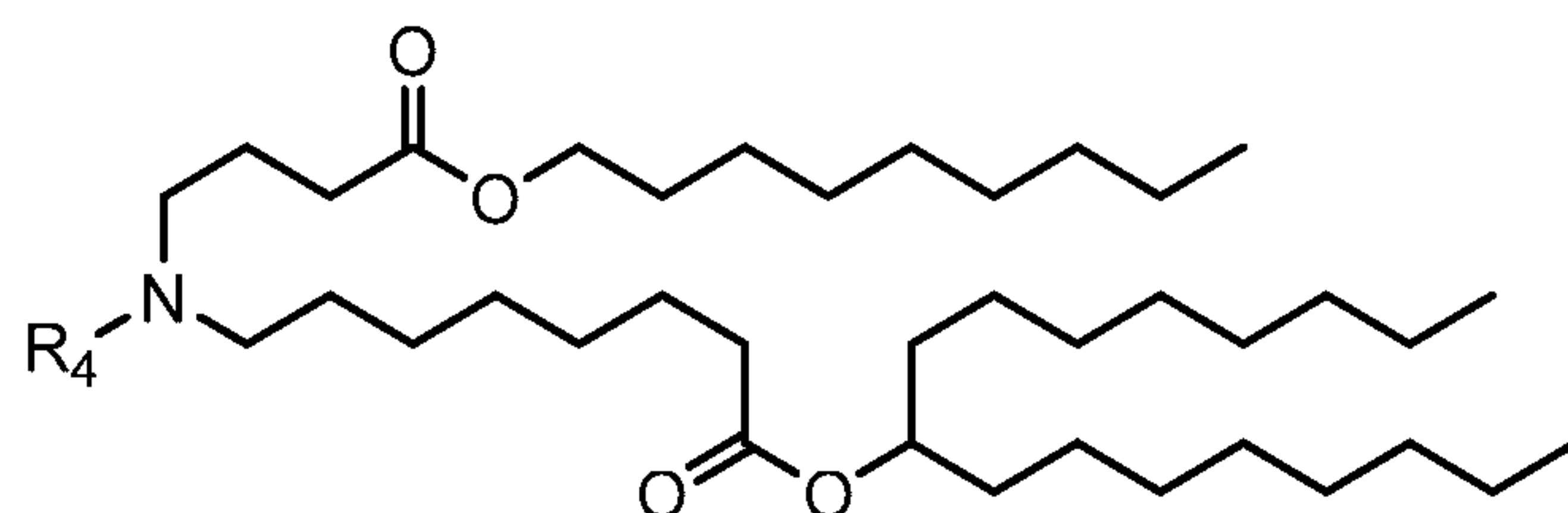
[0596] In some embodiments, the compound of formula (I) is of the formula (IIb),



(IIb),

or a salt thereof, wherein R_4 is as described above.

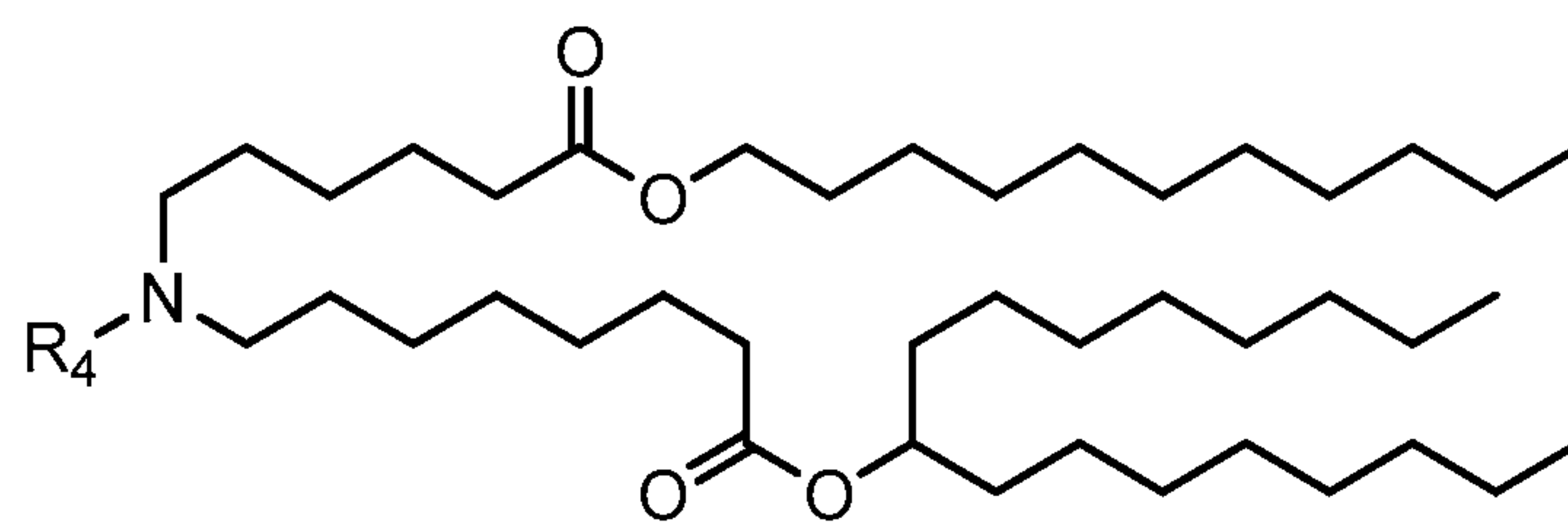
[0597] In some embodiments, the compound of formula (I) is of the formula (IIc),



(IIc),

or a salt thereof, wherein R_4 is as described above.

[0598] In some embodiments, the compound of formula (I) is of the formula (IIe):



(IIe),

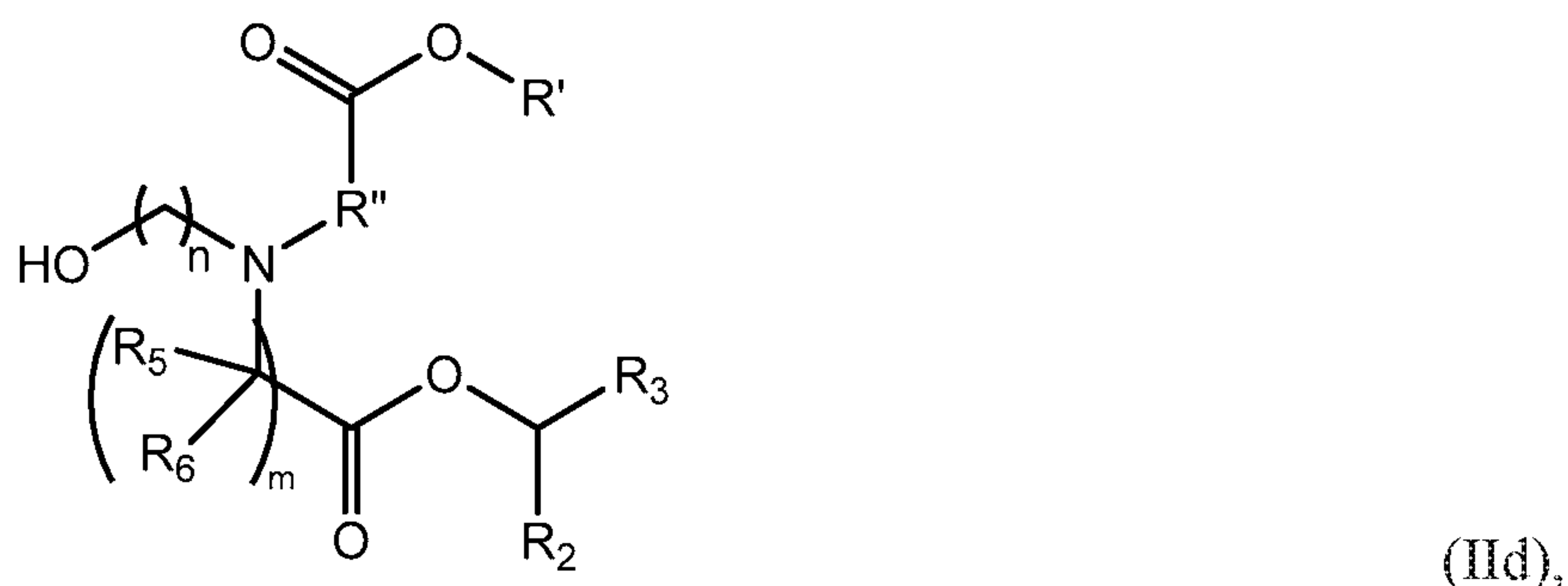
or a salt thereof, wherein R_4 is as described above.

[0599] In some embodiments, the compound of formula (IIa), (IIb), (IIc), or (IIe) comprises an R_4 which is selected from $-(CH_2)_nQ$ and $-(CH_2)_nCHQR$, wherein Q, R and n are as defined above.

[0600] In some embodiments, Q is selected from the group consisting of -OR, -OH, $-O(CH_2)_nN(R)_2$, $-OC(O)R$, $-CX_3$, $-CN$, $-N(R)C(O)R$, $-N(H)C(O)R$, $-N(R)S(O)_2R$, $-N(H)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(H)C(O)N(R)_2$, $-N(H)C(O)N(H)(R)$, $-N(R)C(S)N(R)_2$, $-N(H)C(S)N(R)_2$,

-N(H)C(S)N(H)(R), and a heterocycle, wherein R is as defined above. In some aspects, n is 1 or 2. In some embodiments, Q is OH, -NHC(S)N(R)₂, or -NHC(O)N(R)₂.

[0601] In some embodiments, the compound of formula (I) is of the formula (IId),



or a salt thereof, wherein R₂ and R₃ are independently selected from the group consisting of C₅₋₁₄ alkyl and C₅₋₁₄ alkenyl, n is selected from 2, 3, and 4, and R', R'', R₅, R₆ and m are as defined above.

[0602] In some aspects of the compound of formula (IId), R₂ is C₈ alkyl. In some aspects of the compound of formula (IId), R₃ is C₅-C₉ alkyl. In some aspects of the compound of formula (IId), m is 5, 7, or 9. In some aspects of the compound of formula (IId), each R₅ is H. In some aspects of the compound of formula (IId), each R₆ is H.

[0603] In another aspect, the present application provides a lipid composition (e.g., a lipid nanoparticle (LNP)) comprising: (1) a compound having the formula (I); (2) optionally a helper lipid (e.g. a phospholipid); (3) optionally a structural lipid (e.g. a sterol); (4) optionally a lipid conjugate (e.g. a PEG-lipid); and (5) optionally a quaternary amine compound. In exemplary embodiments, the lipid composition (e.g., LNP) further comprises a polynucleotide encoding an MCM polypeptide, e.g., a polynucleotide encapsulated therein.

[0604] As used herein, the term “alkyl” or “alkyl group” means a linear or branched, saturated hydrocarbon including one or more carbon atoms (e.g., one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or more carbon atoms).

[0605] The notation “C₁₋₁₄ alkyl” means a linear or branched, saturated hydrocarbon including 1-14 carbon atoms. An alkyl group can be optionally substituted.

[0606] As used herein, the term “alkenyl” or “alkenyl group” means a linear or branched hydrocarbon including two or more carbon atoms (e.g., two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or more carbon atoms) and at least one double bond.

[0607] The notation “C₂₋₁₄ alkenyl” means a linear or branched hydrocarbon including 2-14 carbon atoms and at least one double bond. An alkenyl group can include one, two, three, four, or more double bonds. For example, C₁₈ alkenyl can include one or more double bonds. A C₁₈ alkenyl group including two double bonds can be a linoleyl group. An alkenyl group can be optionally substituted.

[0608] As used herein, the term “carbocycle” or “carbocyclic group” means a mono- or multi-cyclic system including one or more rings of carbon atoms. Rings can be three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, or fifteen membered rings.

[0609] The notation “C₃₋₆ carbocycle” means a carbocycle including a single ring having 3-6 carbon atoms. Carbocycles can include one or more double bonds and can be aromatic (e.g., aryl groups). Examples of carbocycles include cyclopropyl, cyclopentyl, cyclohexyl, phenyl, naphthyl, and 1,2-dihydronaphthyl groups. Carbocycles can be optionally substituted.

[0610] As used herein, the term “heterocycle” or “heterocyclic group” means a mono- or multi-cyclic system including one or more rings, where at least one ring includes at least one heteroatom. Heteroatoms can be, for example, nitrogen, oxygen, or sulfur atoms. Rings can be three, four, five, six, seven, eight, nine, ten, eleven, or twelve membered rings. Heterocycles can include one or more double bonds and can be aromatic (e.g., heteroaryl groups). Examples of heterocycles include imidazolyl, imidazolidinyl, oxazolyl, oxazolidinyl, thiazolyl, thiazolidinyl, pyrazolidinyl, pyrazolyl, isoxazolidinyl, isoxazolyl, isothiazolidinyl, isothiazolyl, morpholinyl, pyrrolyl, pyrrolidinyl, furyl, tetrahydrofuryl, thiophenyl, pyridinyl, piperidinyl, quinolyl, and isoquinolyl groups. Heterocycles can be optionally substituted.

[0611] As used herein, a “biodegradable group” is a group that can facilitate faster metabolism of a lipid in a subject. A biodegradable group can be, but is not limited to, -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, an aryl group, and a heteroaryl group.

[0612] As used herein, an “aryl group” is a carbocyclic group including one or more aromatic rings. Examples of aryl groups include phenyl and naphthyl groups.

[0613] As used herein, a “heteroaryl group” is a heterocyclic group including one or more aromatic rings. Examples of heteroaryl groups include pyrrolyl, furyl, thiophenyl, imidazolyl, oxazolyl, and thiazolyl. Both aryl and heteroaryl groups can be optionally substituted. For example, M and M' can be selected from the non-limiting group consisting of optionally

substituted phenyl, oxazole, and thiazole. In the formulas herein, M and M' can be independently selected from the list of biodegradable groups above.

[0614] Alkyl, alkenyl, and cyclyl (e.g., carbocyclyl and heterocyclyl) groups can be optionally substituted unless otherwise specified. Optional substituents can be selected from the group consisting of, but are not limited to, a halogen atom (e.g., a chloride, bromide, fluoride, or iodide group), a carboxylic acid (e.g., $-\text{C}(\text{O})\text{OH}$), an alcohol (e.g., a hydroxyl, $-\text{OH}$), an ester (e.g., $-\text{C}(\text{O})\text{OR}$ or $-\text{OC}(\text{O})\text{R}$), an aldehyde (e.g., $-\text{C}(\text{O})\text{H}$), a carbonyl (e.g., $-\text{C}(\text{O})\text{R}$, alternatively represented by $\text{C}=\text{O}$), an acyl halide (e.g., $-\text{C}(\text{O})\text{X}$, in which X is a halide selected from bromide, fluoride, chloride, and iodide), a carbonate (e.g., $-\text{OC}(\text{O})\text{OR}$), an alkoxy (e.g., $-\text{OR}$), an acetal (e.g., $-\text{C}(\text{OR})_2\text{R}''$), in which each OR are alkoxy groups that can be the same or different and R'' is an alkyl or alkenyl group), a phosphate (e.g., $\text{P}(\text{O})_4^{3-}$), a thiol (e.g., $-\text{SH}$), a sulfoxide (e.g., $-\text{S}(\text{O})\text{R}$), a sulfinic acid (e.g., $-\text{S}(\text{O})\text{OH}$), a sulfonic acid (e.g., $-\text{S}(\text{O})_2\text{OH}$), a thial (e.g., $-\text{C}(\text{S})\text{H}$), a sulfate (e.g., $\text{S}(\text{O})_4^{2-}$), a sulfonyl (e.g., $-\text{S}(\text{O})_2-$), an amide (e.g., $-\text{C}(\text{O})\text{NR}_2$, or $-\text{N}(\text{R})\text{C}(\text{O})\text{R}$), an azido (e.g., $-\text{N}_3$), a nitro (e.g., $-\text{NO}_2$), a cyano (e.g., $-\text{CN}$), an isocyano (e.g., $-\text{NC}$), an acyloxy (e.g., $-\text{OC}(\text{O})\text{R}$), an amino (e.g., $-\text{NR}_2$, $-\text{NRH}$, or $-\text{NH}_2$), a carbamoyl (e.g., $-\text{OC}(\text{O})\text{NR}_2$, $-\text{OC}(\text{O})\text{NRH}$, or $-\text{OC}(\text{O})\text{NH}_2$), a sulfonamide (e.g., $-\text{S}(\text{O})_2\text{NR}_2$, $-\text{S}(\text{O})_2\text{NRH}$, $-\text{S}(\text{O})_2\text{NH}_2$, $-\text{N}(\text{R})\text{S}(\text{O})_2\text{R}$, $-\text{N}(\text{H})\text{S}(\text{O})_2\text{R}$, $-\text{N}(\text{R})\text{S}(\text{O})_2\text{H}$, or $-\text{N}(\text{H})\text{S}(\text{O})_2\text{H}$), an alkyl group, an alkenyl group, and a cyclyl (e.g., carbocyclyl or heterocyclyl) group.

[0615] In any of the preceding, R is an alkyl or alkenyl group, as defined herein. In some embodiments, the substituent groups themselves can be further substituted with, for example, one, two, three, four, five, or six substituents as defined herein. For example, a C_{1-6} alkyl group can be further substituted with one, two, three, four, five, or six substituents as described herein.

[0616] The compounds of any one of formulae (I), (IA), (II), (IIa), (IIb), (IIc), (IId), and (Ile) include one or more of the following features when applicable.

[0617] In some embodiments, R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(\text{CH}_2)_n\text{Q}$, $-(\text{CH}_2)_n\text{CHQR}$, $-\text{CHQR}$, and $-\text{CQ}(\text{R})_2$, where Q is selected from a C_{3-6} carbocycle, 5- to 14- membered aromatic or non-aromatic heterocycle having one or more heteroatoms selected from N, O, S, and P, $-\text{OR}$, $-\text{O}(\text{CH}_2)_n\text{N}(\text{R})_2$, $-\text{C}(\text{O})\text{OR}$, $-\text{OC}(\text{O})\text{R}$, $-\text{CX}_3$, $-\text{CX}_2\text{H}$, $-\text{CXH}_2$, $-\text{CN}$, $-\text{N}(\text{R})_2$, $-\text{C}(\text{O})\text{N}(\text{R})_2$, $-\text{N}(\text{R})\text{C}(\text{O})\text{R}$, $-\text{N}(\text{R})\text{S}(\text{O})_2\text{R}$, $-\text{N}(\text{R})\text{C}(\text{O})\text{N}(\text{R})_2$, $-\text{N}(\text{R})\text{C}(\text{S})\text{N}(\text{R})_2$, and $-\text{C}(\text{R})\text{N}(\text{R})_2\text{C}(\text{O})\text{OR}$, and each n is independently selected from 1, 2, 3, 4, and 5.

