



(51) International Patent Classification:

A61K 39/08 (2006.01) C12N 15/63 (2006.01)
C07K 14/33 (2006.01) A61P 31/04 (2006.01)

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(21) International Application Number:

PCT/US2024/060394

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(22) International Filing Date:

16 December 2024 (16.12.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/610,388 14 December 2023 (14.12.2023) US

(71) Applicants: **THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA** [US/US]; 3600 Civic Center Blvd., 9th Floor, Philadelphia, PA 19104 (US). **THE CHILDREN'S HOSPITAL OF PHILADELPHIA** [US/US]; 3401 Civic Center Blvd., Philadelphia, PA 19104 (US).

(72) Inventors: **ALAMEH, Mohamad**; c/o The Trustees of the University of Pennsylvania, 3600 Civic Center Blvd., 9th Floor, Philadelphia, PA 19104 (US). **ZACKULAR, Joseph**; c/o The Children's Hospital of Philadelphia, 3401 Civic Center Blvd., Philadelphia, PA 19104 (US). **WEISSMAN, Drew**; 614 S. Hancock Street, Philadelphia, PA 19147 (US). **SEMON, Alexa**; 2004 Lombard Street Apt 2R, Philadelphia, PA 19146 (US).

(74) Agent: **DIVITO EVANS, Abigail** et al.; Riverside Law LLP, 175 Strafford Ave, Suite 100, Wayne, PA 19087 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,

(54) Title: CLOSTRIDIODES DIFFICILE VACCINE AND METHODS OF USE

(57) Abstract: The present disclosure provides compositions (e.g., pharmaceutical compositions, e.g., RNA-LNP vaccines) for delivery of *C. difficile* antigens and related technologies (e.g., components thereof and/or methods relating thereto).

WO 2025/129186 A1

Clostridioides difficile VACCINE AND METHODS OF USE

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application No.
5 63/610,388, filed December 14, 2023, which is hereby incorporated by reference herein
in its entirety.

BACKGROUND

C. difficile is a spore-forming anaerobic bacterium that infects the colon,
10 causing a wide range of disorders that vary in severity from mild diarrhea to toxic
megacolon and death. In the United States, *C. difficile* is the most commonly reported
nosocomial pathogen, and *C. difficile* infection (CDI) is a major public health threat
worldwide.

Thus, there is a need for compositions and methods for treating and
15 preventing *C. difficile* infection. This invention fulfills this unmet need.

SUMMARY OF THE INVENTION

In one embodiment, the invention relates to a composition comprising at
20 least one RNA molecule encoding at least one *C. difficile* antigen. In one embodiment, at
least one *C. difficile* antigen comprises *C. difficile* toxin A (TcdA), *C. difficile* toxin B
(TcdB), PPEP-1 (Zmp1), CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant
thereof, or any combination thereof. In one embodiment, the composition comprises a
combination of RNA molecules encoding at least two *C. difficile* antigens, wherein the
25 composition comprises at least two of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66,
ZupT, CspC, or fragments or variants thereof. In one embodiment, the composition
comprises a combination of RNA molecules encoding TcdA, TcdB, PPEP-1, and CdeM,
or fragments or variants thereof. In one embodiment, the composition comprises a
combination of RNA molecules encoding TcdA and TcdB, or fragments or variants
30 thereof. In one embodiment, the composition comprises a combination of RNA
molecules encoding TcdA, TcdB, and PPEP-1, or fragments or variants thereof.

In one embodiment, the fragment of TcdA or TcdB comprises a receptor binding domain (RBD) including a combined repetitive oligopeptide (CROPs) domain.

In one embodiment, the composition comprises an RNA molecule encoding a variant of at least one *C. difficile* antigen comprising an amino acid substitution which disrupts an N-glycosylation site.

In one embodiment, the RNA molecule comprises a nucleic acid sequence encoding an amino acid sequence having at least 80% identity to an amino acid sequence of:

- a) a TcdA antigen comprising an amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;
- b) a TcdB antigen comprising an amino acid sequence of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;
- c) a PPEP-1 antigen comprising an amino acid sequence of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;
- d) a CdeM antigen comprising an amino acid sequence of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;
- e) a CWP84 antigen comprising an amino acid sequence of SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;
- f) a CWP66 antigen comprising an amino acid sequence of SEQ ID NO:28, or a fragment or variant thereof;
- g) a ZupT antigen comprising an amino acid sequence of SEQ ID NO:29, or a fragment or variant thereof; or
- h) a CspC antigen comprising an amino acid sequence of SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a

fragment or variant thereof, encoding a TcdA antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof, encoding a TcdB antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a fragment or variant thereof, encoding a PPEP-1 antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74, or a fragment or variant thereof, encoding a CdeM antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78, or a fragment or variant thereof, encoding a CWP84 antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:79 or SEQ ID NO:80, or a fragment or variant thereof encoding a CWP66 antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:81 or SEQ ID NO:82, or a fragment or variant thereof, encoding a ZupT antigen. In one embodiment, the composition comprises an RNA molecule comprising a nucleic acid sequence having at least 80% identity to SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88, encoding a CspC antigen.

In one embodiment, the composition comprises a combination of RNA molecules, wherein the combination of RNA molecules comprises at least one of:

- a) an RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;
- b) an RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10,

SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) an RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) an RNA molecule encoding SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) an RNA molecule encoding SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) an RNA molecule encoding SEQ ID NO:28, or a fragment or variant thereof;

g) an RNA molecule encoding SEQ ID NO:29, or a fragment or variant thereof;
and

h) an RNA molecule encoding SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

In one embodiment, the composition comprises a combination of RNA molecules encoding at least two amino acid sequences selected from:

a) an RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) an RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) an RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) an RNA molecule encoding SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) an RNA molecule encoding SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) an RNA molecule encoding SEQ ID NO:28, or a fragment or variant thereof;

g) an RNA molecule encoding SEQ ID NO:29, or a fragment or variant thereof;

and

h) an RNA molecule encoding SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

5 In one embodiment, the composition comprises a combination of RNA molecules encoding at least three amino acid sequences selected from:

a) an RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

10 b) an RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) an RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

15 d) an RNA molecule encoding SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) an RNA molecule encoding SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) an RNA molecule encoding SEQ ID NO:28, or a fragment or variant thereof;

20 g) an RNA molecule encoding SEQ ID NO:29, or a fragment or variant thereof; and

h) an RNA molecule encoding SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

25 In one embodiment, the RNA molecule is mRNA, self-replicating RNA, self-amplifying RNA, or circular RNA.

In one embodiment, at least one RNA molecule is a nucleoside modified RNA molecule comprising at least one modified nucleoside. In one embodiment, at least one modified nucleoside comprises pseudouridine, 1-methyl pseudouridine, or 5-methyl-uridine.

30 In one embodiment, the RNA molecule comprises a nucleic acid sequence encoding a leader sequence, a tag or signal peptide. In one embodiment, the RNA

molecule further comprises a linker.