[0618] In another embodiment, R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(\text{CH}_2)_n\text{Q}$, $-(\text{CH}_2)_n\text{CHQR}$, $-\text{CHQR}$, and $-\text{CQ}(\text{R})_2$, where Q is selected from a C_{3-6} carbocycle, a 5-

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 3
CONTENANT LES PAGES 1 À 163

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 3
CONTAINING PAGES 1 TO 163

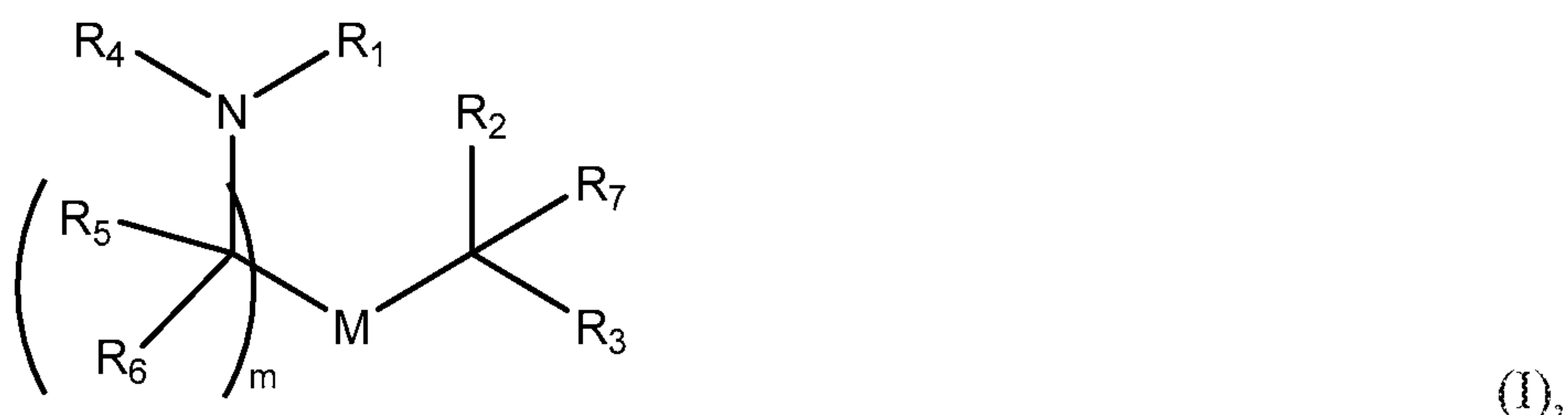
NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

CLAIMS

1. A composition comprising a polynucleotide that comprises an open reading frame (ORF) encoding an MCM polypeptide and a delivery agent, wherein the delivery agent comprises a compound having the formula (I)



or a salt or stereoisomer thereof, wherein

R_1 is selected from the group consisting of C_{5-20} alkyl, C_{5-20} alkenyl, $-R^*YR''$, $-YR''$, and $-R''M'R'$;

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, $-R^*YR''$, $-YR''$, and $-R^*OR''$, or R_2 and R_3 , together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a carbocycle, heterocycle,

$-OR$, $-O(CH_2)_nN(R)_2$, $-C(O)OR$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-N(R)_2$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, and $-C(R)N(R)_2C(O)OR$, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₄ alkyl and C₃₋₁₄ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13; and

provided when R₄ is -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, or -CQ(R)₂, then (i) Q is not -N(R)₂ when n is 1, 2, 3, 4 or 5, or (ii) Q is not 5, 6, or 7-membered heterocycloalkyl when n is 1 or 2.

2. An isolated polynucleotide comprising an ORF that has

(i) at least 99% or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 734,

(ii) at least 98%, at least 99%, or 100% sequence identity to nucleotide 97 to nucleotide 2250 of SEQ ID NO: 732,

(iii) at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 182, 733, and 741;

(iv) at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 735, 736, 738, 743, 744, 748, 749, 750, 754, 755, 758, 762, and 765;

(v) at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence

selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 180, 187, 737, 739, 740, 742, 745, 746, 747, 751, 752, 753, 757, 759, 760, 761, 763, and 764;

(vi) at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 181 and 756;

(vii) at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175;

(viii) at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 164, 166, 167, 168, 169, 170, 172, 177, 178, 179, 195, and 204;

(ix) at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 156, 157, 158, 159, 160, 161, 162, 174 and 176;

(x) at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 155 and 203;

(xi) at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a

sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 64, 66, 71, 91, and 128;

(xii) at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 9, 11, 18, 19, 21, 22, 23, 24, 32, 33, 37, 39, 40, 44, 45, 47, 50, 51, 52, 55, 57, 61, 65, 70, 79, 84, 86, 88, 90, 92, 98, 100, 115, 117, 126, 129, 135, 136, 137, 144, 148, 150, 184, 190, 191, and 206;

(xiii) at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 3, 5, 6, 8, 10, 12, 14, 16, 17, 20, 27, 28, 29, 31, 34, 35, 36, 38, 41, 42, 43, 46, 48, 49, 53, 54, 56, 58, 60, 63, 68, 69, 74, 77, 78, 80, 83, 85, 87, 93, 95, 96, 97, 99, 102, 103, 104, 105, 107, 110, 112, 113, 114, 116, 119, 120, 122, 123, 124, 125, 127, 131, 132, 133, 134, 138, 139, 140, 141, 142, 143, 147, 149, 183, 186, 188, 189, 192, 193, 194, 196, 197, 198, 199, 200, 201, 202, 205, and 207; or

(xiv) at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 97 to nucleotides 2250 of SEQ ID NOs: 1, 2, 4, 7, 13, 15, 25, 26, 30, 59, 62, 67, 72, 73, 75, 76, 81, 82, 89, 94, 101, 106, 108, 109, 111, 118, 121, 130, 145, 146, and 185.

3. The composition of claim 1 or the polynucleotide of claim 2, wherein the ORF further comprises a nucleic acid sequence encoding a mitochondrial transit peptide.

4. The composition of any one of claims 1 to 3 or the polynucleotide of claim 2 or 3, wherein the ORF has

(i) at least 99% or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 734;

(ii) at least 98%, at least 99%, or 100% sequence identity to nucleotide 1 to nucleotide 2250 of SEQ ID NO: 732;

(iii) at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 182 and 733;

(iv) at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 735, 741, 743, 744, 748, 758, 762, and 765;

(v) at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 180, 181, 736, 738, 739, 740, 742, 746, 747, 749, 750, 751, 752, 753, 754, 755, 757, 759, 760, 761, and 763;

(vi) at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 745, 756, and 764;

(vii) at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 154, 165, 171, 173, and 175;

(viii) at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to

nucleotides 2250 of SEQ ID NO: 151, 152, 153, 163, 166, 167, 168, 169, 170, 172, 177, 178, 179, 187, and 204;

(ix) at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NO: 156, 157, 158, 159, 160, 162, 164, 174, 176, 195, and 737;

(x) at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 155, 161, and 203;

(xi) at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 71 and 128;

(xii) at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 4, 6, 8, 9, 11, 19, 22, 23, 24, 32, 33, 37, 40, 44, 45, 47, 51, 61, 64, 65, 66, 79, 84, 86, 90, 91, 92, 100, 101, 112, 115, 117, 126, 129, 135, 136, 146, 148, 184, 190, and 191;

(xiii) at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 2, 3, 5, 7, 10, 12, 13, 14, 15, 16, 18, 20, 21, 26, 27, 28, 29, 31, 34, 36, 38, 39, 41, 42, 43, 46, 48, 49, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 68, 69, 70, 72, 73, 74, 76, 77, 80, 83, 85, 88, 95, 96, 97, 98, 102, 104, 105, 106, 107, 108, 109, 110, 113, 114, 120, 121,

122, 123, 124, 127, 131, 132, 133, 134, 137, 138, 139, 140, 141, 142, 144, 145, 147, 149, 150, 186, 188, 189, 192, 193, 194, 196, 198, 199, 200, 202, 205, 206, and 207; or

(xiv) at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence selected from the group consisting of nucleotides 1 to nucleotides 2250 of SEQ ID NOs: 1, 17, 25, 30, 35, 50, 63, 67, 75, 78, 81, 82, 87, 89, 93, 94, 99, 103, 111, 116, 118, 119, 125, 130, 143, 183, 185, 197, and 201.

5. The composition of any one of claims 1-4 or the polynucleotide of any one of claims 2-4, wherein the polynucleotide comprises at least one chemically modified nucleobase, sugar, backbone, or any combination thereof.

6. The composition of claim 5 or the polynucleotide of claim 5, wherein the at least one chemically modified nucleobase is selected from the group consisting of pseudouracil (ψ), N1-methylpseudouracil (m1 ψ), 2-thiouracil (s2U), 4'-thiouracil, 5-methylcytosine, 5-methyluracil, and any combination thereof.

7. The composition of any one of claims 1-6 or the polynucleotide of any one of claims 2-6, wherein the polynucleotide is mRNA.

8. The composition of any one of claims 1-7 or the polynucleotide of any one of claims 2-7 for use in treating methylmalonic acidemia in a subject in need thereof.

9. A polynucleotide comprising an mRNA encoding an MCM polypeptide for use in treating methylmalonic acidemia in a subject in need thereof.

10. The composition for use of claim 8 or the polynucleotide for use of claim 8 or 9, wherein the polynucleotide increases cellular expression of MCM by at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%.

11. The composition for use of claim 8 or 10 or the polynucleotide for use of any one of claims 8 to 10, wherein the MCM polypeptide expressed from the polynucleotide reduces the level of a metabolite associated with methylmalonic acidemia in a subject in need thereof.

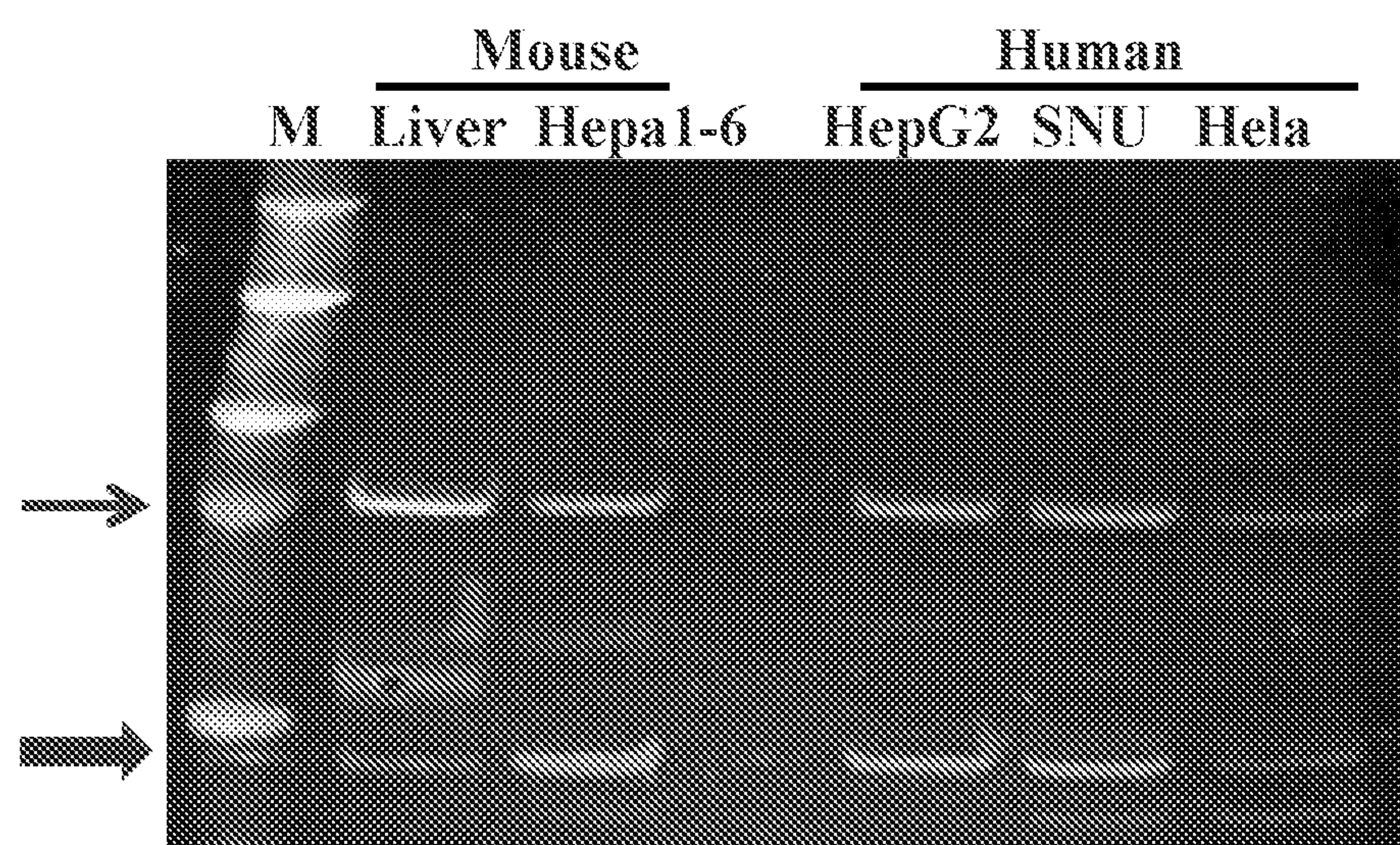
12. The composition for use of claim 11 or the polynucleotide for use of claim 11, wherein the MCM polypeptide expressed from the polynucleotide reduces the level of methylmalonic acid present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

13. The composition for use of claim 11 or the polynucleotide for use of claim 11, wherein the MCM polypeptide expressed from the polynucleotide reduces the level of propionyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

14. The composition for use of claim 11 or the polynucleotide for use of claim 11, wherein the MCM polypeptide expressed from the polynucleotide reduces the level of acetyl-carnitine present in the subject by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or about 100%.

15. The composition for use of any one of claims 11 to 14 or the polynucleotide for use of any one of claims 11 to 14, wherein the level of one or more metabolites associated with methylmalonic acidemia are reduced within one day, within two days, within three days, within four days, within five days, within seven days, within one week, within two weeks, within three weeks, or within one month of the administration of the polynucleotide.

FIG. 1



Mouse α -Methylmalonyl-CoA mutase ($\longrightarrow$)
Rabbit α -Citrate synthetase ($\longrightarrow$)

FIG. 2A

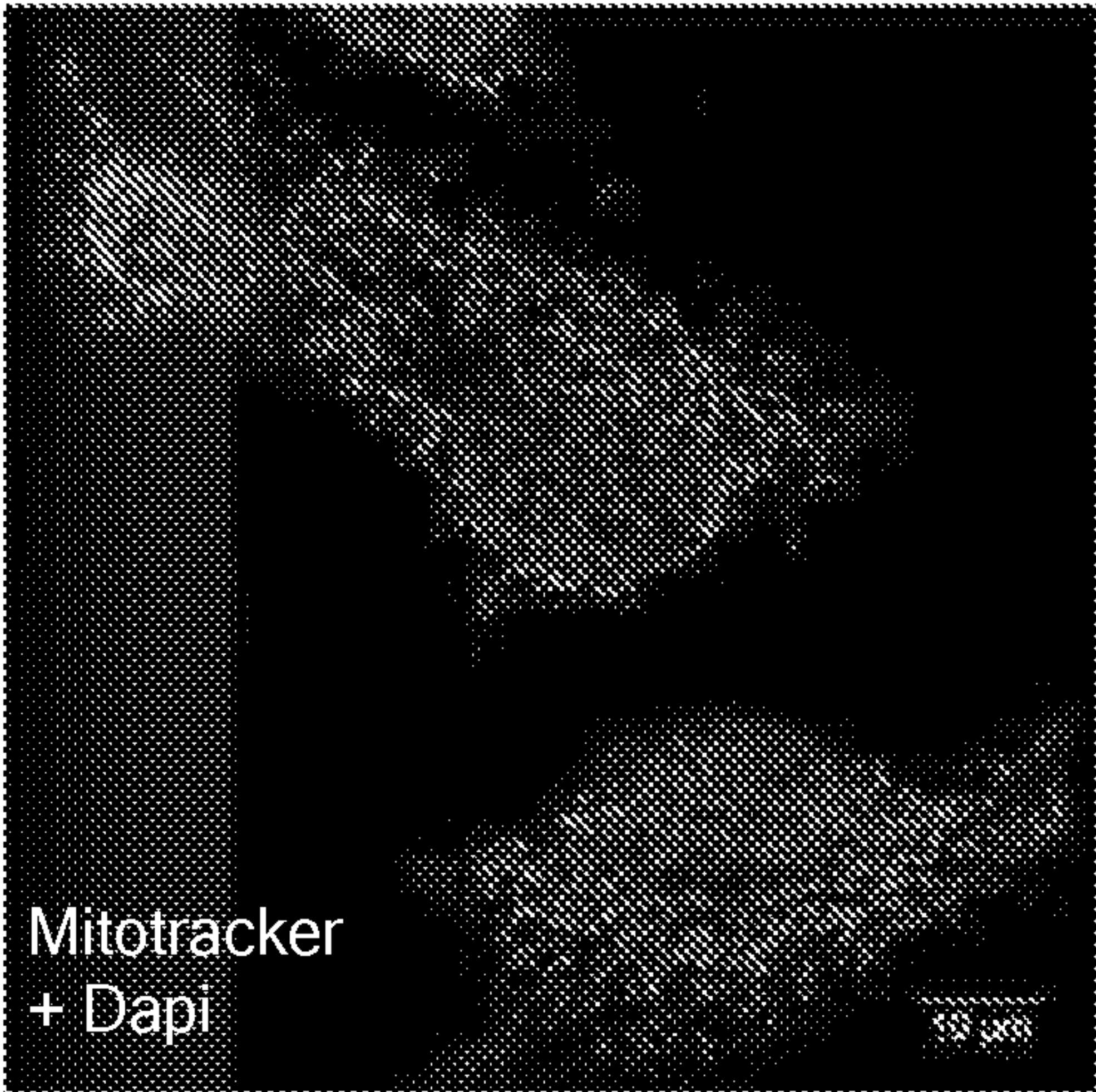


FIG. 2B

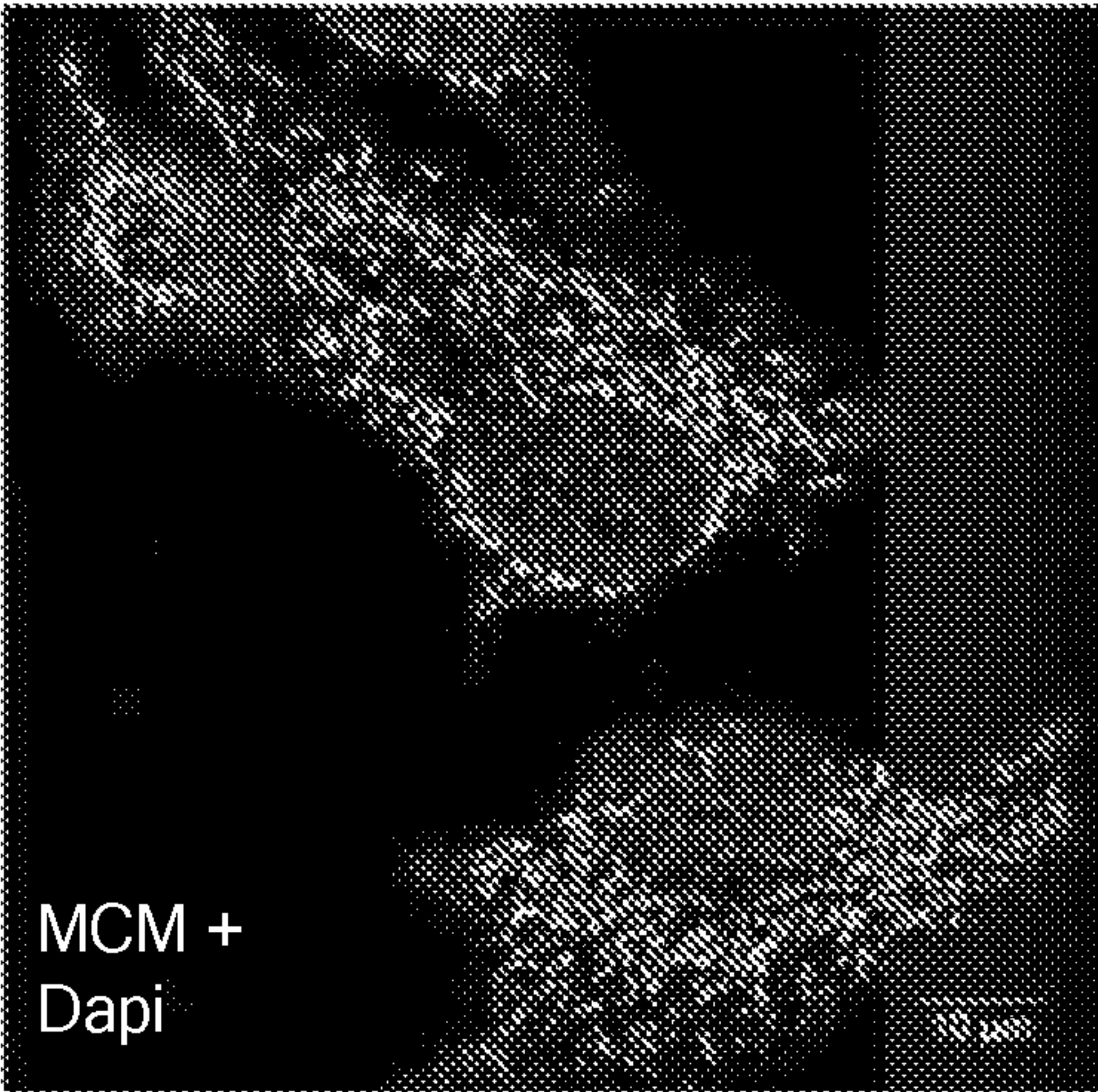
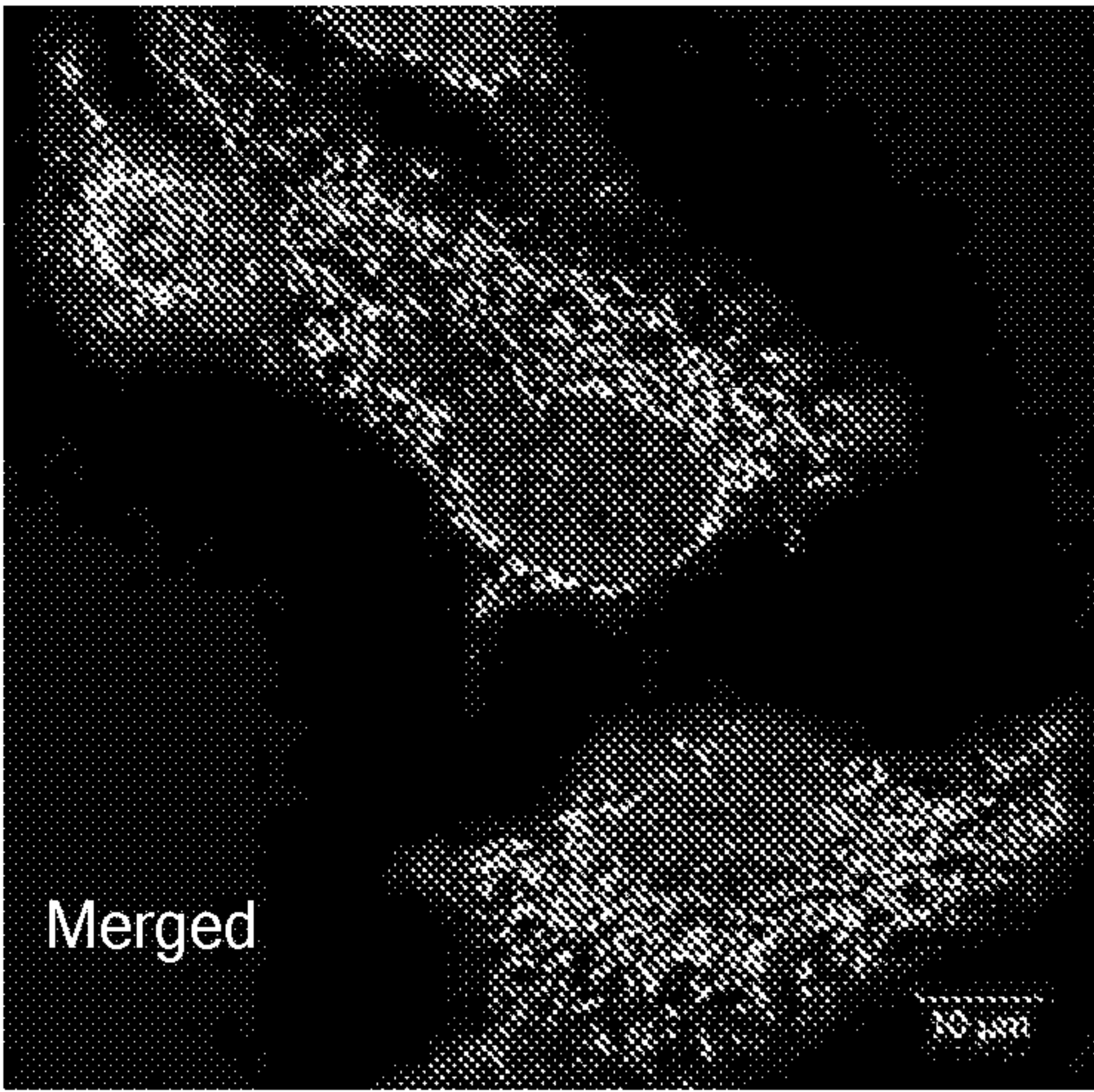


FIG. 2C



<u>KEY</u>	
hMCM:	green
Mitotracker:	red
DAPI:	blue

FIG. 3

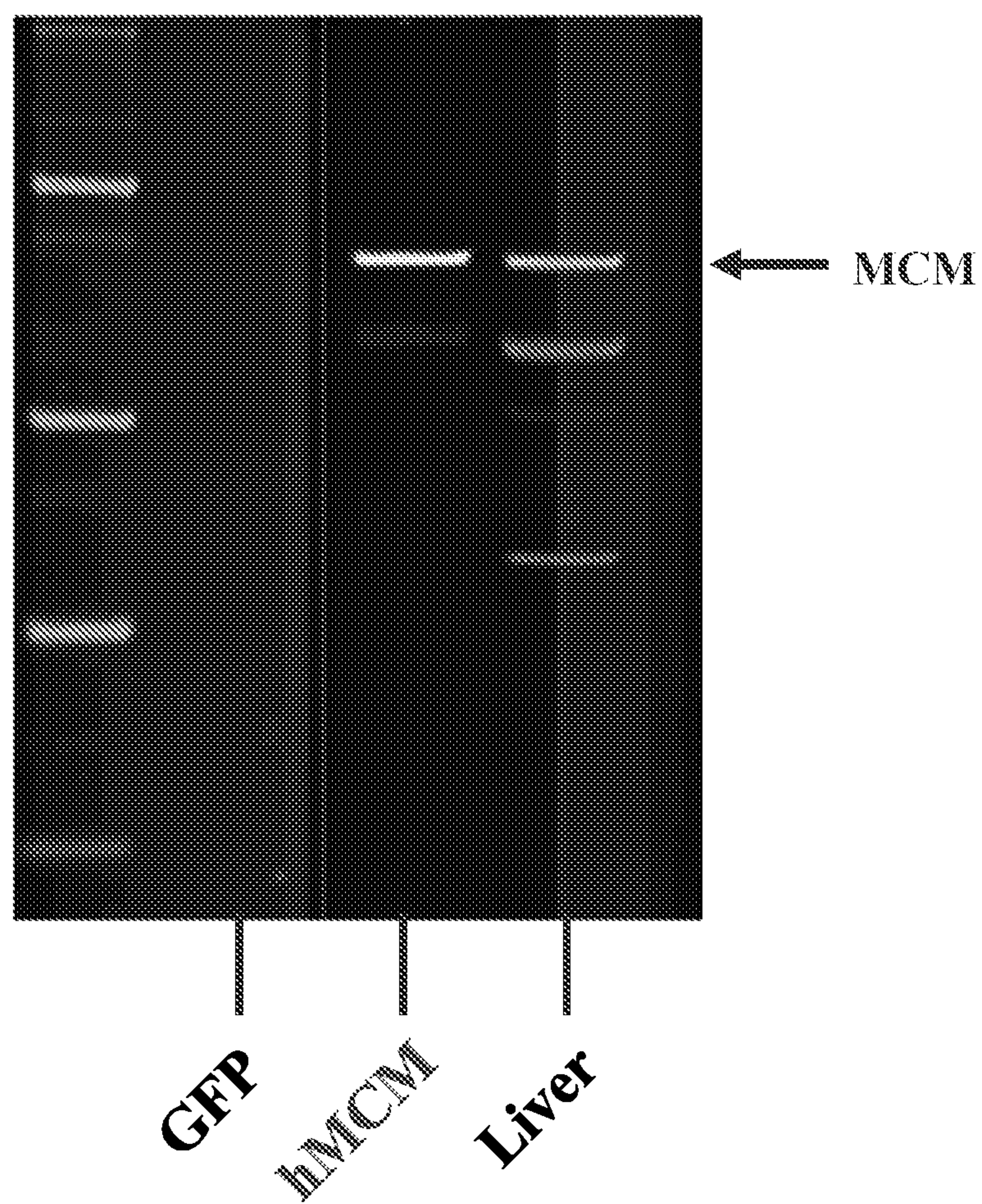


FIG. 4

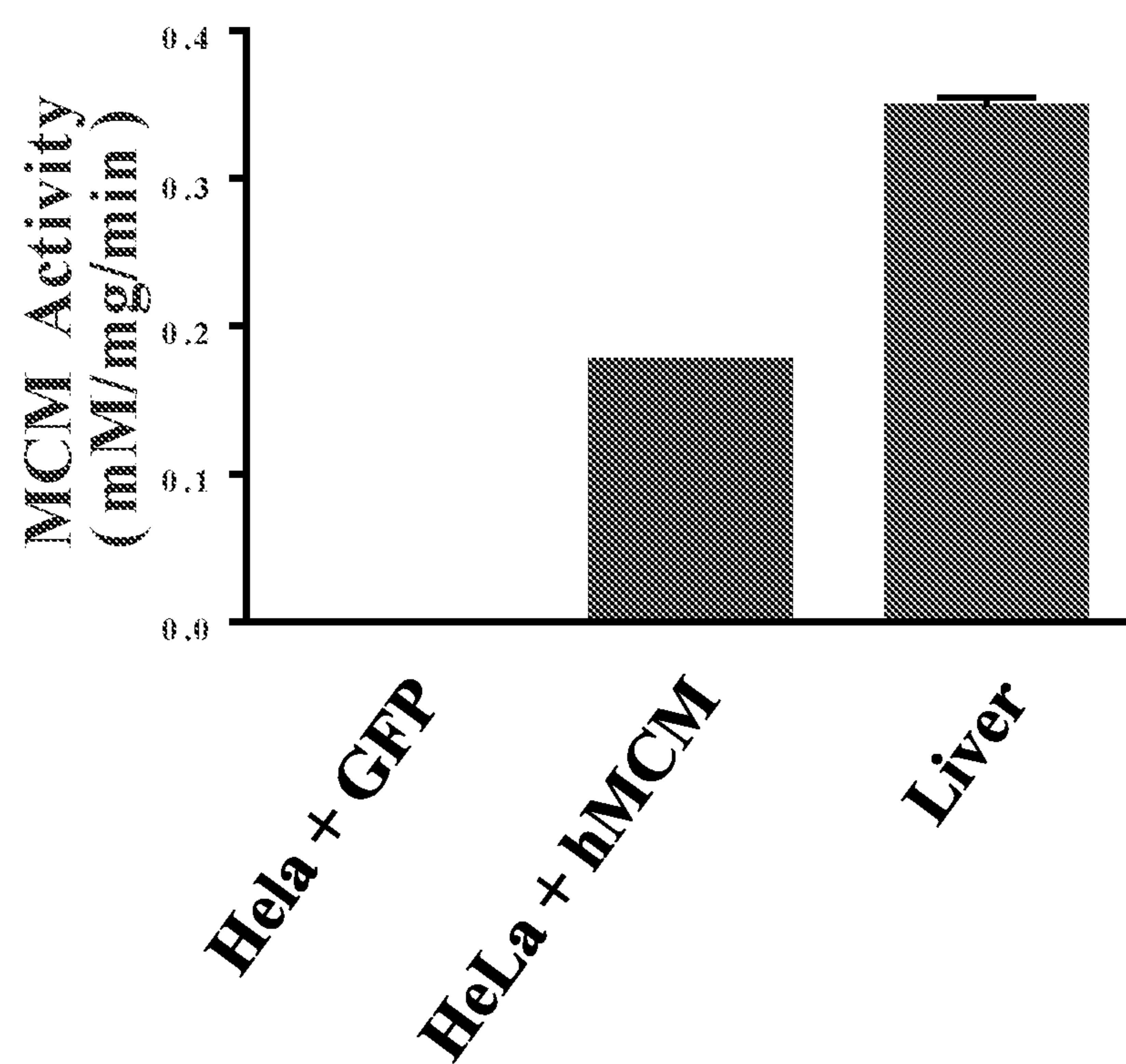
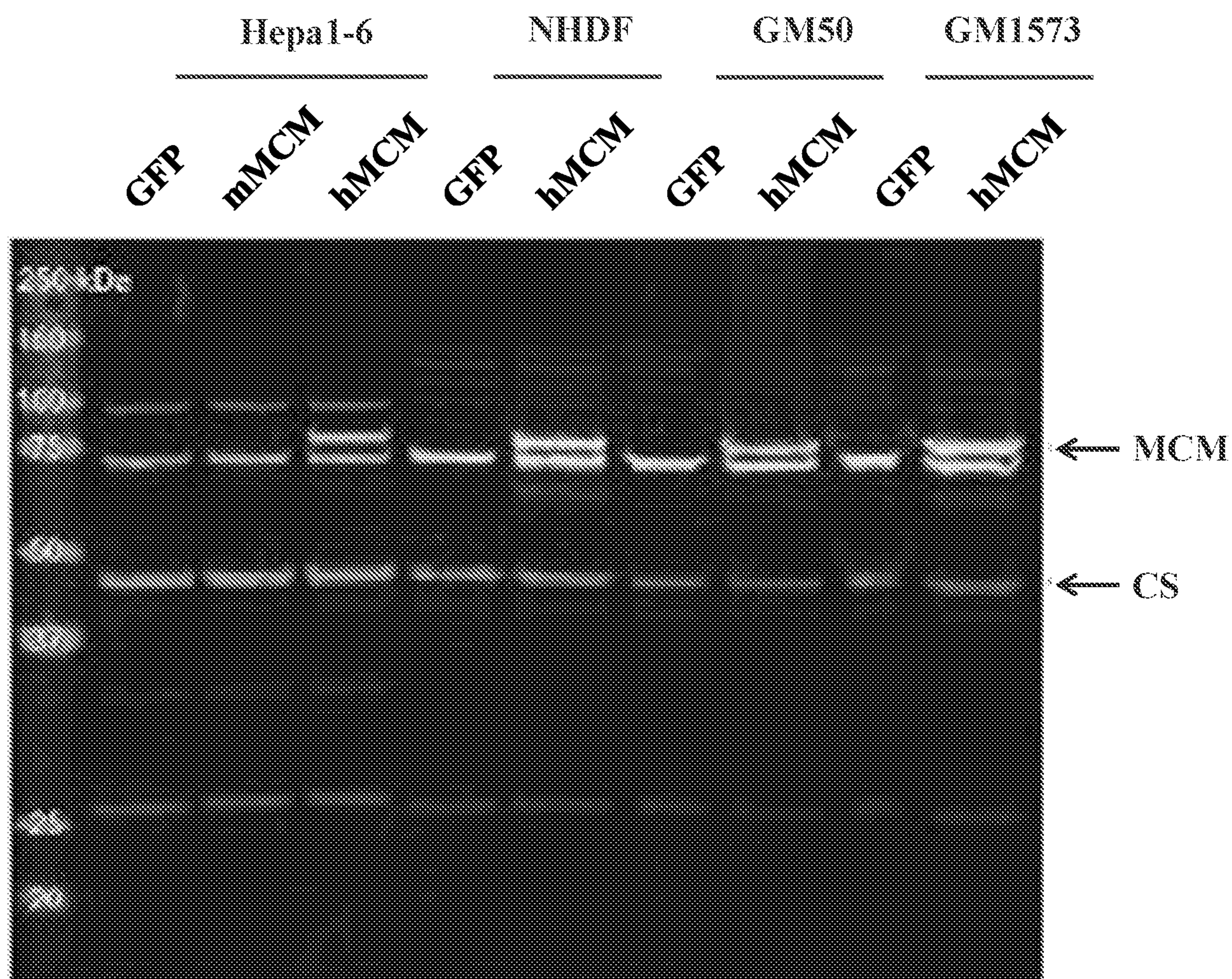


FIG. 5



MCM: Mouse α -Methylmalonyl-CoA mutase (TA505873_2C8)