In one embodiment, the composition comprises one or more lipid nanoparticle (LNP), wherein the one or more RNA molecules are partially or fully encapsulated within the one or more LNP. In one embodiment, the composition

5 comprises a combination of at least two LNPs, wherein each LNP encapsulates an RNA molecule encoding a *C. difficile* antigen. In one embodiment, each LNP encapsulates an RNA molecule encoding TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC.

In one embodiment, the invention relates to a pharmaceutical composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein
10 the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof. In one embodiment, the composition further comprises at least one pharmaceutically acceptable excipient.

In one embodiment, the invention relates to a combination comprising:

15 (i) a first pharmaceutical composition comprising a first RNA molecule, wherein the first RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof; and

(ii) a second pharmaceutical composition comprising a second RNA
20 molecule, wherein the second RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof,

wherein the first RNA molecule and the second RNA molecule are not the same, and wherein the RNA molecules are partially or fully encapsulated within one or more
25 LNPs.

In one embodiment, the invention relates to a combination comprising:

(i) a first pharmaceutical composition comprising a first RNA molecule, wherein the first RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66,
30 ZupT, CspC, a fragment thereof, or a variant thereof;

(ii) a second pharmaceutical composition comprising a second RNA molecule, wherein the second RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof; and

5 (iii) a third pharmaceutical composition comprising a third RNA molecule, wherein the third RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof;

wherein the first RNA molecule, the second RNA molecule, and the third RNA molecule are not the same, and wherein the RNA molecules are partially or fully encapsulated within one or more LNPs.

In one embodiment, the invention relates to a method of inducing an immune response against at least one strain of *C. difficile* infection in a subject, comprising administering a composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof, or a pharmaceutical composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof.

In one embodiment, the invention relates to a method of treating or preventing a disease or disorder associated with *C. difficile* infection in a subject, comprising administering a composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof, or a pharmaceutical composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC. In one embodiment, the composition is administered by an intradermal, subcutaneous, inhalation, intranasal, or intramuscular delivery route. In one embodiment, the method comprises a single administration of the composition. In one embodiment, the

method comprises multiple administrations of the composition. In one embodiment, the composition induces a broad immune response against multiple strains of *C. difficile* in a cell, tissue or subject. In one embodiment, the composition induces a protective immune response in the subject. In one embodiment, the composition challenges and/or hinders *C. difficile* colonization in a subject. In one embodiment, the disease or disorder is colitis or diarrhea. In one embodiment, administration does not affect intestinal microbiota in the subject.

In one embodiment, the composition is for use in the treatment or prevention of a *C. difficile* infection in a subject comprising administering one or more doses of the composition to a subject. In one embodiment, the composition is for use in the treatment or prevention of a *C. difficile* infection in a subject comprising administering one or more doses of the composition to a subject.

In one embodiment, the composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof, further comprises a delivery vehicle. In one embodiment, the delivery vehicle is poly(ethylenimine) PEI, chitosan, Janus dendrimers, dendrimers, lipopolymers, or a one component system.

In one embodiment, the composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof, further comprises an adjuvant. In one embodiment, the adjuvant is alum, ASO3, squalene, liposomes, chitosan, Matrix M, a cytokine, a nucleic acid molecule encoding an adjuvant molecule, or any combination thereof. In one embodiment, the nucleic acid molecule encodes IL-12, IL-21, or APRIL.

In one embodiment, the composition comprises an RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, operably linked to nucleotide sequence encoding an oligomerization domain. In one embodiment, the oligomerization domain comprises ferritin, or a fragment or variant thereof.

In one embodiment, the invention relates to a composition comprising a fusion molecule comprising a fusion of a *C. difficile* antigen to an oligomerization domain. In one embodiment, the oligomerization domain comprises ferritin, or a fragment or variant thereof. In one embodiment, the *C. difficile* antigen comprises an amino acid sequence having at least 80% identity to an amino acid sequence of:

a) a TcdA antigen comprising an amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) a TcdB antigen comprising an amino acid sequence of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) a PPEP-1 antigen comprising an amino acid sequence of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) a CdeM antigen comprising an amino acid sequence of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) a CWP84 antigen comprising an amino acid sequence of SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) a CWP66 antigen comprising an amino acid sequence of SEQ ID NO:28, or a fragment or variant thereof;

g) a ZupT antigen comprising an amino acid sequence of SEQ ID NO:29, or a fragment or variant thereof; or

h) a CspC antigen comprising an amino acid sequence of SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

The following detailed description of embodiments provided herein will be better understood when read in conjunction with the appended drawings. It should be understood that embodiments provided in the present disclosure are not limited to the precise arrangements and instrumentalities of the embodiments shown in the drawings.

Figure 1A through Figure 1L depict an approach, construct design, and

physicochemical characterization of exemplary RNA-LNP formulations used in the assessment of immunogenicity and potency of an exemplary multivalent vaccine against *C. difficile*. Figure 1A depicts a maximum likelihood Phylogenetic tree illustrating the amino acid sequence identity of TcdA, TcdB, and PPEP-1 from the RNA constructs compared with 137 unique representative *C. difficile* strains. Strains are grouped across five distinct phylogenomic clades, represented by different colors. Figure 1B-Figure 1E depicts the results of example experiments where mice were immunized intramuscularly (i.m.) once (Figure 1C) or twice (Figure 1B, Figure 1D, and Figure 1E) with 1 or 5 μ g of bivalent RNA-LNPs (TcdA/TcdB) (green), trivalent RNA-LNPs (TcdA/TcdB/PPEP-1) (purple), or trivalent recombinant protein with alum (gray). Age-matched naïve mice served as unvaccinated controls. Figure 1B depicts the results of example experiments wherein two weeks post-boost, TcdA-, TcdB-, and PPEP-1-specific antibodies in sera (left, middle, and right, respectively), were measured by ELISA. Figure 1C depicts the results of example experiments wherein two weeks after prime, total number of T follicular helper (Tfh) cells in draining lymph nodes (dLN) were measured by flow cytometry. Figure 1D and Figure 1E depict the results of example experiments wherein two weeks post-boost, total germinal center (GC) B cells (Figure 1D) and antigen-specific B cells (Figure 1E) were measured in the spleen by flow cytometry. Figure 1F and Figure 1G depicts the results of example experiments wherein two weeks after immunization, mice were challenged intraperitoneally (i.p.) with 625 ng recombinant TcdA (Figure 1F) or 125 ng recombinant TcdB (Figure 1G) and monitored for survival. Figure 1B to Figure 1G n = 5 to 14 mice per group, two independent experiments. Data are represented as mean $\pm$ SEM with fold change (Figure 1B), mean $\pm$ SD (Figure 1C to Figure 1E), or percent survival (Figure 1F and Figure 1G). Statistics by one-way analysis of variance (ANOVA) with Tukey's multiple comparisons (Figure 1C to Figure 1E) or log-rank (Mantel-Cox) test (Figure 1F and Figure 1G). Stats in Figure 1F and Figure 1G are shown between 5 μ g RNA-LNP vaccines against other conditions. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Figure 1H depicts a schematic representation of exemplary RNA constructs used in this study. Elements encoded in the representation include a signal peptide (or secretion signal), exemplary modified putative N-glycosylation sites (*), percent GC in an open reading frame and a polyA tail. TcdA,