CS: Rabbit α -Citrate synthetase

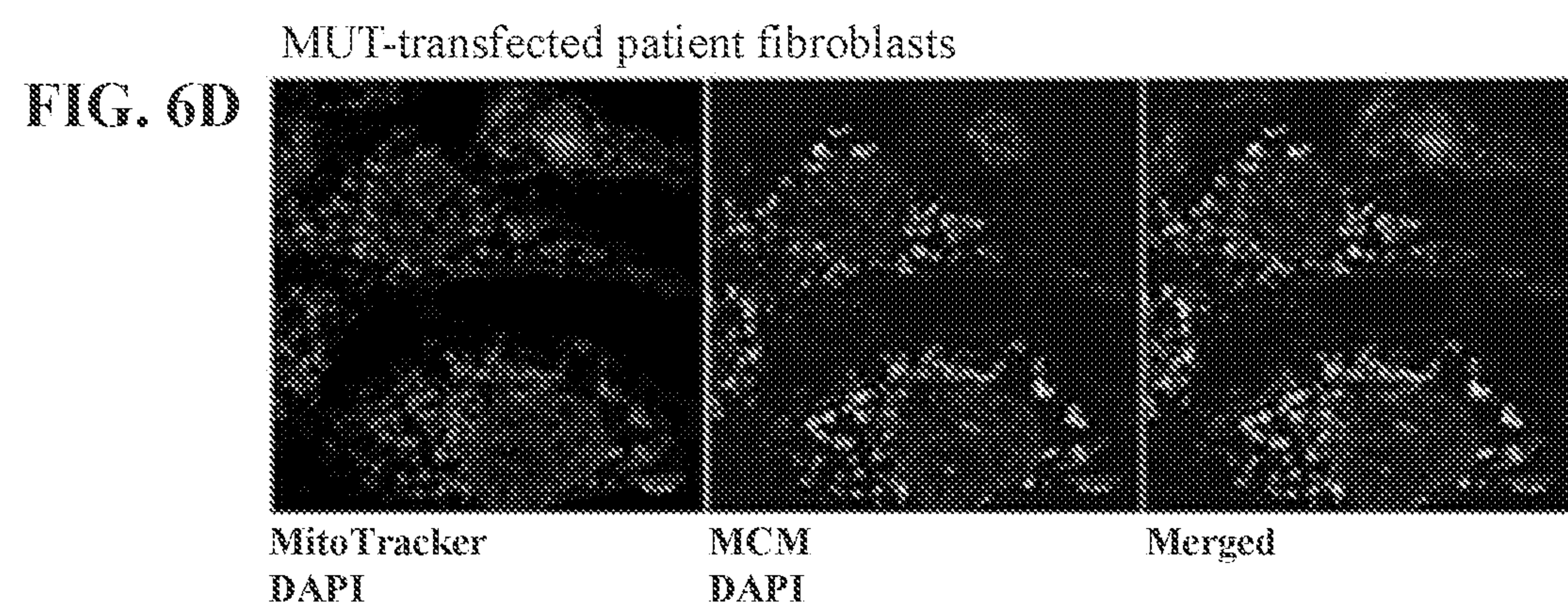
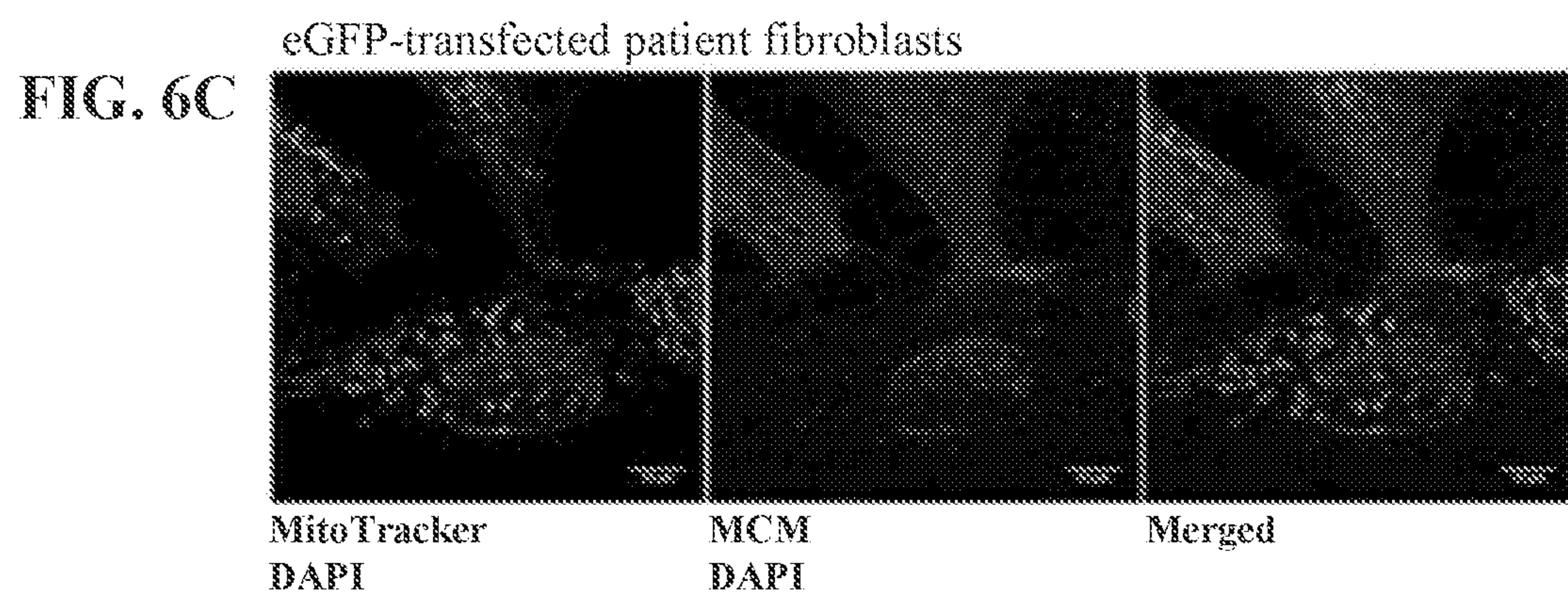
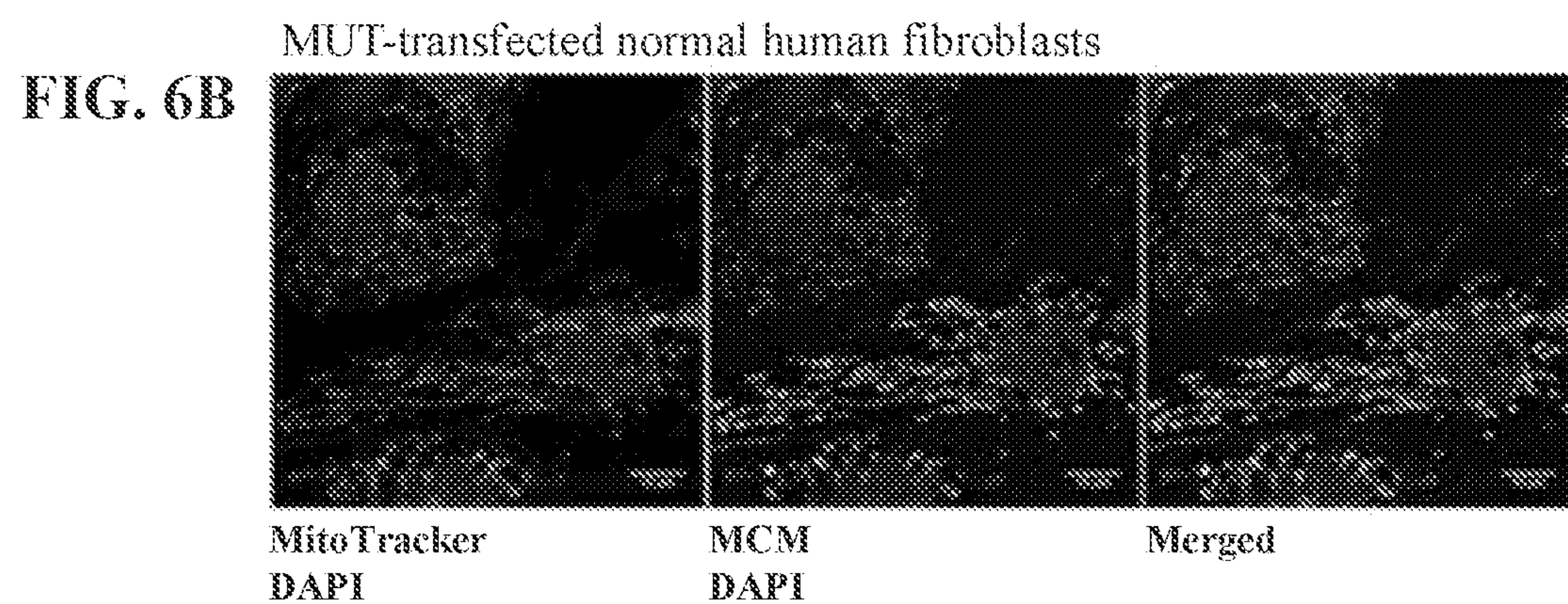
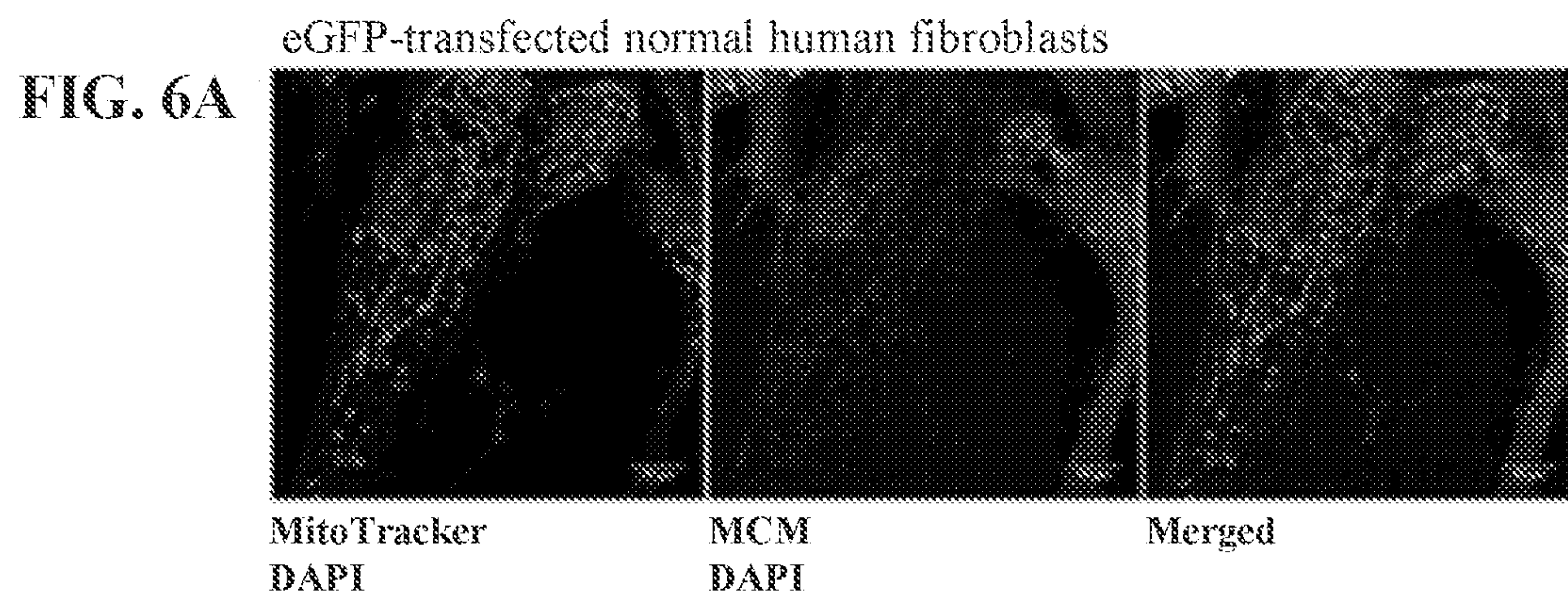