Toxin A; TcdB, Toxin B; PPEP-1, Pro-Pro endopeptidase-1. Figure 1I shows a table summarizing exemplary LNP compositions used in this study. DSPC, distearoyl-sn-glycero-3-phosphocholine; PEG-DMG, dimyristoyl-rac-glycero-3-methoxypolyethyleneglycol. Figure 1J shows a table summarizing LNP physicochemical parameters, acceptance criteria, and batch to batch consistency of exemplary LNP compositions as described herein. Figure 1K depicts results of example experiments wherein Neuro2a cells were transfected with RNA-LNP (LNP), RNA, or untreated (NT). 24 hours post-transfection cells were lysed, and protein expression was detected via Western Blot with primary antibodies against TcdA and TcdB. PPEP-1 was detected using a polyclonal serum from immunized mice. Figure 1L depicts the results of example experiments demonstrating viability of human umbilical vein endothelial cells (HUVEC), kidney epithelial cells (VERO), colorectal adenocarcinoma cells (CACO-2), and primary muscle cells transfected with Luciferase (Luc), TcdA, or TcdB RNA-LNP or treated with recombinant TcdA or TcdB. Dotted line represents 80% toxicity. Data represented as mean $\pm$ SD. Statistics comparing treatments to untransfected by one-way ANOVA with Dunnett's multiple comparisons test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Figure 2A through Figure 2D depict the results of example experiments demonstrating potency of humoral, cellular, and polyfunctional immune response elicited by an exemplary TcdA/TcdB/PPEP-1 RNA-LNP. Mice were immunized twice with 1 μ g trivalent RNA-LNPs. Splenocytes were isolated two weeks after the immunization and stimulated with either peptide pools (Figure 2A and Figure 2B) or TcdA recombinant protein (Figure 2C and Figure 2D). Peptide-responsive CD8⁺ (Figure 2A and Figure 2C) and CD4⁺ (Figure 2B and Figure 2D) T cells were assessed after stimulation with peptide pools for 6 hours, followed by stimulation with brefeldin A, monensin, and anti-CD107a, via intracellular cytokine staining. Cell frequencies were measured using flow cytometry. $n = 3-5$ mice per group. Data represented as mean $\pm$ SD. Statistics by one-way ANOVA with Tukey's multiple comparisons test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Figure 3A and Figure 3B depict example experiments showing dose dependent antigen-specific B cell responses in dLN to an exemplary TcdA/TcdB/PPEP-1 RNA-LNP. C57Bl/6J mice were immunized once with either 1 μ g (low) or 5 μ g (high) of

an exemplary bivalent RNA-LNP, an exemplary trivalent RNA-LNP, a trivalent recombinant protein with alum adjuvant, or 5 µg of luciferase RNA-LNP. Draining lymph nodes were collected two weeks after the immunization or from age matched unvaccinated controls (naïve), and (Figure 3A) total germinal center (GC) B cells and (Figure 3B) TcdA- (left), TcdB- (middle), and PPEP-1- (right) specific B cells were measured by flow cytometry and calculated per 10,000 cells. n = 5 mice per group, 2 independent experiments. Data represented as mean ± SD. Statistics by one-way ANOVA with Tukey's multiple comparisons test. * p <0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

Figure 4A through Figure 4D depict results from example experiments demonstrating Tfh and B cell subset gating and antigen probe staining. Figure 4A depicts gating of Tfh cells as singlets, live, TCRβ+B220-CD4+CD8- CD62L-PD-1+CXCR5+. Figure 4B depicts gating of spleen isotype-switched B cells as singlets, live, dumpnegative, CD19+B220+, and IgD-IgM-. Figure 4C depicts antigen probe staining for TcdA, TcdB, and PPEP-1 on isotype switched splenic B cells from naïve (top row) and trivalent LNP-treated (bottom row) mice. Figure 4D depicts Gating of germinal center B cells within isotype switched B cells in the spleen.

Figure 5A and Figure 5B depict results from example experiments, wherein naïve mice were injected i.p. with recombinant TcdA or TcdB treated with serum from unvaccinated (control) or monovalent mRNA-LNP vaccinated mice and monitored for survival. n = 5 mice per group. Data are represented as percent live animals. Red line represents LD100. Statistical analysis in by Log-rank (Mantel-Cox) test. *p <0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

Figure 6A through Figure 6E depict results from example experiments demonstrating exemplary TcdA/TcdB/PPEP-1 RNA-LNPs do not induce changes to the intestinal microbiota composition before or after antibiotic treatment. Stool samples were collected from mice that were immunized with bivalent or trivalent RNA-LNPs or unimmunized mice for 16S rRNA gene sequencing prior to immunization (d-35, pre-vaccine), two weeks post-immunization (d-7, post-vaccine), and one-week post-cefoperazone (abx) treatment (d0, post-antibiotics). Figure 6A depicts a schematic of experimental design. Mice were immunized twice with 1µg of bivalent or trivalent RNA-

LNPs. Seven days after the second immunization, mice were treated with cefoperazone in their drinking water for 5 days, followed by a 2-day recovery period on normal water.

Figure 6B depicts the Shannon diversity measurement of fecal samples taken from different vaccine timepoints. Figure 6C depicts a principal component analysis at the

5 operational taxonomic unit level. Each dot represents a unique sample from within each experimental group (color) taken at unique timepoints following vaccination (shape). Figure 6D depicts Shannon diversity measurement of fecal samples collected post-antibiotic treatment. Figure 6E depicts a principal component analysis at the operational taxonomic unit level. Each dot represents a unique sample from within each experimental

10 group (color) taken post-antibiotic treatment.

Figure 7 depicts results from example experiments demonstrating exemplary TcdA/TcdB/PPEP-1 RNA-LNPs do not induce changes to the intestinal microbiota composition before or after antibiotic treatment. The relative abundance of other Clostridia spp. separated by the vaccine timepoint at which the sample was

15 collected are shown. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Number of mice per group: untreated ($n=10$), antibiotics only ($n=28$), bivalent ($n=25$) and trivalent ($n=25$).

Figure 8A and Figure 8B depict the results of example experiments demonstrating the phylum composition of intestinal microbiota following RNA-LNP vaccination. Mice were immunized twice with 1 μ g of bivalent or trivalent RNA-LNPs.

20 Seven days after the second immunization, mice were treated with cefoperazone (abx) in their drinking water for 5 days, followed by a 2-day recovery period on normal water. Stool samples were collected for 16S rRNA gene sequencing prior to immunization (d-35, pre-vaccine), two weeks post-immunization (d-7, post-vaccine), and one-week post-antibiotics treatment (d0, post-antibiotics). Percent abundance of phylum composition

25 (Figure 8A) pre-vaccine and post-vaccine and (Figure 8B) post-antibiotic treatment, depicted by the colors of each of the bars.

Figure 9A through Figure 9K depict the results of example experiments demonstrating exemplary TcdA/TcdB/PPEP-1 RNA-LNP vaccines protect against lethal CDI and induce mucosal antibody responses. Figure 9A depicts a schematic of

30 experimental design. Mice were immunized i.m. with 1 μ g bivalent (green) or trivalent (purple) RNA-LNPs. Seven days after boost, they were placed on a 5-day course of

cefoperazone (abx) in their drinking water. Seven days after the start of abx, they were challenged with 10,000 spores of *C. difficile* (VPI10463). Mice were monitored for survival (Figure 9B) weight loss (Figure 9C) over the course of infection. Weight loss, animal behavior, and stool consistency was scored from 0-4 and represented as a

5 cumulative clinical sickness score (Figure 9D). Figure 9E depicts Ceca pathology scores on day 2 post-infection. Figure 9F and Figure 9G depict *C. difficile* CFUs (Figure 9F) and *C. difficile* toxin titers (Figure 9G) in stools over time. Dotted lines indicate assay limit of detection (LoD). Figure 9H and Figure 9I depict Mucosal IgA (Figure 9H) and IgG (Figure 9I) titers 2 weeks after boost or infection were measured by ELISA. For *C.d* only,

10 mice received cefoperazone followed by 100,000 spores of *C. difficile* CD196. Vaccine-only mice received two 1 µg trivalent RNA-LNP immunizations. Vaccine + *C.d* mice were immunized then infected with *C. difficile* VPI 10463. *C.d* + vaccine mice were infected with CD196 then immunized twice with 1 µg trivalent RNA-LNP. ND, not detectable. n = 4 to 5 mice per group, two independent experiments. Data are represented

15 as mean ± SD (Figure 9C to Figure 9E), mean ± SEM (Figure 9F to Figure 9G), or box and whiskers (Figure 9H and Figure 9I). Statistics by log-rank (Mantel-Cox) (Figure 9B), two-way ANOVA with Tukey's multiple comparisons (Figure 9C and Figure 9D), one-way ANOVA with Tukey's multiple comparisons (Figure 9E), Mann-Whitney (Figure 9F and Figure 9G), or Kruskal-Wallis multiple comparisons test (Figure 9H and Figure 9I).

20 *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Ceca were harvested on day 2 post-infection, fixed in formalin, and stained with hemoxyl and eosin for blinded scoring by a pathologist. Edema, cellular infiltration and epithelial damage for the cecum and colon was scored from 0–4 and represented as a cumulative pathology score (Figure 9J). Representative images of tissue (Figure 9K).

25 Figure 10A and Figure 10B depict the results of example experiments demonstrating exemplary TcdA/TcdB/PPEP-1 RNA-LNP induce mucosal antibody responses. Data in Figure 10A were analyzed by Kruskal-Wallis multiple comparisons test. Data in Figure 10B were analyzed by Wilcoxon matched pairs signed rank test. *p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001. Color denotes from which group

30 unvaccinated group significantly differs. Data were combined from two independent experiments.

Figure 11A through Figure 11C depict results from example experiments demonstrating exemplary RNA-LNP vaccines against TcdA/TcdB/PPEP-1 elicit antigen-specific antibodies across different mouse genetic backgrounds. Figure 11A shows Balb/c mice immunized IM twice with 1 μ g (low) of an exemplary trivalent RNA-LNP or a trivalent recombinant protein with alum as adjuvant. Serum was collected two weeks after a last immunization or age-matched unvaccinated controls, and TcdA (left), TcdB (middle), and PPEP-1 (right) specific IgG antibodies were measured by endpoint ELISA. Figure 11B shows Balb/c (left) mice immunized IM once with either 1 μ g (low) or 5 μ g (high) of an exemplary trivalent RNA-LNP. Serum was collected two weeks after a last immunization or age-matched unvaccinated controls, and TcdA, TcdB, and PPEP-1 specific IgG antibodies were measured by endpoint ELISA. Symbols represent individual mice. Brackets indicated fold change. Figure 11C depicts results from example experiments showing increased valency does not affect antigen-specific antibody responses to TcdA/TcdB/PPEP-1. C57BL/6J mice were immunized once with 1 μ g of monovalent TcdA/TcdB/PPEP-1 (blue), bivalent (green), or trivalent (purple) RNA-LNPs. Sera were collected two weeks after the last immunization, and TcdA, TcdB, and PPEP-1 specific IgG antibodies were measured by endpoint ELISA. Data represented as mean $\pm$ SEM. Five mice were included per group. Statistical analysis by Kruskal-Wallis test with Dunn's multiple comparison test (TcdA and TcdB) or Mann Whitney test (PPEP-1). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Figure 12A through Figure 12E depict example experiments showing dose dependent antigen-specific antibody responses to multivalent RNA-LNPs. Figure 12A depicts the results of example experiments wherein C57BL/6J mice were immunized with either exemplary i.m. once with either 1 μ g or 5 μ g of trivalent RNA-LNP. Figure 12B and Figure 12C depict the results of example experiments demonstrating mice were immunized i.m. once (prime, P) or twice (boost, B) with either 1 μ g or 5 μ g of bivalent (Figure 12B) or trivalent (Figure 12C) RNA-LNP vaccines. Sera were collected two weeks after the indicated immunization and TcdA-, TcdB-, and PPEP-1-specific IgG antibodies were measured by endpoint ELISA. $n = 5$ mice per group, 2 independent experiments. Data are represented as mean $\pm$ SEM. Brackets indicate fold change. Statistics by Mann-Whitney test (Figure 12A) or Wilcoxon test (Figure 12B and Figure

12C). P values shown when $p < 0.05$. Figure 12D and Figure 12E depict the results of example experiments wherein Mice were immunized once (Figure 12D) or twice (Figure 12E) with 1 μ g of trivalent RNA-LNP (purple), or trivalent recombinant protein with alum (grey). Sera were collected two weeks after the last immunization or from age-matched unvaccinated controls (white), and TcdA, TcdB, and PPEP-1 specific IgA antibodies were measured by endpoint ELISA. $n = 5-9$ mice per group, 2 independent experiments. Data are represented as mean $\pm$ SEM. Brackets indicate fold change. Statistics by Kruskal-Wallis test with Dunn's multiple comparison test.

Figure 13 depicts the gating strategy for CD4⁺ and CD8⁺ T cell intracellular cytokine staining. Representative flow cytometry plots of unstimulated splenocytes stained.

Figure 14A- Figure 14B depict results from example experiments depicting humoral immune response protects mice from lethal dose 100 (LD₁₀₀) of TcdA and TcdB. Mice were injected i.p. with increasing doses of recombinant TcdA (Figure 14A) or TcdB (Figure 14B) and monitored for survival.