FIG. 7

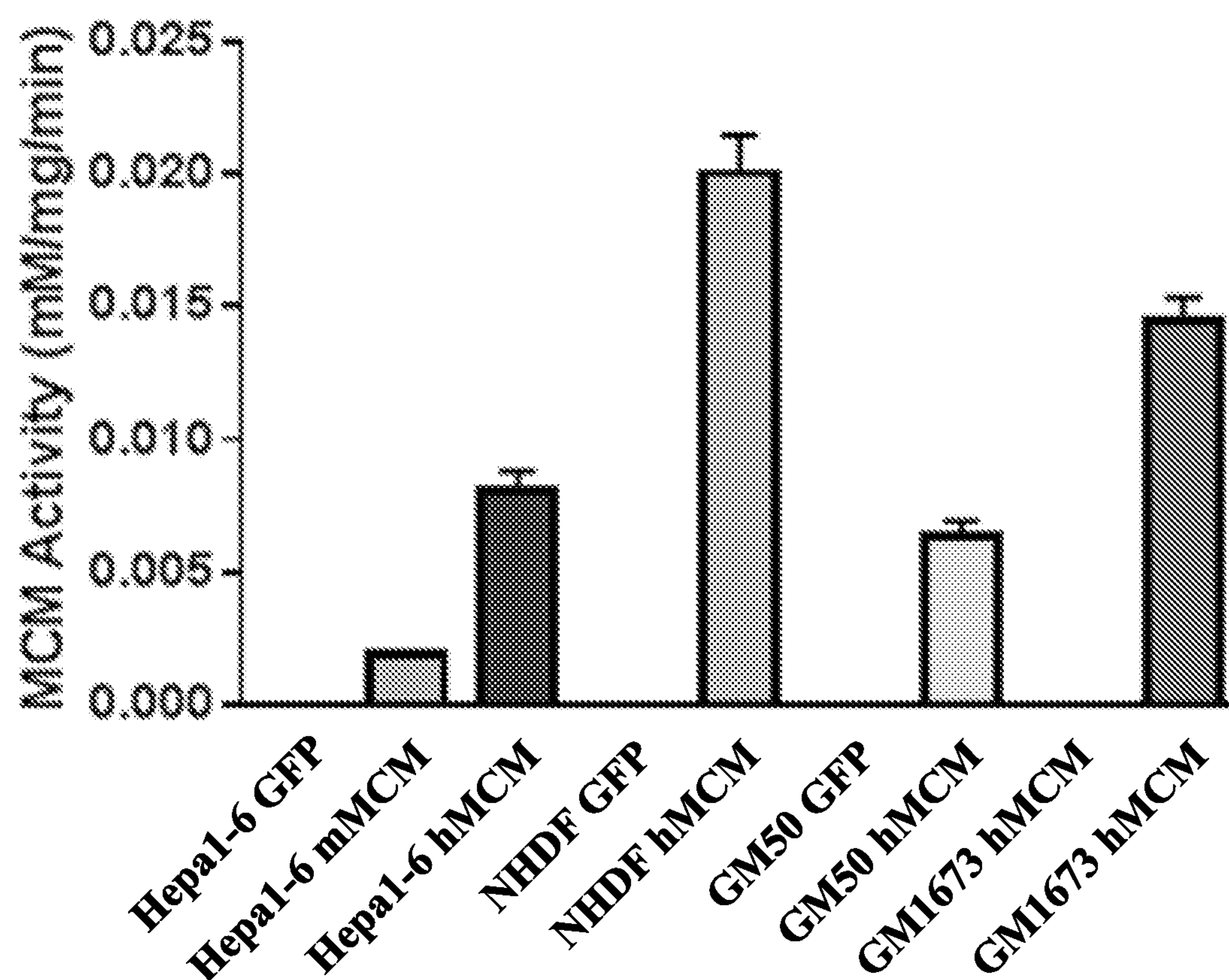


FIG. 8A

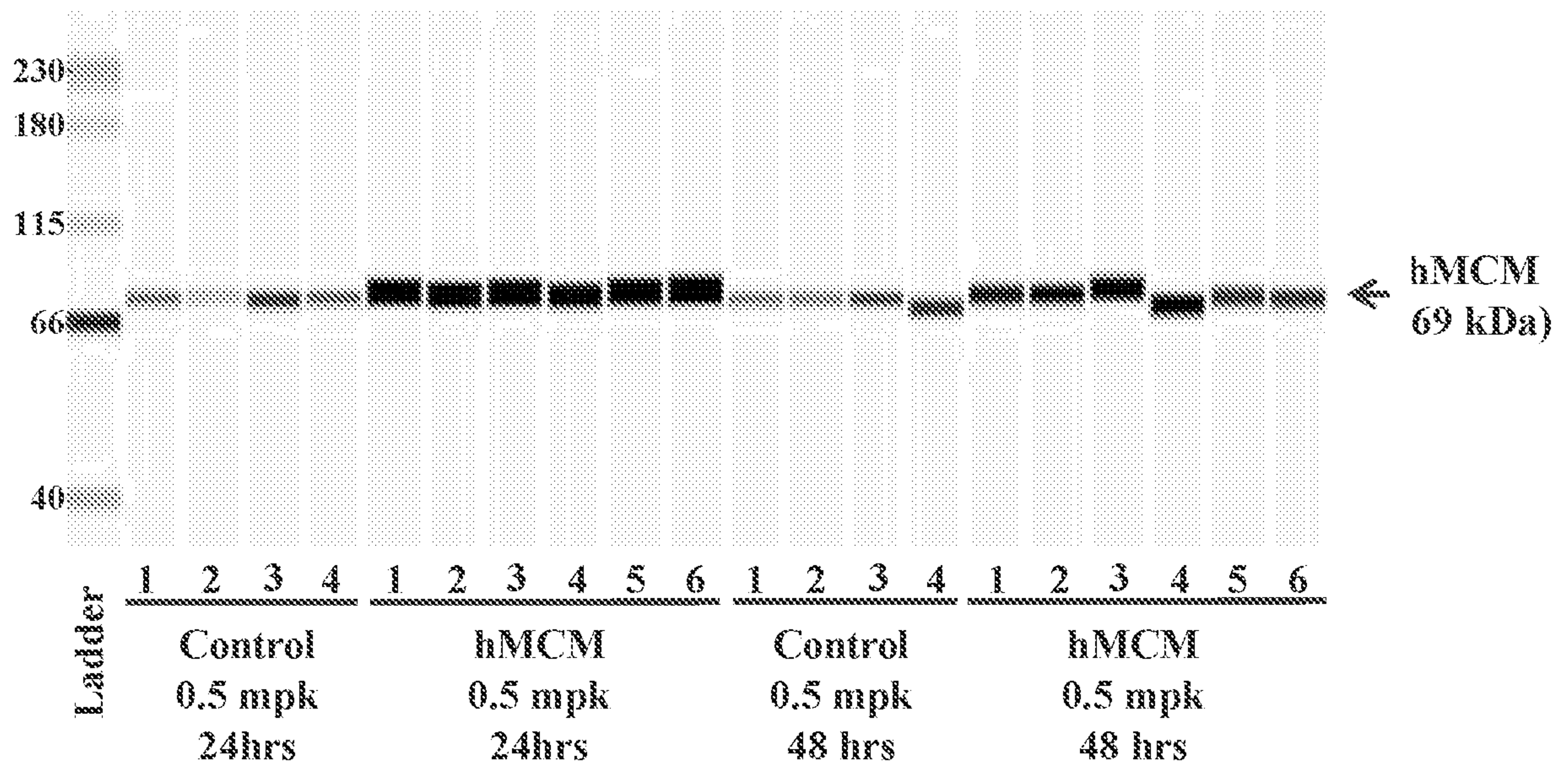


FIG. 8B

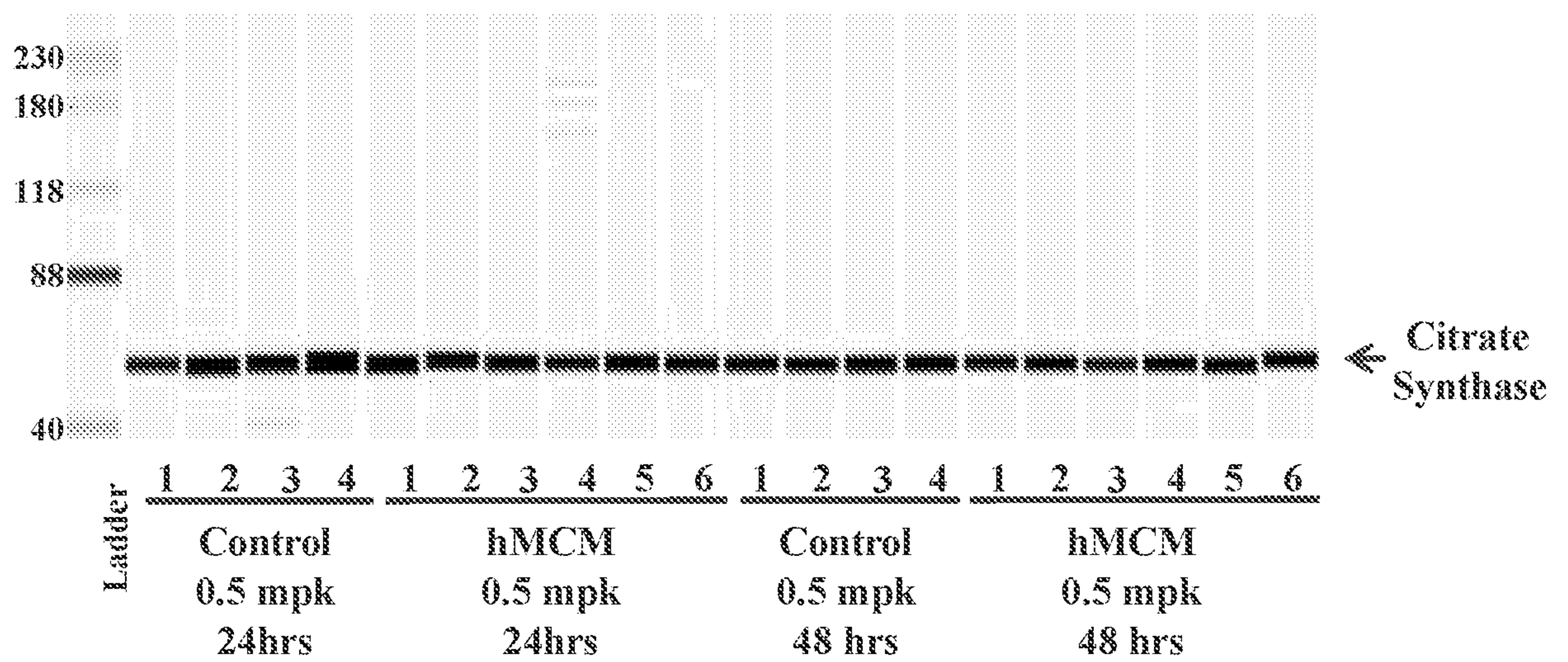


FIG. 9

ATGCTGCGGGCCAAGAACCAGCTGTTCTGCTGAGCCCTCACTACCTGCGGCAGGTGAA
GGAGAGCAGCGGCAGCCGGCTGATCCAGCAGCGGCTGCTGCACCAGCAGCAGCCCCTGC
ACCCCGAGTGGGCCGCCCTGGCCAAGAAGCAGCTGAAGGGCAAGAACCCCGAGGACCTG
ATCTGGCACACGCCCGAGGGCATCAGCATCAAGCCCCTGTACAGCAAGCGGGACACCAT
GGACCTGCCCCGAGGAGCTGCCCCGGCGTGAAGCCCTTACCCGGGGCCCCCTACCCACCA
TGTACACCTTCCGGCCCTGGACCATCCGGCAGTACGCCGGCTTCAGCACCGTGGAGGAG
AGCAACAAGTTCTACAAGGACAACATCAAGGCCGGCCAGCAGGGCCTGAGCGTGGCCTT
CGACCTGGCCACCCACCGGGGCTACGACAGCGACAACCCACGGGTGCGGGGCGACGTGG
GCATGGCCGGCGTGGCCATCGACACCGTGGAGGACACCAAGATCCTGTTGACGGGCATC
CCTCTGGAGAAGATGAGCGTGAGCATGACCATGAACGGCGCCGTGATCCCCGTGCTGGC
CAACTTCATCGTGACCGGCGAGGAGCAGGGCGTGCCCAAGGAGAAGCTGACCGGCACCA
TCCAGAACGACATCCTGAAGGAGTTCATGGTGCGGAACACCTACATCTTCCCTCCCGAG
CCCAGCATGAAGATCATCGCCGACATCTTCGAGTACACCGCCAAGCACATGCCCAAGTT
CAACAGCATCAGCATCAGCGGCTACCACATGCAGGAGGCCGGCGCCGACGCCATCCTGG
AGCTGGCCTACACCCTGGCCGACGGCCTGGAGTACAGCCGGACCGGCCTGCAGGCCGGC
CTGACCATCGACGAGTTCGCGCCCCGGCTGAGCTTCTTCTGGGGCATCGGCATGAACTT
CTACATGGAGATCGCCAAGATGCGGGCCGGCCGGCGGCTGTGGGGCCACCTGATCGAGA
AGATGTTCCAGCCCAAGAACAGCAAGAGCCTGCTGCTGCGGGCCCACTGCCAGACCAGC
GGCTGGAGCCTGACCGAGCAGGACCCCTACAACAACATCGTGCGGACCGCCATCGAGGC
CATGGCCGGCGTGTTCGGCGGCACCCAGAGCCTGCACACCAACAGCTTCGACGAGGCC
TGGGCCTGCCCACCGTGAAGAGCGCCCGGATCGCCCGGAACACCCAGATCATCATCCAG
GAGGAGAGCGGCATCCCCAAGGTGGCCGACCCCTGGGGCGGCAGCTACATGATGGAGTG
CCTGACCAACGACGTGTACGACGCCGCCCTGAAGCTGATCAACGAGATCGAGGAGATGG
GCGGCATGGCCAAGGCCGTGGCCGAGGGCATCCCCAAGCTGCGGATCGAGGAGTGCGCC
GCCCGGCGGCAGGCCCGGATCGACAGCGGCAGCGAGGTGATCGTGCGGTGAACAAGTA
CCAGCTGGAGAAGGAGGACGCCGTGGAGGTGCTGGCCATCGACAACACCAGCGTGCGGA
ACCGGCAGATCGAGAAGCTGAAGAAGATCAAGAGCAGCCGGGACCAGGCCCTGGCCGAG
CGGTGCCTGGCCGCCCTGACCGAGTGCGCCGCCAGCGGCGACGGCAACATCCTGGCCCT
GGCCGTGGACGCCAGCCGGGCCCGGTGCACCGTGGGCGAGATCACCGACGCCCTGAAGA
AGGTGTTCCGGCGAGCACAAGGCCAACGACCGGATGGTGAGCGGCGCCTACCGGCAGGAG
TTCGGCGAGAGCAAGGAGATCACCAGCGCCATCAAGCGGGTGCACAAGTTTCATGGAGCG
GGAGGGCCGGCGGGCCCCGGCTGCTGGTGGCCAAGATGGGCCAGGACGGCCACGACCGGG
GCGCCAAGGTGATCGCCACCGGCTTCGCCGACCTGGGCTTCGACGTGGACATCGGCCCA
CTGTTCCAGACGCCCCGGGAGGTGGCCCAGCAGGCCGTGGACGCCGACGTGCACGCCGT
GGGCGTGAGCACCTGGCCGCCGGCCACAAGACCCTGGTGCCCGAGCTGATCAAGGAGC
TGAACAGCCTGGGCCGGGCCGACATCCTGGTGATGTGCGGCGGCGTGATCCCGCCCCAG
GACTACGAGTTCCTGTTTCGAGGTGGGCGTGAGCAACGTGTTTCGGCCCCGGCACCCGGAT
CCCCAAGGCCGCCGTGCAGGTGCTGGACGACATCGAGAAGTGCCTGGAGAAGAAGCAGC
AGAGCGTG

SEQ ID NO: 732

10/21

FIG. 10

ATGCTGCGGGCCAAGAACCAGCTGTTCTGCTGAGCCCGCACTACCTGCGGCAGGTGAA
GGAGAGCAGCGGCAGCCGGCTGATCCAGCAGCGCCTCCTCCACCAGCAGCAGCCCTCC
ACCCCGAGTGGGCGCCCTCGCCAAGAAGCAGCTCAAGGGCAAGAACCCCGAGGACCTC
ATCTGGCACACGCCCCGAGGGCATCTCCATCAAGCCCCTCTACTCCAAGCGCGACACCAT
GGACCTCCCCGAGGAGCTCCCCGGCGTCAAGCCCTTCACCCGCGGGCCCCCTACCCACCA
TGTACACCTTCCGCCCCCTGGACCATCCGCCAGTACGCCGGCTTCTCCACCGTCGAGGAG
TCCAACAAGTTCTACAAGGACAACATCAAGGCCGGCCAGCAGGGCCTCTCCGTCGCCCTT
CGACCTCGCCACCCACCGCGGGCTACGACTCCGACAACCCGCGCGTCCGCGGGCGACGTCG
GCATGGCCGGCGTCGCCATCGACACCGTCGAGGACACCAAGATCCTCTTCGACGGGCATC
CCTCTCGAGAAGATGTCCGTCTCCATGACCATGAACGGCGCCGTCATCCCCGTCTCTCGC
CAACTTCATCGTCACCCGGCGAGGAGCAGGGCGTCCCCAAGGAGAAGCTCACCGGCACCA
TCCAGAACGACATCCTCAAGGAGTTTATGGTCCGCAACACCTACATCTTCCCTCCCGAG
CCCTCCATGAAGATCATCGCCGACATCTTCGAGTACACCGCCAAGCACATGCCCAAGTT
CAACTCCATCTCCATCTCCGGCTACCACATGCAGGAGGCCGGCGCCGACGCCATCTCTCG
AGCTCGCCTACACCCTCGCCGACGGCCTCGAGTACTCCCGCACCGGCCTCCAGGCCGGC
CTCACCATCGACGAGTTCCGCGCCCCGCTCTCCTTCTTCTGGGGCATCGGCATGAACTT
CTACATGGAGATCGCCAAGATGCGCGCCGGCCGCGCCTCTGGGGCCACCTCATCGAGA
AGATGTTCCAGCCCAAGAACTCCAAGTCCCTCCTCCTCCGCGCCCACTGCCAGACCTCC
GGCTGGTCCCTCACCGAGCAGGACCCCTACAACAACATCGTCCGCACCGCCATCGAGGC
CATGGCCGGCGTCTTCGGCGGCACCCAGTCCCTCCACACCAACTCCTTCGACGAGGCC
TCGGCCTCCCCACCGTCAAGTCCGCCCCGCATCGCCCGCAACACCCAGATCATCATCCAG
GAGGAGTCCGGCATCCCCAAGGTGCGCCGACCCCTGGGGCGGCTCCTACATGATGGAGTG
CCTCACCAACGACGTCTACGACGCCGCCCTCAAGCTCATCAACGAGATCGAGGAGATGG
GCGGCATGGCCAAGGCCGTGCGCCGAGGGCATCCCCAAGCTCCGCATCGAGGAGTGCGCC
GCCCGCCGCCAGGCCCGCATCGACTCCGGCTCCGAGGTTCATCGTCCGGCGTCAACAAGTA
CCAGCTCGAGAAGGAGGACGCCGTGAGGTCCTCGCCATCGACAACACCTCCGTCCGCA
ACCGCCAGATCGAGAAGCTCAAGAAGATCAAGTCTCCTCCCGCGACAGGCCCTCGCCGAG
CGCTGCCTCGCCGCCCTCACCGAGTGCGCCGCCCTCCGGCGACGGCAACATCCTCGCCCT
CGCCGTGACGCCTCCCGCGCCCGCTGCACCGTCCGGCGAGATCACCGACGCCCTCAAGA
AGGTCTTCGGCGAGCACAAGGCCAACGACCGCATGGTCTCCGGCGCCTACCGCCAGGAG
TTCGGCGAGTCCAAGGAGATCACCTCCGCCATCAAGCGCGTCCACAAGTTTCATGGAGCG
CGAGGGCCCGCCGCCCGCGCCTCCTCGTCCGCAAGATGGGCCAGGACGGCCACGACCGCG
GCGCCAAGGTCATCGCCACCGGCTTCGCCGACCTCGGCTTCGACGTGACATCGGCCCA
CTCTTCAGACGCCCCCGGAGGTCGCCCGAGCAGGCCGTGACGCGCGACGTCCACGCCGT
CGGCGTCTCCACCCTCGCCGCCCGGCCACAAGACCCCTCGTCCCCGAGCTCATCAAGGAGC
TCAACTCCCTCGGCCGCCCGGACATCCTCGTCATGTGCGGGCGGCGTCATCCCTCCCCAG
GACTACGAGTTCTCTTCGAGGTCGGCGTCTCCAACGTCTTCGGCCCCGGCACCCGCAT
CCCCAAGGCCGCCGTCCAGGTCCTCGACGACATCGAGAAGTGCCTCGAGAAGAAGCAGC
AGTCCGTC

SEQ ID NO: 733

FIG. 11

ATGCTGCGGGCCAAGAACCAGCTGTTCTGCTGAGCCCCACTACCTGCGGCAGGTGAA
GGAGAGCAGCGGCAGCCGGCTGATCCAGCAGCGGCTGCTGCACCAGCAGCAGCCCCTGC
ACCCCGAGTGGGCCGCCCTGGCCAAGAAGCAGCTGAAGGGCAAGAACCCCGAGGACCTG
ATCTGGCACACCCCCGAGGGCATCAGCATCAAGCCCCTGTACAGCAAGCGGGACACCAT
GGACCTGCCCCGAGGAGCTGCCCCGGCGTGAAGCCCTTACCCGGGGCCCCCTACCCACCA
TGTACACCTTCCGGCCCTGGACCATCCGGCAGTACGCCGGCTTCAGCACCGTGGAGGAG
AGCAACAAGTTCTACAAGGACAACATCAAGGCCGGCCAGCAGGGCCTGAGCGTGGCCTT
CGACCTGGCCACCCACCGGGGCTACGACAGCGACAACCCCCGGGTGCGGGGCGACGTGG
GCATGGCCGGCGTGGCCATCGACACCGTGGAGGACACCAAGATCCTGTTGACGGGCATC
CCCCTGGAGAAGATGAGCGTGAGCATGACCATGAACGGCGCCGTGATCCCCGTGCTGGC
CAACTTCATCGTGACCGGCGAGGAGCAGGGCGTGCCCAAGGAGAAGCTGACCGGCACCA
TCCAGAACGACATCCTGAAGGAGTTCATGGTGCGGAACACCTACATCTTCCCCCCCCGAG
CCCAGCATGAAGATCATCGCCGACATCTTCGAGTACACCGCCAAGCACATGCCCAAGTT
CAACAGCATCAGCATCAGCGGCTACCACATGCAGGAGGCCGGCGCCGACGCCATCCTGG
AGCTGGCCTACACCCTGGCCGACGGCCTGGAGTACAGCCGGACCGGCCTGCAGGCCGGC
CTGACCATCGACGAGTTCGCCCCCGGCTGAGCTTCTTCTGGGGCATCGGCATGAACTT
CTACATGGAGATCGCCAAGATGCGGGCCGGCCGGCGGCTGTGGGGCCACCTGATCGAGA
AGATGTTCCAGCCCAAGAACAGCAAGAGCCTGCTGCTGCGGGCCCCACTGCCAGACCAGC
GGCTGGAGCCTGACCGAGCAGGACCCCTACAACAACATCGTGCGGACCGCCATCGAGGC
CATGGCCGGCGTGTTCGGCGGCACCCAGAGCCTGCACACCAACAGCTTCGACGAGGCC
TGGGCCTGCCCACCGTGAAGAGCGCCCGGATCGCCCGGAACACCCAGATCATCATCCAG
GAGGAGAGCGGCATCCCCAAGGTGGCCGACCCCTGGGGCGGCAGCTACATGATGGAGTG
CCTGACCAACGACGTGTACGACGCCGCCCTGAAGCTGATCAACGAGATCGAGGAGATGG
GCGGCATGGCCAAGGCCGTGGCCGAGGGCATCCCCAAGCTGCGGATCGAGGAGTGCGCC
GCCCGGCGGCAGGCCCGGATCGACAGCGGCAGCGAGGTGATCGTGCGGTGAACAAGTA
CCAGCTGGAGAAGGAGGACGCCGTGGAGGTGCTGGCCATCGACAACACCAGCGTGCGGA
ACCGGCAGATCGAGAAGCTGAAGAAGATCAAGAGCAGCCGGGACCAGGCCCTGGCCGAG
CGGTGCCTGGCCGCCCTGACCGAGTGCGCCGCCAGCGGCGACGGCAACATCCTGGCCCT
GGCCGTGGACGCCAGCCGGGCCCCGGTGCACCGTGGGCGAGATCACCGACGCCCTGAAGA
AGGTGTTCCGGCGAGCACAAGGCCAACGACCGGATGGTGAGCGGCGCCTACCGGCAGGAG
TTCGGCGAGAGCAAGGAGATCACCAGCGCCATCAAGCGGGTGCACAAGTTTCATGGAGCG
GGAGGGCCGGCGGGCCCCGGCTGCTGGTGGCCAAGATGGGCCAGGACGGCCACGACCGGG
GCGCCAAGGTGATCGCCACCGGCTTCGCCGACCTGGGCTTCGACGTGGACATCGGCCCC
CTGTTCCAGACCCCCCGGGAGGTGGCCCAGCAGGCCGTGGACGCCGACGTGCACGCCGT
GGGCGTGAGCACCTGGCCGCCGGCCACAAGACCCTGGTGCCCGAGCTGATCAAGGAGC
TGAACAGCCTGGGCCCCGCCGACATCCTGGTGATGTGCGGCGGCGTGATCCCCCCCCAG
GACTACGAGTTCCTGTTTCGAGGTGGGCGTGAGCAACGTGTTTCGGCCCCGGCACCCGGAT
CCCCAAGGCCGCCGTGCAGGTGCTGGACGACATCGAGAAGTGCCTGGAGAAGAAGCAGC
AGAGCGTG

SEQ ID NO: 734

FIG. 12

ATGCTGAGAGCCAAGAACCAGCTGTTCTGCTGAGCCCTCACTACCTGAGACAGGTAAA
GGAGAGCAGCGGCTCAAGGCTGATCCAACAGAGGTTGCTGCACCAGCAGCAGCCCCTTC
ATCCCGAGTGGGCAGCTCTGGCCAAGAAACAGTTGAAGGGCAAGAACCCCGAGGACCTG
ATTTGGCATACCCCCGAGGGCATCAGCATCAAGCCTCTGTACTCCAAGCGCGACACCAT
GGACCTCCCCGAGGAACTGCCCCGGGGTGAAACCATTTACACGCGGGCCCCCTACCCACCA
TGTACACCTTCCGGCCATGGACGATTAGGCAGTACGCTGGCTTCAGCACGGTGGAGGAG
AGCAATAAGTTCTACAAGGACAACATCAAGGCAGGGCAGCAGGGGGCTTAGCGTGGCCTT
TGACCTGGCCACACACCGGGGGTACGACAGCGACAATCCCAGGGTGCGGGGCGACGTTG
GCATGGCCGGCGTTGCTATCGATACCGTGGAGGATACGAAGATACTGTTTGACGGGATC
CCCCTGGAGAAAATGAGCGTGAGCATGACAATGAACGGCGCTGTGATCCCTGTCCTCGC
TAACTTCATCGTTACCGGCGAGGAACAGGGAGTACCTAAGGAAAAGCTGACTGGAAC TA
TTCAGAATGACATTCTCAAGGAATTCATGGTCAGGAACACTTATATCTTTCCCCCTGAA
CCGTCTATGAAAATCATAGCAGACATTTTTTGAGTATACCGCCAAGCACATGCCCAAGTT
TAACAGCATCTCCATCTCTGGCTACCACATGCAGGAGGCAGGCGCCGACGCTATTCTGG
AGCTGGCGTACACTCTCGCCGACGGACTGGAGTACTCTCGCACTGGACTCCAGGCCGGG
CTGACCATCGATGAGTTCGCACCGAGGCTGTCTTCTTCTGGGGCATTGGAATGAACTT
TTACATGGAGATCGCTAAGATGCGAGCAGGCCGGAGACTGTGGGGCCACCTGATTGAAA
AAATGTTCCAACCTAAGAATTCGAAGAGCCTGCTGTTGAGAGCCCCTGACCAGACCAGC
GGCTGGTCGCTGACCGAGCAGGACCCCTACAATAACATAGTGCGGACCGCTATTGAAGC
CATGGCCGCGCTTTTTGGCGGCACTCAGAGCCTGCACACCAACAGCTTTGACGAGGCC
TGGGCCTGCCCCTGTGAAGAGTGCCAGAATCGCCAGGAATACCCAGATTATCATTCAG
GAGGAGTCAGGAATACCCAAAGTGGCCGACCCCTGGGGGGGTAGCTACATGATGGAGTG
CCTGACAAACGACGTGTACGATGCCGCCCTCAAGCTGATCAACGAGATCGAGGAAATGG
GCGGCATGGCCAAGGCTGTTGCCGAAGGCATCCCCAAGCTCCGGATTGAAGAGTGTGCG
GCCAGGAGACAGGCTAGGATAGACAGCGGCAGCGAAGTGATCGTCGGCGTTAACAAATA
CCAGCTGGAAAAGGAAGATGCTGTGGAGGTGCTGGCCATCGACAAACACCTCCGTTCCGA
ACAGACAAATCGAGAAGCTGAAGAAAATCAAGAGCAGCAGAGACCAGGCCCTTGCTGAG
CGCTGTCTGGCCGCACTAACCGAATGCGCCGCCAGCGGCGATGGGAATATCCTGGCCCT
GGCCGTGGACGCTTCTAGGGCAAGATGCACAGTGGGCGAGATCACAGACGCTCTGAAGA
AGGTCTTCGGCGAGCATAAAGCCAACGACCGGATGGTGAGCGGAGCATAACCGGCAGGAG
TTCGGCGAGAGCAAAGAAATTACCTCTGCCATCAAACGCGTGCATAAGTTTATGGAGCG
GGAGGGCAGGAGACCGCGGCTGTTGGTGGCCAAAATGGGACAAGATGGCCACGATCGCG
GTGCCAAGGTGATCGCCACCGGGTTCGCTGACCTGGGCTTCGACGTTGATATAGGCCCT
CTTTTCCAGACACCTCGGGAGGTGGCTCAACAGGCTGTGGATGCCGACGTCCACGCAGT
GGGCGTGAGCACCTGGCCGCGCCGCCACAAGACCCCTGGTGCCCGAGTTAATCAAGGAGT
TGAATTCACCTGGGCAGACCTGACATCCTGGTGATGTGCGGCGGCGTCATCCCCCGCAG
GATTACGAGTTCCTGTTTCGAGGTGGGAGTGAGCAACGTGTTTCGGTCCTGGCACACGGAT
CCCCAAGGCCGCGTCCAGGTTCTGGACGACATCGAGAAGTGCCTGGAAAAGAAACAGC
AGTCCGTG

SEQ ID NO: 151

FIG. 13

ATGCTGAGAGCCAAAAATCAGCTCTTCCTACTGAGCCCTCATTACCTGAGGCAGGTGAA
GGAGAGCTCAGGGAGCCGGCTCATCCAGCAGAGACTGCTCCATCAGCAGCAGCCGCTGC
ACCCGGAGTGGGCCGCCCTGGCCAAGAAACAACCTGAAGGGCAAGAACCCAGAGGACCTT
ATTTGGCACACCCCAGAGGGCATTAGCATCAAGCCTCTCTATAGCAAGAGAGACACCAT
GGACCTCCCCGAGGAGCTGCCCCGGCGTGAAGCCTTTACCCCGTGGCCCCCTACCCAACCA
TGTACACATTCAGGCCCTGGACTATCCGGCAGTACGCGGGCTTCTCTACCGTGGAGGAG
TCCAACAAGTTCTACAAGGACAATATTAAAGCTGGCCAGCAGGGACTGAGCGTGGCCTT
CGATTTGGCCACCCACAGAGGATACGACAGCGACAACCCAGGGTGAGGGGCGACGTTG
GAATGGCCGGCGTGGCCATTGATACTGTGAGGACACCAAGATCCTGTTGATGGGATA
CCCCTCGAAAAAATGTCTGTGTCCATGACAATGAATGGGGCCGTGATCCCAGTGTTGGC
CAATTTTATCGTGACAGGAGAGGAGCAGGGCGTGCCCAAGGAGAAGCTGACCGGCACCA
TCCAGAACGACATCTTGAAAGAATTATGTTAGAAATACCTACATCTTTCCTCCCGAA
CCCAGCATGAAGATCATCGCCGATATCTTCGAGTATACCGCCAAACACATGCCCAAATT
CAACTCCATTAGCATATCCGGATACCATATGCAGGAGGCTGGCGCCGACGCCATCCTGG
AGCTGGCGTACACCCTGGCCGATGGCCTGGAATATTCTAGGACCGGCCTGCAGGCTGGA
CTCACAATCGATGAGTTCGCCCCCGCCTGAGCTTCTTCTGGGGCATCGGCATGAACTT
CTACATGGAGATCGCTAAGATGAGAGCGGGACGCGCCTGTGGGCTCACCTGATCGAAA
AAATGTTCCAACCTAAGAATAGCAAGTCTCTGCTCCTGCGGGCACACTGTCAAACCTCC
GGCTGGTCCCTTACAGAACAGGACCCGTACAACAACATCGTGAGAACCGCCATCGAAGC
CATGGCGGCCGTGTTGCGGGGACCCAGTCACTCCACACCAATAGCTTCGACGAGGCCC
TGGGCTTGCCCTACTGTCAAGTCTGCCAGGATCGCCAGGAACACTCAGATTATCATCCAG
GAGGAGAGCGGCATTCCTAAGGTAGCTGACCCCTGGGGGGGAAGCTACATGATGGAATG
CCTCACCACGACGTGTACGATGCCGCCCTGAAGCTGATTAATGAGATCGAGGAAATGG
GGGGCATGGCCAAGGCTGTAGCCGAGGGAATCCCCAAGCTGAGAATCGAGGAGTGCGCC
GCCAGGCGGCAGGCCCGCATCGATAGTGGCAGCGAGGTGATCGTGGGCGTTAACAAGTA
CCAGCTCGAGAAGGAAGACGCCGTAGAGGTGTTGGCAATCGACAATACCTCCGTGAGAA
ATAGGCAGATCGAGAAGCTAAAGAAGATAAAGAGCAGCCGGGACCAGGCTCTGGCAGAA
CGGTGCCTGGCCGCCCTCACTGAATGCGCCGCTTCTGGCGATGGCAACATTCTGGCCCT
TGCCGTGATGCCAGCCGGGCCGATGCACTGTGGGAGAGATCACAGACGCTCTGAAGA
AGGTGTTCCGGCGAGCATAAAGCAAACGATAGAATGGTGTCCGGCGCCTACAGGCAGGAG
TTTGGCGAGAGCAAGGAGATCACATCTGCTATTAAGAGAGTGCATAAGTTTCATGGAGAG
AGAGGGACGGCGGCCTAGACTGCTGGTAGCAAAGATGGGCCAGGACGGCCACGATCGGG
GAGCAAAGGTCATTGCTACCGGGTTCGCCGACCTGGGCTTCGACGTGGACATTGGCCCC
CTCTTCCAGACCCCCCGCAAGTTGCTCAGCAGGCTGTGGATGCTGACGTGCACGCTGT
GGGCGTCAGCACACTGGCCGCTGGACATAAGACCCTGGTTCCCGAGCTGATCAAGGAGC
TGAATAGCCTGGGTGCGCCAGACATCCTCGTGATGTGCGGGGGTGTGATCCCCCCTCAG
GATTACGAGTTCCTGTTGAGGTTGGAGTCTCTAATGTGTTTCGGCCCCGGAACAAGAAT
CCCGAAGGCCGCTGTGCAAGTCCTGGACGACATCGAAAAGTGCCTCGAGAAGAAGCAGC
AGTCTGTG

SEQ ID NO: 152

FIG. 14

ATGCTGAGGGCCAAGAACCAGCTGTTCTCTCTCTCCTCACTACCTGAGACAGGTGAA
GGAGTCCAGCGGTAGCAGACTGATCCAGCAGAGGCTGCTTCACCAGCAGCAGCCCCCTCC
ATCCGGAGTGGGCTGCACTGGCAAAAAGCAGCTTAAGGGAAAAAACCCCGAGGACCTG
ATTTGGCACACCCCAGAGGGAATTAGCATTAACCCCTCTACAGCAAACGGGACACGAT
GGACCTGCCCCGAGGAACTGCCCCGGCGTGAAGCCCTTCACACGGGGGCCATACCCACAA
TGTATACATTCCGTCTTGGACCATTCGCCAATACGCCGGGTTCAGCACGGTGGAGGAA
TCCAACAAGTTCTACAAGGACAACATCAAGGCCGGCCAGCAAGGCCTCAGCGTCGCCTT
TGACTTGGCCACCCACAGAGGCTACGACTCCGACAACCCAGAGTGAGAGGCGACGTGG
GTATGGCCGGCGTGGCCATCGACACTGTTGAGGACACCAAAATCCTATTTGACGGCATA
CCCCTGGAAAAGATGTCCGTGAGCATGACAATGAATGGCGCTGTGATCCCCGTGCTCGC
TAACTTCATCGTGACCGGCGAGGAGCAGGGTGTCCCTAAGGAGAACTTACCGGCACAA
TCCAGAACGACATTCTCAAGGAGTTTATGGTGCGCAACACCTACATTTTCCCCCAGAG
CCTAGCATGAAGATCATTGCCGACATCTTCGAGTACACCGCCAAGCATATGCCAAAGTT
TAACTCCATTTCAATCTCTGGCTACCATATGCAGGAAGCTGGCGCTGACGCCATCTTAG
AGCTGGCCTACACCTTGGCCGACGGCCTCGAATACAGCAGGACCGGCCTTCAGGCCGGT
CTGACCATTGACGAATTTGCACCCAGACTGAGCTTCTTCTGGGGCATCGGCATGAACTT
CTATATGGAGATCGCTAAAATGAGAGCCGGCCGGAGACTGTGGGCTCACCTGATCGAGA
AAATGTTTCAACCCAAGAACTCCAAAAGCCTTCTGCTCAGGGGCCATTGTCAAACCTCT
GGATGGAGCCTGACTGAGCAGGACCCGTATAACAACATCGTCAGGACCGCCATCGAGGC
CATGGCTGCGGTGTTTGGCGGAACCCAGTCCCTACACACAAATAGCTTTGACGAAGCCC
TGGGGCTGCCAACCGTGAAAAGCGCCAGAATCGCCCGGAATACTCAGATAATCATCCAG
GAGGAGTCCGGGATCCCCAAGGTTGCCGACCCCTGGGGGGGAAGTTACATGATGGAGTG
CTTGACCAACGATGTGTACGATGCTGCTCTGAAGCTGATTAACGAGATTGAGGAGATGG
GAGGCATGGCTAAGGCTGTGCGCGAGGGCATCCCTAAGCTGAGAATCGAGGAGTGCGCC
GCCCCGCCGGCAGGCCAGGATTGACAGCGGATCTGAGGTGATCGTGGGGGTGAACAAGTA
CCAGCTGGAGAAGGAGGACGCCGTTGAGGTTCTGGCTATCGACAACACCAGCGTCAGAA
ACAGGCAGATCGAGAAGCTGAAGAAAATCAAAGCAGCAGAGACCAGGCCCTGGCAGAA
AGATGCCTGGCTGCCCTGACCGAGTGCGCTGCTAGCGGAGACGGGAACATCCTGGCTCT
GGCCGTCGACGCTTCCAGAGCCAGGTGTACCGTCCGGCGAGATCACCGACGCCCTGAAGA
AAGTGTTCCGGCGAGCACAAGGCCAACGACCGCATGGTGTCCGGCGCCTACAGGCAGGAG
TTCGGCGAGTCCAAGGAGATTACCTCCGCCATCAAACGGGTGCACAAATTCATGGAGAG
AGAGGGCAGACGGCCACGCTTGCTCGTGGCCAAGATGGGGCAGGACGGCCACGATAGGG
GCGCTAAGGTTATCGCTACAGGCTTTGCTGATCTGGGCTTCGACGTGGACATCGGCCCC
CTGTTTCAGACCCCTAGAGAGGTAGCTCAGCAGGCCGTCGACGCCGACGTGCACGCCGT
GGGCGTCAGCACACTGGCAGCCGGACACAAGACCCTCGTGCCCCGAGCTTATCAAGGAAC
TGAACCTCCTTGGGGCGGGCCCGACATCCTCGTCATGTGCGGCGGAGTGATTCCCCCCCAG
GACTATGAGTTCCTGTTCGAGGTCGGCGTAAGCAACGTCCTTTGGCCCCGGAACTCGGAT
CCCCAAGGCCGCGAGTACAGGTGCTGGACGACATCGAAAAGTGCCTGGAGAAGAAGCAAC
AGTCTGTG

SEQ ID NO: 153

FIG. 15

ATGCTGCGCGCTAAAAACCAACTATTCCTGCTGTCCCCCATTACCTGAGACAGGTGAA
GGAGAGCTCCGGAAGCAGGCTGATCCAACAGAGACTTCTGCACCAGCAGCAGCCCCTGC
ATCCCGAGTGGGCGCTCTGGCTAAGAAACAACCTGAAGGGAAAGAACCCCGAGGACCTG
ATCTGGCACACCCCAGAGGGCATATCCATCAAGCCCCTGTATTCTAAGAGGGACACTAT
GGACCTGCCCCGAGGAGCTCCCCGGCGTCAAGCCTTTCACACGCGGACCTTACCCACAA
TGTACACATTCCGACCATGGACCATTAGGCAGTACGCAGGATTCAGCACCGTTGAGGAG
TCCAATAAATTCTATAAGGACAATATTAAGGCCGACAACAGGGCCTGAGTGTGGCATT
CGACCTGGCCACACACAGAGGTTACGACTCCGATAATCCTAGGGTGAGAGGCGACGTGG
GCATGGCTGGAGTGGCCATCGATAACCGTCGAGGACACAAAGATCCTGTTTCGACGGCATC
CCCTTGAGAAAATGTCCGTGAGCATGACCATGAACGGCGCCGTGATTCCCGTGCTGGC
CAACTTTATTGTAACCGGCGAGGAGCAGGGCGTCCCCAAGGAGAAGCTGACCGGCACCA
TCCAGAACGATATCCTCAAGGAGTTCATGGTACGGAATACCTACATTTTTCCCCCGAA
CCCTCCATGAAGATCATCGCCGACATCTTCGAGTACACCGCCAAGCACATGCCCAAGTT
CAACTCCATCAGTATTAGCGGCTATCACATGCAGGAGGCCGCGCCGACGCCATCCTGG
AACTGGCCTACACACTGGCCGACGGTCTGGAGTACAGCAGAACCGGCCTGCAAGCAGGC
CTGACCATCGACGAGTTCGCCCCCTCGACTGAGCTTCTTCTGGGGCATTGGCATGAACTT
TTACATGGAGATCGCCAAAATGCGCGCTGGCAGGCGACTCTGGGGCCACCTGATAGAAA
AAATGTTCCAGCCCAAGAACTCCAAGTCCCTGCTCCTCAGGGCCCACTGTCAGACTAGC
GGCTGGAGCCTCACCGAGCAGGACCCATATAATAACATTGTAAGAACAGCCATCGAGGC
CATGGCCGCGAGTGTTCGGAGGGACCCAGAGCCTGCATACCAACAGCTTCGATGAGGCC
TGGGACTGCCCACCTGTAAAGTCCGCTAGAATAGCCCGGAACACCCAGATCATTTATCCAG
GAAGAGTCCGGGATCCCCAAGGTGGCCGACCCATGGGGCGGCAGCTACATGATGGAGTG
TCTCACCAACGACGTTTACGATGCTGCCCTGAAGCTTATCAACGAAATCGAGGAGATGG
GCGGCATGGCCAAAGCCGTTCGCCGAGGGGATTCCAAAGCTCCGGATCGAGGAGTGCGCC
GCCAGAAGGCAGGCCAGAATCGACTCTGGAAGCGAGGTGATTGTGGGTGTGAATAAGTA
TCAGCTTGAGAAGGAAGACGCTGTAGAGGTGCTGGCTATAGATAACACCAGCGTCAGGA
ACAGGCAGATTGAAAAGCTGAAGAAGATCAAGTCGAGCAGAGATCAGGCGCTGGCCGAG
AGATGCCTGGCCGCCCTGACTGAGTGCGCCGCCTCTGGCGATGGCAACATCCTCGCACT
GGCCGTGGACGCTAGCCGGGCCCCGGTGTACCGTGGGCGAGATAACCGATGCCCTGAAGA
AGGTGTTCCGAGAGCACAAGGCCAACGACAGAATGGTGTCTGGCGCATATAGACAGGAG
TTCGGCGAGAGCAAGGAGATTACGAGCGCCATCAAGAGAGTGCACAAGTTTCATGGAGAG
GGAGGGCAGGAGACCACGGCTGCTGGTTGCCAAGATGGGCCAAGACGGTTCACGACAGAG
GGGCTAAAGTGATTGCCACCGGCTTCGCCGACCTGGGCTTCGACGTGGACATCGGTCCA
CTGTTCCAGACCCCTAGGGAGGTGGCGCAGCAGGCCGTGGACGCCGACGTGCACGCCGT
CGGCGTGAGCACACTGGCTGCCGGCCACAAGACACTAGTTCCTGAGCTGATCAAGGAAC
TCAACAGCCTGGGCCGGCCTGATATACTGGTGATGTGCGGGGGGGTGATCCCCCCTCAG
GATTACGAGTTTCTGTTTGAGGTGGGCGTGTCCAACGTCTTCGGCCCCGGGACCAGAAAT
CCCAAAGGCCGCGAGTGCAGGTATTAGACGACATTCAGAAGTGCCTGGAGAAGAAACAGC
AGTCCGTT