Figure 15A- Figure 15F depict results from example experiments demonstrating one immunization with an exemplary TcdA/TcdB/PPEP-1 RNA-LNP protects mice from lethal CDI. C57BL6/J mice were immunized IM once (prime) or twice (prime + boost) with either 1 μ g of exemplary TcdA/TcdB (bivalent) or 1 μ g of exemplary TcdA/TcdB/PPEP-1 (trivalent), 5 μ g PPEP-1 monovalent (blue) RNA-LNPs or 1 μ g trivalent recombinant protein + alum (grey) RNA-LNP vaccine. Seven days after the second immunization, mice were treated with cefoperazone and challenged with 10,000 spores of *C. difficile* (VPI 10463). All animals were monitored daily for (Figure 15A, Figure 15C, Figure 15E) survival and (Figure 15B, Figure 15D, Figure 15F) weight loss. $n = 5$ mice per group, two independent experiments. Data are represented as percent live animals (Figure 15A, Figure 15C, Figure 15E) or mean $\pm$ SD (Figure 15B, Figure 15D, Figure 15F). Statistics by Log-rank (Mantel-Cox) test (Figure 15A, Figure 15C, Figure 15E), two-way ANOVA (Figure 15B, Figure 15D) or mixed effect model with Tukey's multiple comparisons test (Figure 15F). Color denotes from which group *Cd* + bivalent group significantly differs. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Figure 16A and Figure 16B depict results from example experiments

demonstrating low dose of exemplary TcdA/TcdB RNA-LNP vaccines protect from lethal CDI. Low dose PPEP-1 does not. C57BL6/J mice were immunized IM once with either 1 µg of an exemplary TcdA/TcdB (bivalent) RNA-LNP vaccine, 5 µg of an exemplary TcdA/TcdB (bivalent) RNA-LNP vaccine, or 5 µg of an exemplary PPEP-1 (monovalent) RNA-LNP vaccine and challenged with 10,000 spores of *C. difficile* (VPI 10463). All animals were monitored daily for survival and weight loss. Figure 16A depicts survival as a percentage of live animals per group. Figure 16B depicts weight loss as a percentage of starting weight. Symbols represent mean and error bars represent $\pm$ SD. n=5 mice/group in one independent experiment.

Figure 17A through Figure 17C depict results from example experiments demonstrating exemplary TcdA/TcdB/PPEP-1 RNA-LNP vaccines protect against infection with diverse *C. difficile* strains. C57BL6/J mice were immunized IM with either 1 µg of an exemplary TcdA/TcdB (bivalent) RNA-LNP vaccine or 1 µg of an exemplary TcdA/TcdB/PPEP-1 (trivalent) RNA-LNP vaccine. Unimmunized mice served as negative controls. All animals were challenged with 100,000 spores of *C. difficile* (CD196) and monitored daily for survival, weight loss, behavior, stool consistency. Figure 17A depicts survival as a percentage of live animals per group. Figure 17 depicts weight loss as a percentage of starting weight. Symbols represent mean and error bars represent $\pm$ SD. Figure 17C depicts clinical sickness score measured as a composite score of weight loss, behavior, and stool consistency. n = 5 mice/group. Data are representative of two independent experiments. Color denotes from which group unvaccinated group significantly differs. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

Figure 18A and 18B depict results from example experiments demonstrating the characterization of antigen-specific, polyfunctional T cells. Peptide-responsive CD4⁺ (Figure 18A) and CD8⁺ (Figure 18B) T cells in spleen were assessed by flow cytometry after stimulation with peptide pools for 6h in the presence of brefeldin A, monensin, and anti-CD107a, and subsequent extra and intracellular staining. Boolean gating was used to determine expression of multiple intracellular cytokines. SPICE 6 was used for visualization.

Figure 19A and Figure 19B depict a model of vaccination against *C. difficile* infection. Figure 19A depicts acute *C. difficile* infection in unvaccinated mice. *C.*

difficile toxins target the intestinal epithelium and cause severe pathology and gastrointestinal disease leading to morbidity and mortality. Figure 19B depicts RNA-LNP vaccination against *C. difficile* virulence factors. Immunization with RNA-LNP vaccines encoding TcdA, TcdB, and PPEP-1 elicit systemic and mucosal antigen-specific immune responses. When vaccinated mice are infected with *C. difficile*, anti-toxin IgG and IgA protect mice from disease and inclusion of PPEP-1 as an immunogen improves decolonization of toxigenic *C. difficile* from the gastrointestinal tract.

Figure 20A through Figure 20K depicts the results of example experiments demonstrating RNA-LNP vaccines provide long-term protection against CDI. Figure 20A and Figure 20B depict antigen-specific antibodies in sera (Figure 20A) and feces (Figure 20B) measured by ELISA at 14 and 40+ days after the last immunization with 1 µg of bivalent or trivalent RNA-LNPs. Figure 20C depicts antigen-specific memory B cells in the spleen 40+ days post-immunization measured by flow cytometry. Figure 20D to Figure 20G depict the results of example experiments wherein mice were infected with 10,000 spores of *C. difficile* (VPI 10463) 40 days after last immunization. Figure 20D depicts a schematic of experimental design. Figure 20E depicts survival, Figure 20F depicts weight loss, and Figure 20G depicts clinical sickness monitored over course of infection. Figure 20H to Figure 20K depict the results of example experiments wherein mice were infected with 10,000 *C. difficile* VPI 10463 spores two weeks after the last immunization and re-infected with 10,000 VPI 10463 spores 26 weeks after initial infection. Figure 20H depicts a schematic of experimental design. Figure 20I depicts survival, Figure 20J depicts weight loss, and Figure 20K depicts clinical scores. ND, not detectable. n = 4 to 10 mice per group. Data are represented as box and whiskers (Figure 20A and Figure 20B), mean ± SD (Figure 20C, Figure 20F to Figure 20G, and Figure 20J and Figure 20K), or percent survival (Figure 20E) and (Figure 20I). Statistics by Mann-Whitney (Figure 20A and Figure 20B), log-rank (Mantel-Cox) (Figure 20E and Figure 20I), or two-way ANOVA with Tukey's multiple comparisons test (Figure 20F, Figure 20G, Figure 20J, and Figure 20K). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

Figure 21A through Figure 21J depict the results of example experiments demonstrating RNA-LNP vaccine targeting vegetative and spore proteins protects against

CDI and elicits antibodies in non-human primates. Figure 21A depicts the amino acid sequence identity of CdeM from our RNA construct compared to 137 *C. difficile* strains across five clades. Figure 21B depicts antigen-specific antibodies in sera 14 days after the last immunization with 1 µg of RNA-LNPs measured by ELISA. Figure 21C to Figure 21I depict the results of example experiments wherein mice were immunized twice (i.m.) with 1 µg of tetravalent (red), spore trivalent (TcdA/ TcdB/CdeM) (blue), vegetative trivalent (TcdA/TcdB/PPEP-1) (purple), or CdeM monovalent (orange) RNA-LNPs. Two weeks after the last immunization, mice were treated with cefoperazone and infected with 10,000 *C. difficile* spores (VPI 10463). Figure 21C depicts a survival curve, Figure 21D depicts weight loss, and Figure 21E clinical scores. Figure 21F to Figure 21I depict *C. difficile* CFUs (Figure 21F and Figure 21H) and toxin titers (Figure 21G and Figure 21I) in feces over time. Figure 21J depicts antigen-specific IgG titers in a macaque at baseline (naïve), after prime (day 21), and boost (day 35) immunizations with 200 µg of tetravalent RNA-LNP. ND, not detectable. n = 5 mice per group, two independent experiments. Data represented as mean ± SD (Figure 21A, Figure 21D, and Figure 21E), percent survival (Figure 21C), or mean ± SEM (Figure 21B and Figure 21F-Figure 21I). Data in Figure 21J show fold change compared with control macaque. Statistics by log-rank (Mantel-Cox) (Figure 21C), two-way ANOVA with Tukey's multiple comparisons test (Figure 21D and Figure 21E), or Kruskal-Wallis multiple comparisons test (Figure 21B, and Figure 21F to Figure 21I). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