SEQ ID NO: 154

FIG. 16A

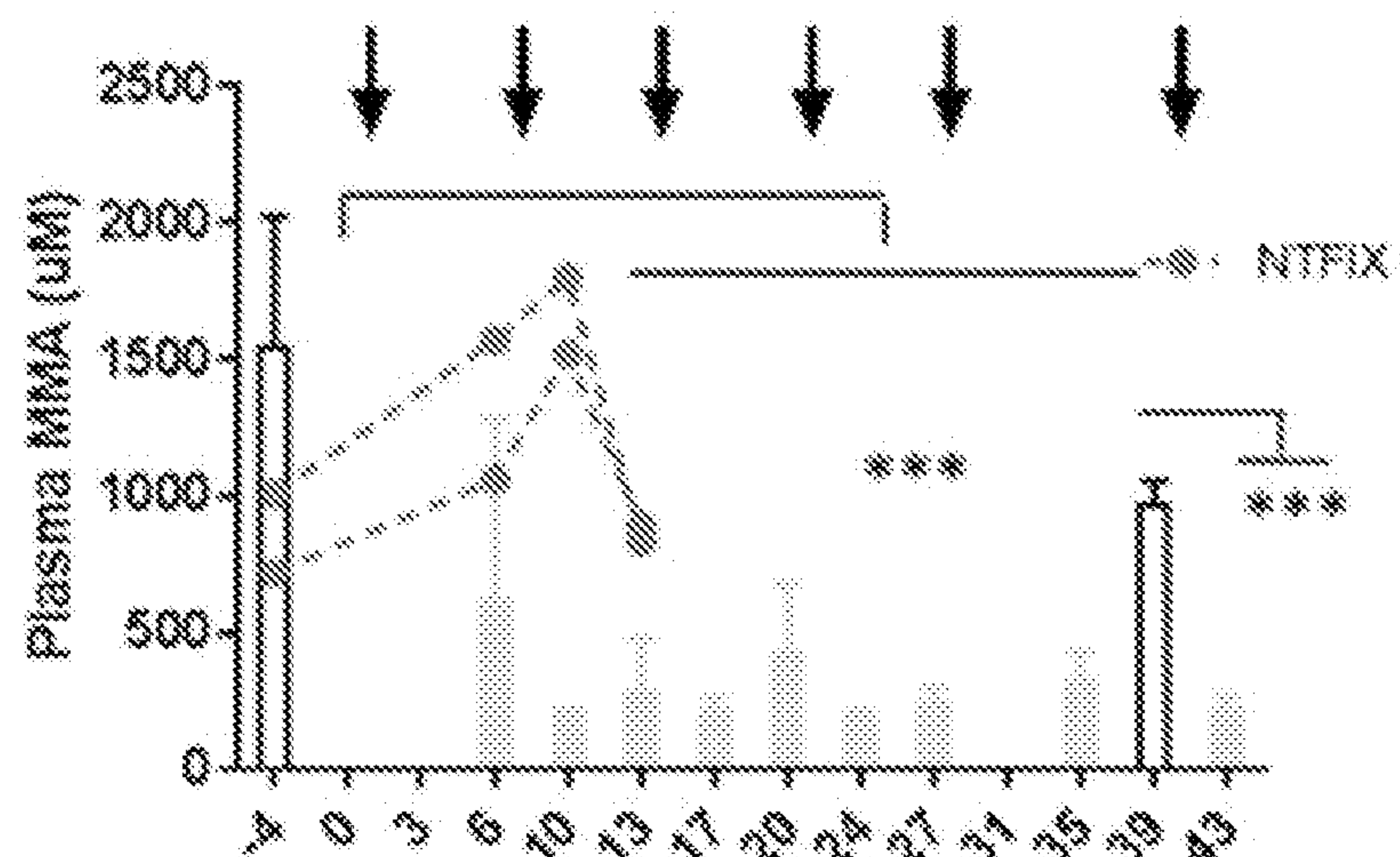


FIG. 16B

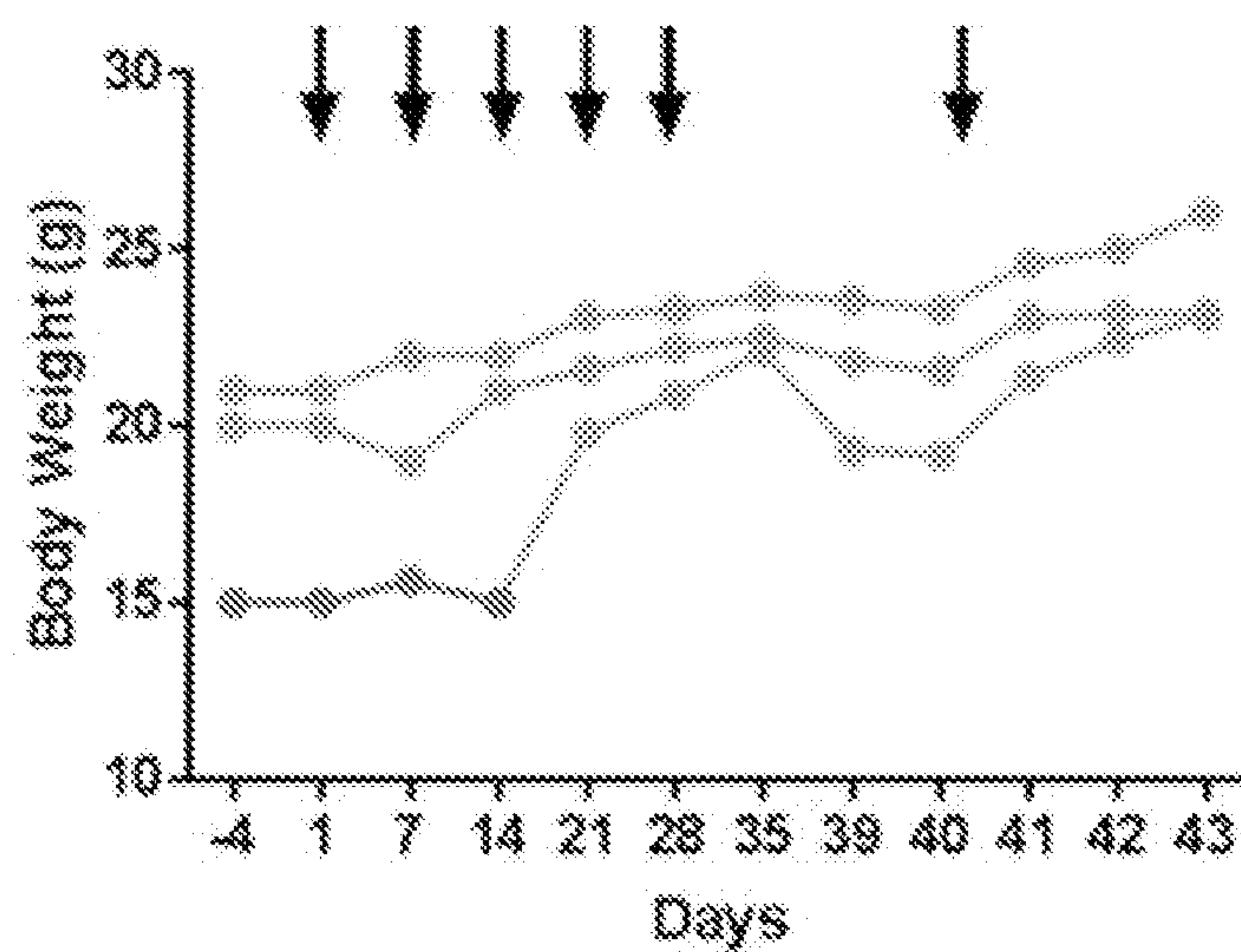
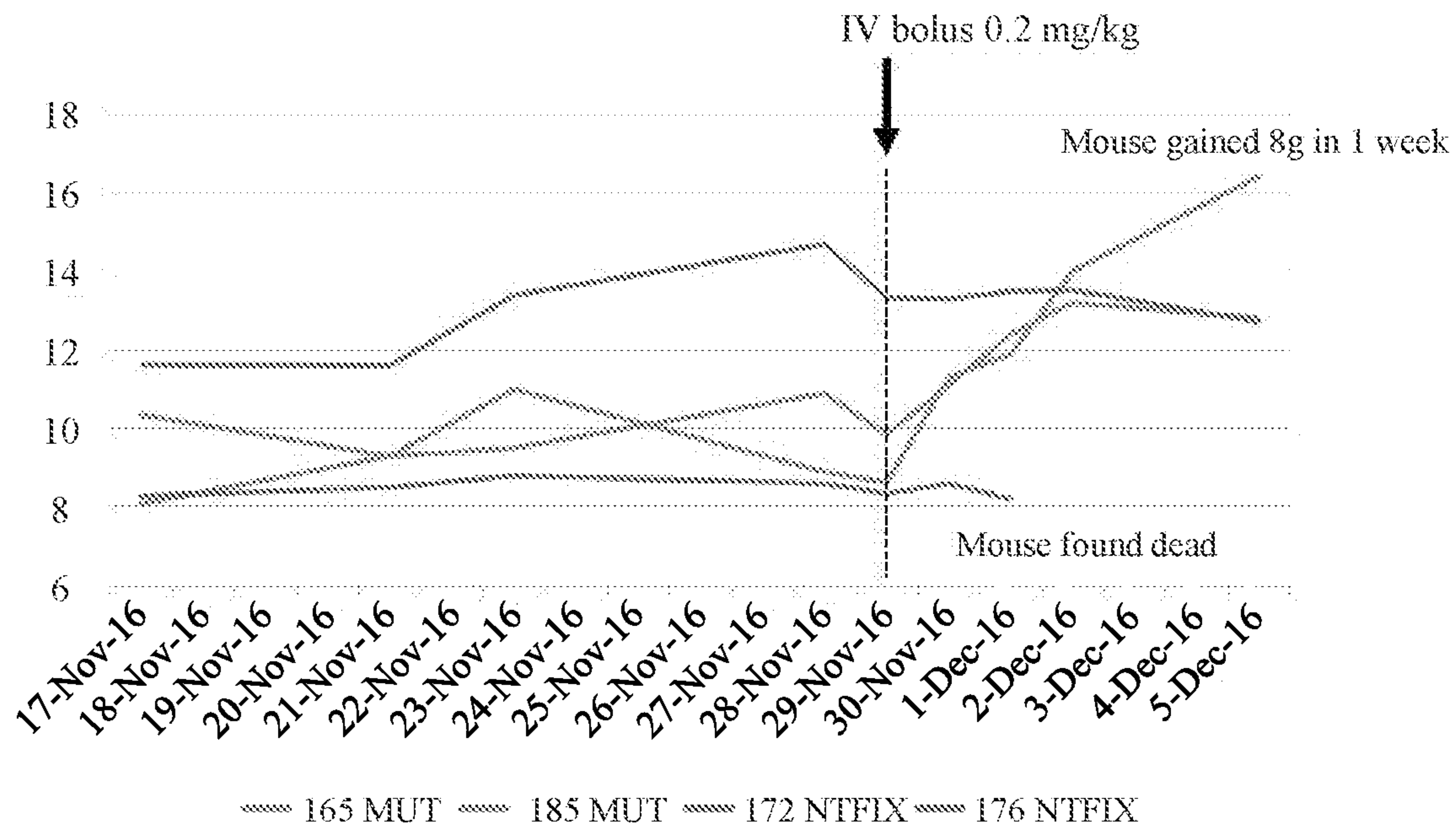
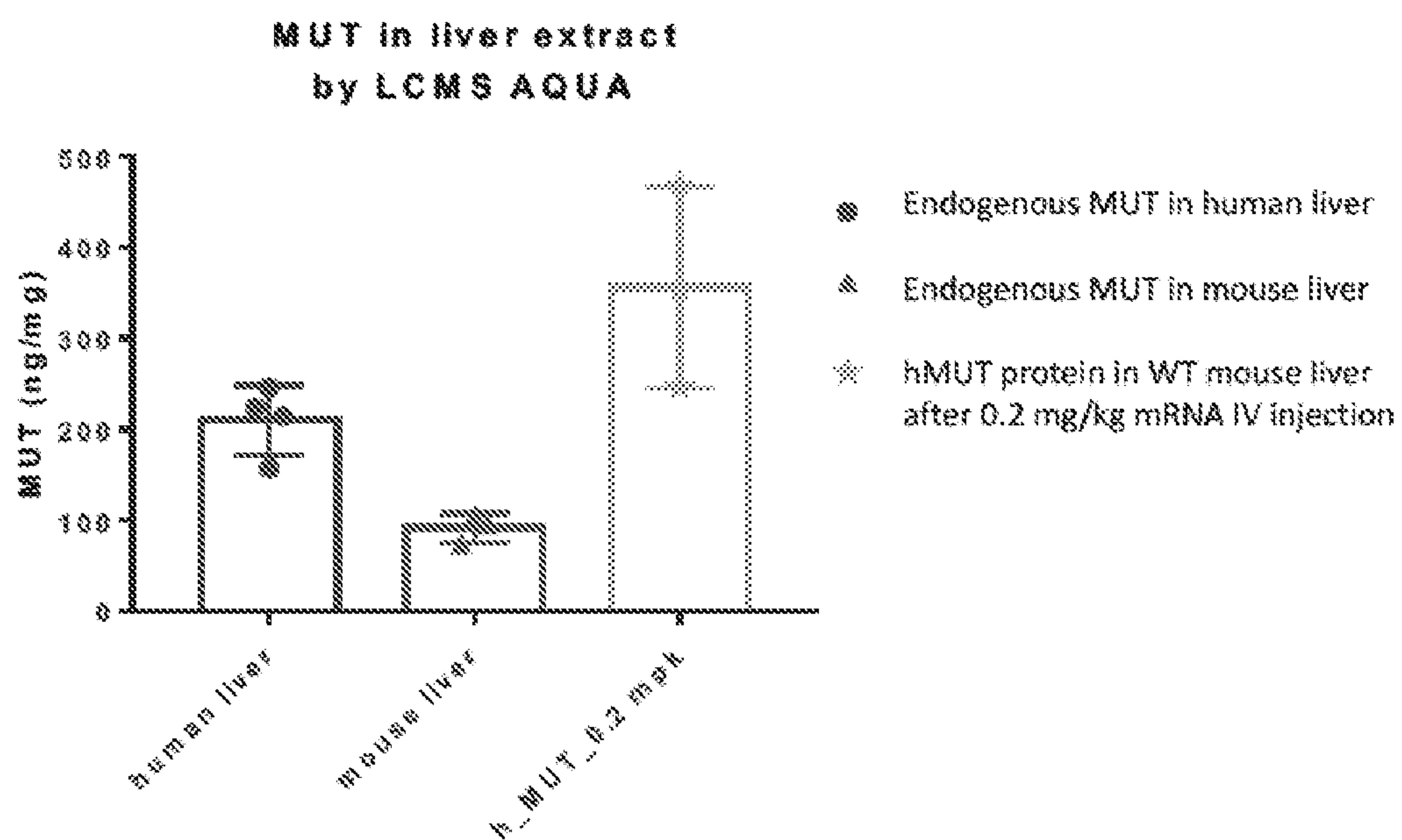


FIG. 16C



Study Design: Weekly IV injections (hMUT or NTFIX mRNA) into MCK mice

FIG. 17



19/21

FIG. 18A

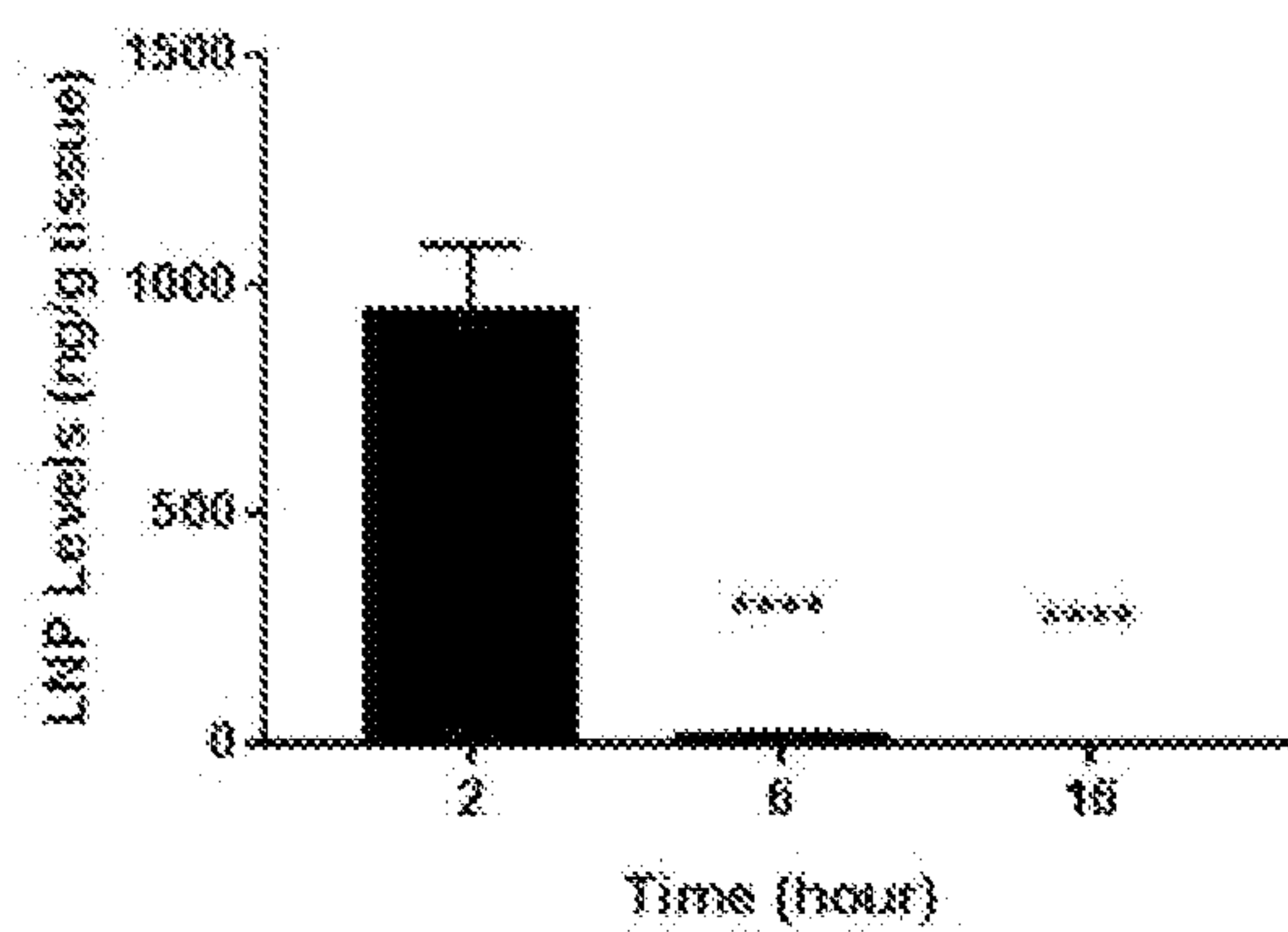


FIG. 18B

hMUT mRNA in liver (bDNA)

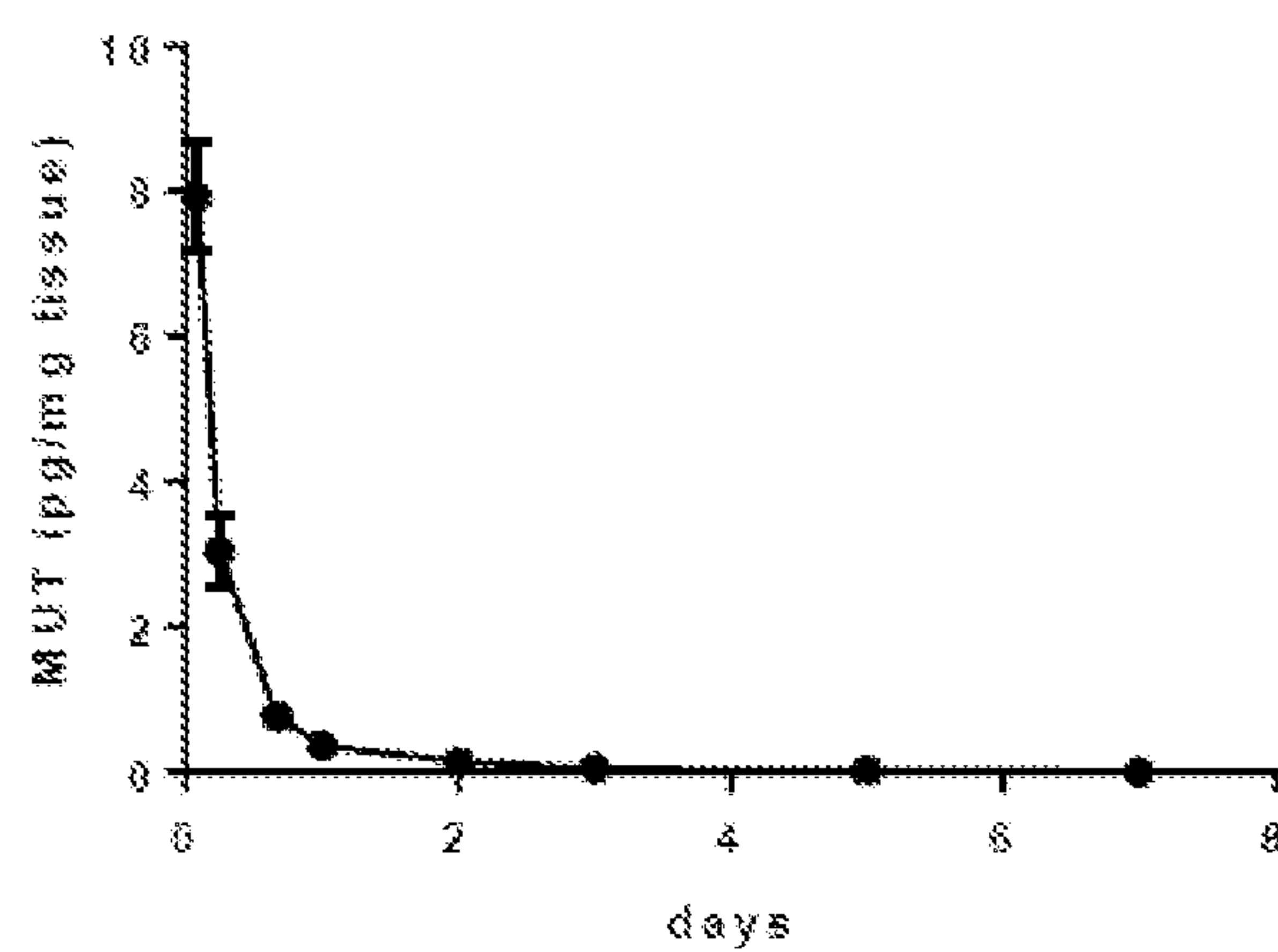


FIG. 18C

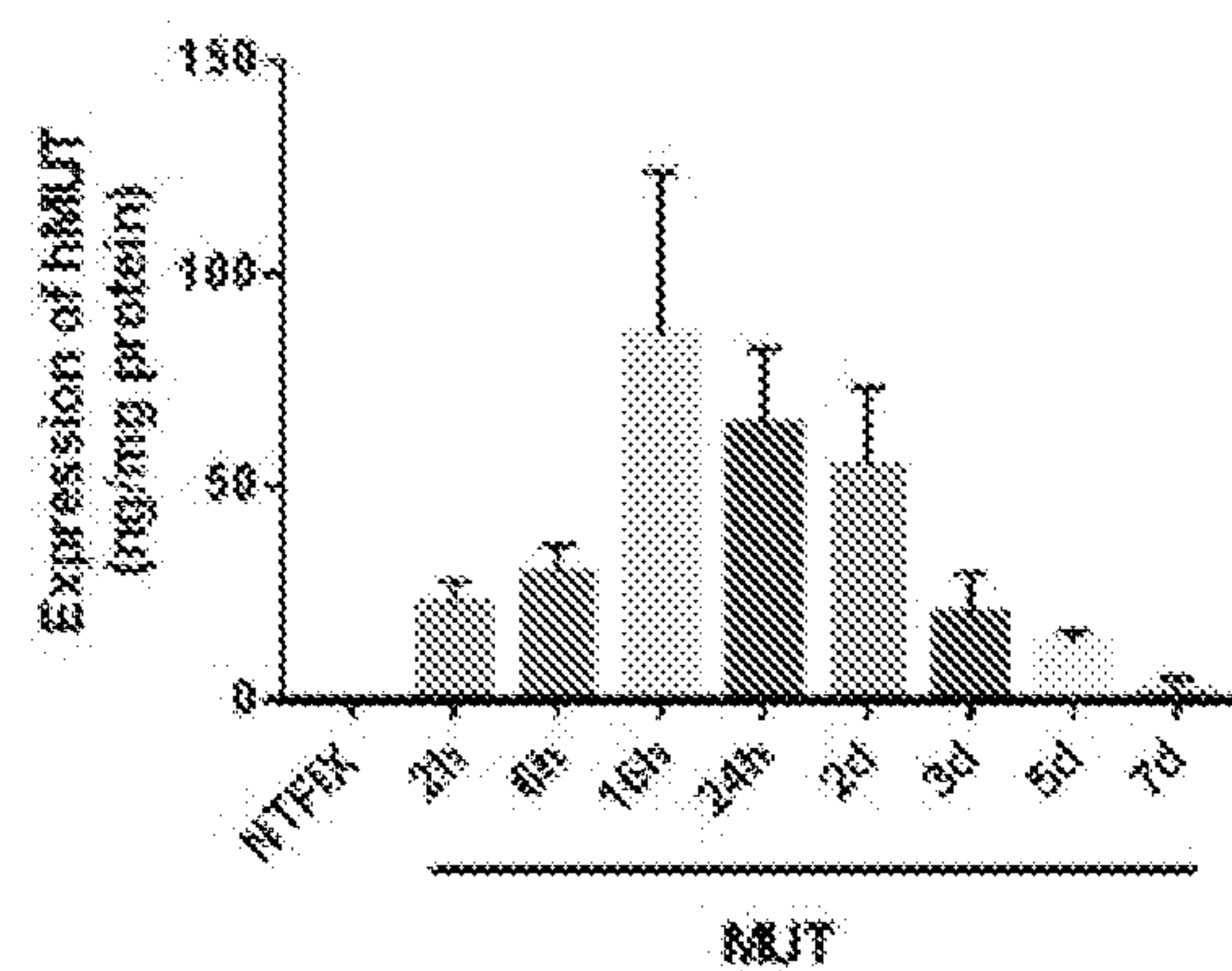


FIG. 19A

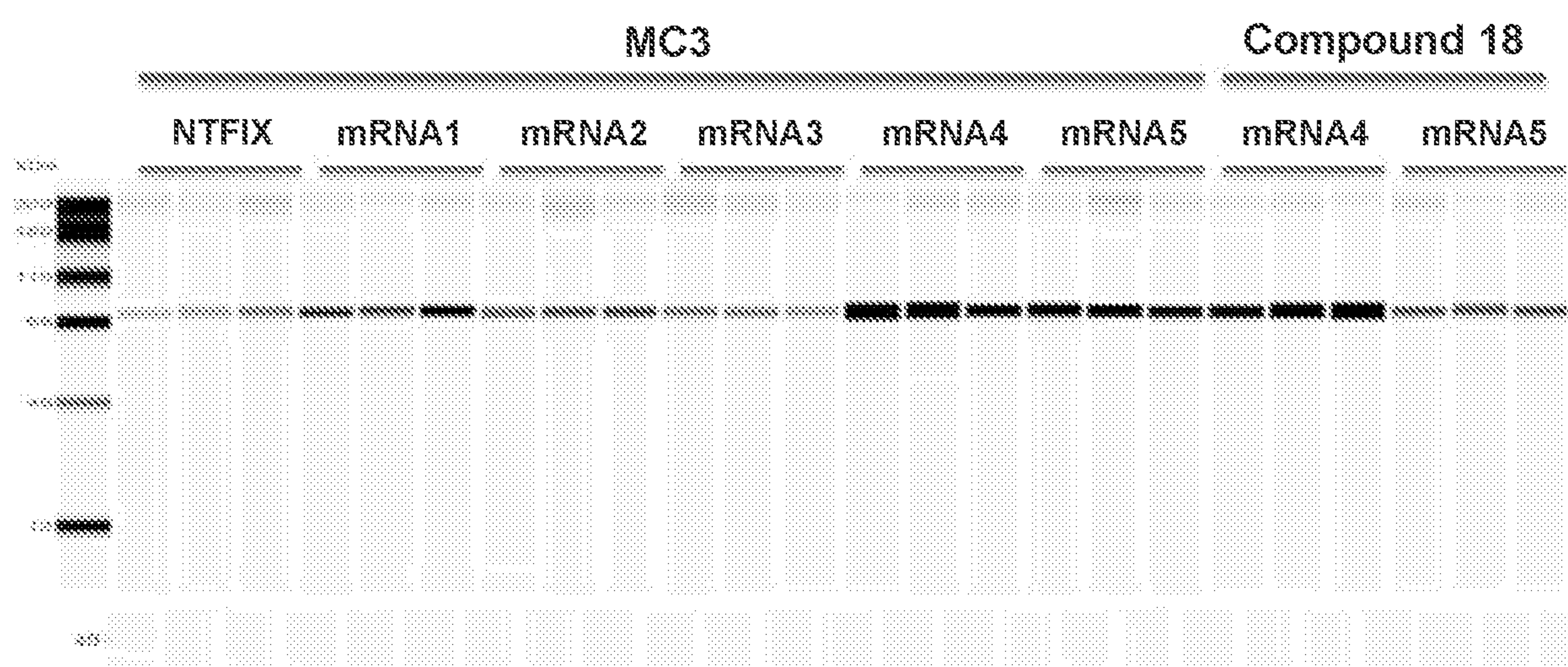


FIG. 19B

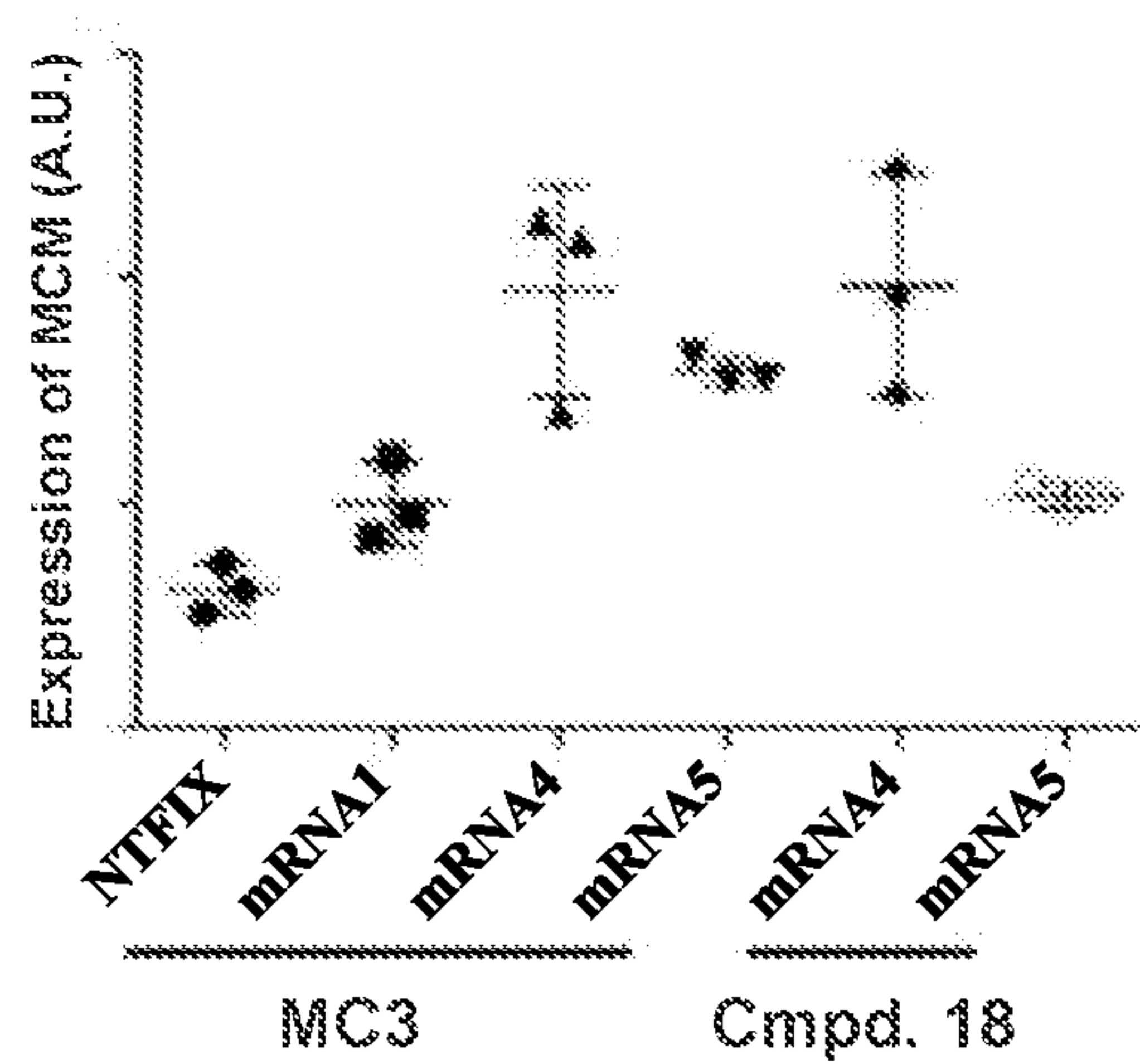


FIG. 20A

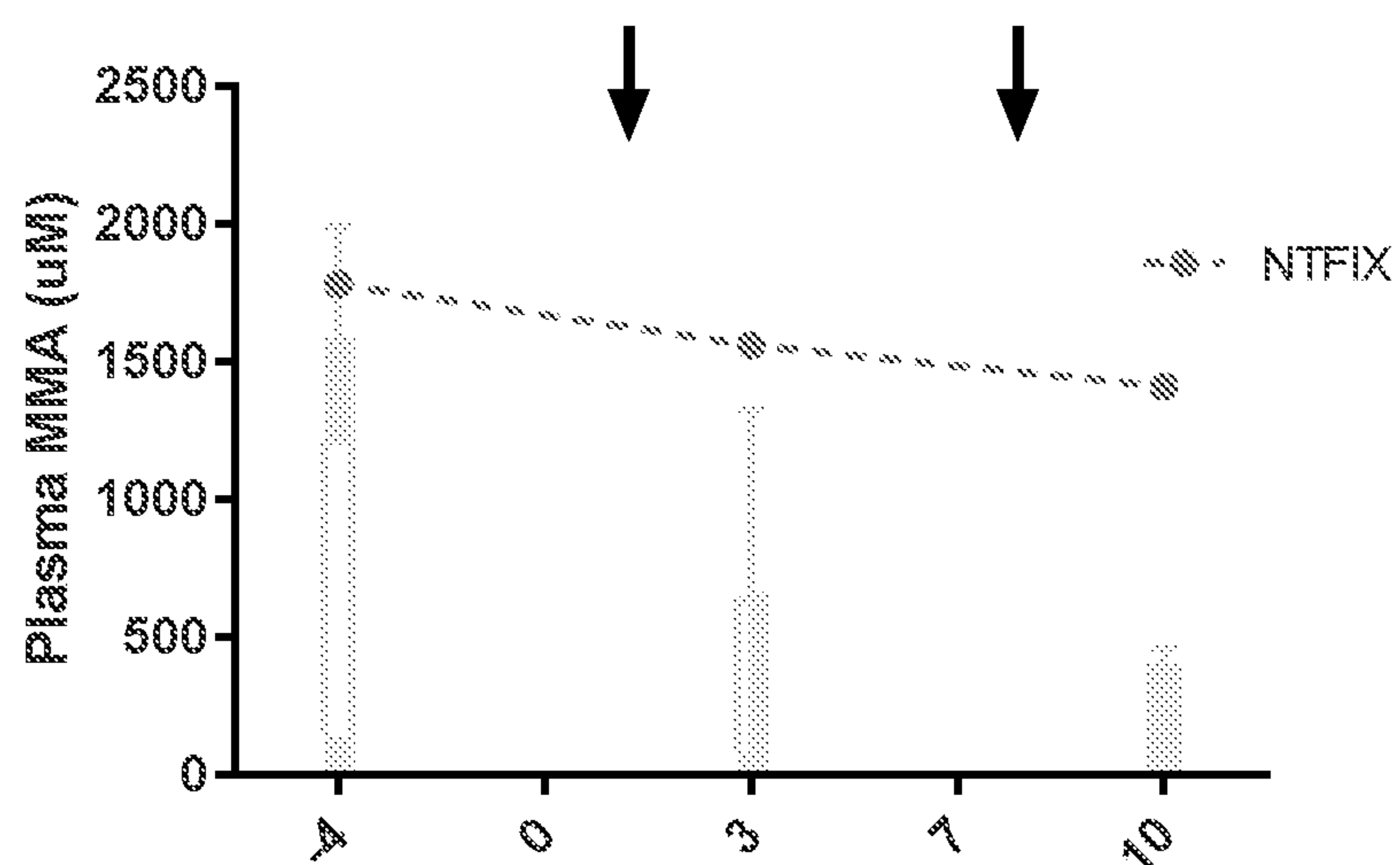


FIG. 20B