Figure 22A through Figure 22D depict the results of example experiments demonstrating trivalent RNA-LNP vaccine elicits antigen-specific antibodies in hamsters. Syrian golden hamsters were immunized with 3 or 6µg of trivalent RNA-LNPs (blue), 6µg of trivalent recombinant protein with alum (grey) or 6µg of empty LNP (eLNP, black) Figure 22A depicts the experimental schematic. Figure 22B depicts weight changes after vaccination. Figure 22C depicts serum and Figure 22D depicts mucosal TcdA-, TcdB-, and PPEP-1-specific IgG antibodies were measured by endpoint ELISA. Data are represented as mean ± SD (Figure 22A) or mean ± SEM (Figure 22B and Figure 22C). n = 6 hamsters per group. Statistics by Kruskal-Wallis test with Dunn's multiple comparison test. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

Figure 23 depicts the results of example experiments demonstrating vaccination elicits immune response independent of infection model. Systemic IgG titers 2 weeks after boost or infection. For *C.d* only, mice received cefoperazone followed by 100,000 spores of CD196. Vaccine only mice received two 1µg trivalent RNA-LNP immunizations. Vaccine + *C.d*, mice were immunized then infected with VPI 10463. *C.d* + vaccine, mice were infected with CD196 then immunized twice with 1µg trivalent RNA-LNP. ND, not detectable. N = 4-5 mice per group, 2 independent experiments. Data are represented as box and whiskers.

Figure 24A and Figure 24B depict the results of example experiments demonstrating RNA-LNP vaccination against CdeM elicits antigen specific immune response. Figure 24A depicts the schematic representation of RNA construct used in this study. Elements encoded in the representation include the signal peptide (SP), modified putative N-glycosylation sites (*), and the polyA tail. CdeM, *C. difficile* exosporium morphogenic protein. Figure 24B depicts Mucosal IgA (left) and IgG (right) titers 2 weeks after boost or infection. Vaccine only received two 1µg tetravalent RNA-LNP immunizations. Vaccine + *C.d*, mice were infected with VPI 10463 after immunization with tetravalent RNA-LNP. Connecting line represents paired samples. N = 4-5 mice per group, 2 independent experiments. Significance by Wilcoxon matched-pairs signed rank test.

Figure 25 depicts anti-toxin IgG and IgA titers measured on 8 weeks old female BL6 mice immunized with of 1µg of bivalent TcdA and TcdB CROP RNA-LNP vaccines two weeks before breeding. Serum samples were collected at 14 days post-delivery of the first litter.

Figure 26 depicts the anti-toxin IgG and IgA titers in stomach content measured on pups 2 weeks after birth. The dotted lines show the Antibody titers in serum of dams at the day of stomach content collection.

Figure 27 depicts anti-toxin IgG and IgA titers in feces measured on pups 2 weeks after birth. The dotted line shows the antibody titers in serum of dams at the day of feces collection.

Figure 28 depicts Anti-toxin IgG and IgA titers measured on serum from control animals (dams) and in the stomach and feces of pups from control animals.

Figure 29 depicts the results of example experiments wherein hamsters were injected with 2µg/antigen at days 0 and 21. Serum was collected at days 21 and 36. Endpoint ELISA titers were measured on serum samples.

5

DETAILED DESCRIPTION

C. difficile is a spore-forming anaerobic bacterium that infects the colon, causing a wide range of disorders that vary in severity from mild diarrhea to toxic megacolon and death (H. Dudukgian, et al., *J Gastrointest Surg.* 2010;(14):315-322; C. P. Kelly, J. T. LaMont, *Annu Rev Med.* 1998;(49):375-390; V. K. Viswanathan, et al., *Gut Microbes.* 2010;(1):234-242). In the United States, *C. difficile* is the most commonly reported nosocomial pathogen, and *C. difficile* infection (CDI) is a major public health threat worldwide (F. C. Lessa, et al., *N Engl J Med.* 2015;(372):2369-2370). The Centers for Disease Control and Prevention recently classified *C. difficile* as a threat-level urgent pathogen (CDC, *Antibiotic resistance threats in the United States.* 2019), which is reinforced by high recurrence rates (H. Dudukgian, et al., *J Gastrointest Surg.* 2010;(14):315-322; V. K. Viswanathan, et al., *Gut Microbes.* 2010;(1):234-242; F. C. Lessa, et al., *N Engl J Med.* 2015;(372):2369-2370), increasing CDI-associated healthcare costs (F. C. Lessa, et al., *N Engl J Med.* 2015;(372):2369-2370), on-going emergence of antibiotic-resistance (H. Huang, et al., *Int J Antimicrob Agents.* 2009;(34):516-522; N. Khanafer, et al., *Anaerobe.* 2016;(37):43-48), and marked rise in community-acquired infection (E. Ofori, et al., *J Hosp Infect.* 2018;(99):436-442). These factors underscore the need for the development of novel preventative and therapeutic measures.

The pathogenesis of CDI is primarily driven by the action of two potent toxins, which cause robust damage to the epithelium and lead to a hyperinflammatory immune response (M. C. Abt, et al., *Nat Rev Microbiol.* 2016;(14):609-620; B. Nibbering, et al., *Front Microbiol.* 2021;(12):804949). These large multi-domain toxins, toxin A (TcdA) and toxin B (TcdB), target Rho family small GTPases, leading to inactivation, subsequent actin condensation, and eventual cell death via either apoptosis or necrosis (M. C. Abt, et al., *Nat Rev Microbiol.* 2016;(14):609-620; S. L. Kordus, et al., *Nat Rev Microbiol.* 2022;(20):285-298; W. K. Smits, et al., *Nat Rev Dis Primers.*

2016;(2):16020). While the immune response to CDI is crucial for protection against translocation of microbiota and for mucosal barrier healing, this hyperinflammatory reaction also contributes significantly to disease pathogenesis and can impede recovery (R. E. El Feghaly, et al., *Clin Infect Dis.* 2013;(56):1713-1721; T. S. Steiner, et al., *Clin Diagn Lab Immunol.* 1997;(4):719-722; H. Yu et al., *Clin Vaccine Immunol.* 2017;(24)). Clinical observations in humans and preclinical work suggest that immunological therapies have promise as prevention and treatment options for CDI (J. F. Bruxelle, et al., *Adv Exp Med Biol.* 2018;(1050):197-225). Previous clinical trials have focused on vaccine candidates targeting toxin A and B (M. Henderson, et al., *Vaccines (Basel).* 2017;(5)); however, there remain substantial opportunities for enhancing current vaccine approaches in the treatment and prevention of CDI. For instance, antibodies generated against the secreted toxins will not directly impede *C. difficile* colonization of the intestine or promote clearance, raising concerns about the possibility of persistent asymptomatic carriage and increased transmission of disease (T. V. Riley, et al., *Vaccine.* 2019;(37):7300-7306). To address this challenge, several research groups have explored the potential of cell surface proteins, spore coat proteins and surface layer polysaccharides as vaccine antigens (S. Pechine, et al., *Front Microbiol.* 2018;(9):1009). While multiple immunogens have shown promise in preventing and treating CDI, none have achieved the goal of ensuring long-lasting clearance of *C. difficile*.

20 The unprecedented speed of development and high levels of protection demonstrated in clinical and preclinical studies have firmly established mRNA-LNP vaccines as a leading platform for the advancement of vaccines and therapeutics against emerging infectious diseases, including lethal bacterial infections (E. Kon, et al., *Sci Adv.* 2023;(9):eadg1036). However, there are limited mRNA-LNP vaccines for bacterial
25 pathogens and no mRNA-LNP vaccines against enteric pathogens to date.

 The present disclosure relates to compositions and methods for inducing an immune response against *C. difficile* infection (CDI) in a subject. In some embodiments, the present disclosure provides a composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen. In some embodiments, the present
30 disclosure provides a composition comprising at least one mRNA molecule encoding at least one *C. difficile* antigen.

In some embodiments, at least one *C. difficile* antigen comprises toxin A (TcdA) or a fragment or variant thereof. In some embodiments, a fragment of TcdA is a fragment comprising a receptor binding domain (RBD) including a combined repetitive oligopeptide (CROPs) domain. In some embodiments, the TcdA antigen comprises SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7.

In some embodiments, at least one *C. difficile* antigen comprises toxin B (TcdB), or a fragment or variant thereof. In some embodiments, a fragment of TcdB is a fragment comprising a RBD including a combined repetitive oligopeptide (CROPs) domain. In some embodiments, the TcdB antigen comprises SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14.

In some embodiments, at least one *C. difficile* antigen comprises Pro-Pro endopeptidase 1 (PPEP-1/Zmp1), or a fragment or variant thereof. In some embodiments, the PPEP-1 antigen comprises SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21.

In some embodiments, at least one *C. difficile* antigen comprises exosporium morphogenic protein CdeM (CdeM), or a fragment or variant thereof. In some embodiments, the CdeM antigen comprises SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25.

In some embodiments, at least one *C. difficile* antigen comprises cell surface protein cwp84 (CWP84), or a fragment or variant thereof. In some embodiments, the CWP84 antigen comprises SEQ ID NO:26 or SEQ ID NO:27.

In some embodiments, at least one *C. difficile* antigen comprises cell surface-associated protein CWP66 (CWP66), or a fragment or variant thereof. In some embodiments, the CWP66 comprises SEQ ID NO:28.

In some embodiments, at least one *C. difficile* antigen comprises putative zinc (Zn) transporter ZupT (ZupT), or a fragment or variant thereof. In some embodiments, the ZupT antigen comprises SEQ ID NO:29.

In some embodiments, at least one *C. difficile* antigen comprises bile acid germinant receptor pseudoprotease CspC (CspC), or a fragment or variant thereof. In

some embodiments, the CspC antigen comprises SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32.

In some embodiments, the present disclosure provides a composition comprising a combination of at least two RNA molecules (e.g., a combination of at least two IVT mRNA molecules) encoding a combination of at least two *C. difficile* antigens, or fragments or variants thereof. In some embodiments, the disclosure provides a composition comprising a combination of 2, 3, 4, 5, 6, 7, or 8 of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, or fragments or variants thereof.

In some embodiments, the present disclosure provides a composition comprising a combination of RNA molecules (e.g., a combination of IVT mRNA molecules) encoding a combination of TcdA and TcdB, or fragments or variants thereof. In some embodiments, the present disclosure provides a composition comprising a combination of RNA molecules (e.g., a combination of IVT mRNA molecules) encoding a combination of TcdA, TcdB, and PPEP-1, or fragments or variants thereof

In some embodiments, an RNA molecule is a nucleoside modified RNA or mRNA molecule. In some embodiments, RNA molecules described herein (e.g., nucleoside-modified RNA or mRNA molecules) are encapsulated in one or more lipid nanoparticles (LNPs). In some embodiments, a composition comprises at least one LNP comprising at least one nucleoside-modified RNA or mRNA molecule encoding at least one *C. difficile* antigen (e.g., TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, fragments, or variants thereof, or any combination thereof). In some embodiments, a composition comprises a combination of LNPs comprising a combination of nucleoside-modified RNA or mRNA molecules encoding a combination of at least two *C. difficile* antigens. For example, in some embodiments, a composition comprises a combination of LNPs comprising a combination of nucleoside-modified RNA or mRNA molecules encoding a combination of at least one of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, and CspC, or a fragment or variant thereof, and at least one additional *C. difficile* antigen. In some embodiments, a composition comprises a combination of LNPs comprising a combination of nucleoside-modified RNA or mRNA molecules encoding a combination of at least two of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, and CspC, or a fragment or variant thereof. In some

embodiments, a composition comprises one or more LNP comprising one or more nucleoside-modified RNA molecule encoding a combination of TcdA, TcdB, PPEP-1, CdeM, or fragments or variants thereof. In some embodiments, a composition comprises one or more LNP(s) comprising one or more nucleoside-modified RNA or mRNA molecule encoding a combination of TcdA and TcdB, or fragments or variants thereof. In some embodiments, a composition comprises one or more LNP comprising one or more nucleoside-modified RNA molecule encoding a combination of TcdA, TcdB, and PPEP-1, or fragments or variants thereof.

In some embodiments, a composition comprises a combination of a first LNP molecule comprising a nucleoside-modified RNA molecule encoding TcdA and a second LNP molecule comprising a nucleoside-modified RNA molecule encoding TcdB. In some embodiments, a composition comprises a combination of a first LNP molecule comprising a nucleoside-modified RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof, and a second LNP molecule comprising a nucleoside-modified RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof. In some embodiments, a composition comprises a first LNP molecule comprising a nucleoside-modified RNA molecule comprising SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, and a second LNP molecule comprising a nucleoside-modified RNA molecule comprising SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof.

In some embodiments, a composition comprises a first LNP molecule comprising a nucleoside-modified RNA molecule encoding TcdA, a second LNP molecule comprising a nucleoside-modified RNA molecule encoding TcdB, and a third LNP molecule comprising a nucleoside-modified RNA molecule encoding PPEP-1. In some embodiments, a composition comprises a first LNP molecule comprising a nucleoside-modified RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID

NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof, a second LNP molecule comprising a nucleoside-modified RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof, and a
5 third LNP molecule comprising a nucleoside-modified RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21 or a fragment or variant thereof. In some embodiments, a composition comprises a first LNP molecule comprising a nucleoside-modified RNA molecule comprising SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37,
10 SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, a second LNP molecule comprising a nucleoside-modified RNA molecule comprising SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a
15 fragment or variant thereof, and a third LNP molecule comprising a nucleoside-modified RNA molecule comprising SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66.

In some embodiments, the disclosure relates to methods of treating or
20 preventing a disease or disorder associated with *C. difficile* using the modified *C. difficile* antigens described herein, or compositions comprising the modified *C. difficile* antigens described herein.

Definitions

25 Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs.

As used herein, each of the following terms has the meaning associated with it in this section.

30 The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, “an

element” means one element or more than one element.

“About” as used herein when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of $\pm 20\%$, $\pm 10\%$, $\pm 5\%$, $\pm 1\%$, or $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods.

The term “antibody,” as used herein, refers to an immunoglobulin molecule, which specifically binds with an antigen. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoreactive portions of intact immunoglobulins. Natural antibodies are typically tetramers of immunoglobulin molecules. The term “antibody” as used herein encompasses antibody fragments, which refers to a portion of an intact antibody and refers to the antigenic determining variable regions of an intact antibody. Antibodies or antibody fragments as described herein may exist in a variety of forms including, for example, polyclonal antibodies, monoclonal antibodies, Fab, F(ab)₂, Fab’, F(ab’)₂, and Fv fragments, linear antibodies, scFv antibodies, and multispecific antibodies as well as single chain antibodies and humanized antibodies (Harlow et al., 1999, In: Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY; Harlow et al., 1989, In: Antibodies: A Laboratory Manual, Cold Spring Harbor, New York; Houston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; Bird et al., 1988, Science 242:423-426).

The term “antigen” or “Ag,” as used herein, is defined as a molecule that binds to an antibody or a T cell receptor. Any macromolecule, including virtually all proteins or peptides, can serve as an antigen. Furthermore, antigens can be derived from recombinant or genomic DNA or RNA. Furthermore, one skilled in the art will understand that an antigen need not be encoded solely by a full-length nucleotide sequence of a gene. The present disclosure provides, but is not limited to, the use of partial nucleotide sequences. Moreover, an antigen need not be encoded by a “gene” at all. An antigen can be generated, synthesized, or can be derived from a biological sample. Such a biological sample can include, but is not limited to a tissue sample, a tumor sample, a cell, or a biological fluid.

The term “immunogen,” as used herein, is intended to denote a substance

of matter, which is capable of inducing an immune response in an individual. This immune response may involve either antibody production, or the activation of specific immunogenically-competent cells, or both. In some embodiments, an immunogen elicits a humoral response. In some embodiments, an immunogen elicits a cellular response. In some embodiments, the immune response is an adaptive immune response. Any DNA or RNA, which comprises a nucleotide sequence or a partial nucleotide sequence encoding a protein that elicits an immune response therefore encodes an “immunogen” as that term is used herein. In some embodiments, the immune response significantly engages pathogenic agents that share immunological features with the immunogen. “Immunogen” refers to any substance introduced into the body in order to generate an immune response. That substance can be a physical molecule, such as a protein, or can be encoded by a vector, such as DNA, RNA, or a virus. In some embodiments, a composition described herein is an immunogen. In some embodiments, a RNA molecule described herein is an immunogen. In some embodiments, an RNA molecule described herein encodes an immunogen. In some embodiments, a composition (e.g., a pharmaceutical composition, immunogenic composition, vaccine) described herein comprises an immunogen.

The term “immune response,” as used herein, means a process involving the activation and/or induction of an effector function in, by way of non-limiting examples, a T cell, B cell, natural killer (NK) cell, and/or an antigen-presenting cell (APC). Thus, an immune response, as would be understood by the skilled artisan, includes, but is not limited to, any detectable antigen-specific activation and/or induction of a helper T cell or cytotoxic T cell activity or response, production of antibodies, antigen presenting cell activity or infiltration, macrophage activity or infiltration, neutrophil activity or infiltration, and the like.

The term “immunogenic composition,” as used herein, refers to any molecule that induces an immune response upon administration. An immunogenic composition may comprise an antigen (e.g., a peptide or polypeptide), a nucleic acid encoding an antigen, a cell expressing or presenting an antigen or cellular component, a virus expressing or presenting an antigen or cellular component, or any combination thereof.

The term “vaccine,” as used herein, refers to an immunogenic composition

that provides protective immunity upon inoculation into a subject.

The term “encoding,” as used herein, refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an RNA (e.g., mRNA), to serve as templates for synthesis of other polymers and
5 macromolecules in biological processes having either a defined sequence of nucleotides (e.g., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of RNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical
10 to the RNA sequence and is usually provided in sequence listings, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

The term “vector,” as used herein, refers to a composition of matter which comprises an isolated nucleic acid and which can be used to deliver the isolated nucleic
15 acid to the interior of a cell. Numerous vectors are known in the art including, but not limited to, linear polynucleotides, polynucleotides associated with ionic or amphiphilic compounds, plasmids, and viruses. Thus, the term “vector” includes an autonomously replicating plasmid or a virus. The term should also be construed to include non-plasmid and non-viral compounds which facilitate transfer of nucleic acid into cells, such as, for
20 example, polylysine compounds, liposomes, and the like. Examples of viral vectors include, but are not limited to, adenoviral vectors, adeno-associated virus vectors, retroviral vectors, and the like.

The term “expression vector,” as used herein, refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operatively
25 linked to a nucleotide sequence to be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or in an in vitro expression system. Expression vectors include all those known in the art, such as cosmids, plasmids (e.g., naked or contained in liposomes), RNA, and viruses (e.g., lentiviruses, retroviruses, adenoviruses, and adeno-
30 associated viruses) that incorporate the recombinant polynucleotide.

The terms “homologous” or “identical,” as used herein in reference to

nucleotide or amino acid sequences, refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared X 100. For example, if 6 of 10 of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology. Generally, a comparison is made when two sequences are aligned to give maximum homology.

The terms “substantially homologous” or “substantially identical,” as used herein in reference to amino acid or nucleotide sequences, refers to a sequence having a degree of identity with respect to a second sequence of at least 60%, of at least 65%, of at least 70%, of at least 75%, of at least 80%, of at least 85%, of at least 90%, of at least 91%, of at least 92%, of at least 93%, of at least 94%, of at least 95%, of at least 96%, of at least 97%, of at least 98%, of at least 99%, or of at least 99.5%. The identity between two sequences can be determined by using the BLASTP algorithm for amino acid sequences or the BLASTN algorithm for nucleotide sequences (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894, Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990)).

The term “variant,” as used herein with respect to a nucleic acid, refers (i) a portion or fragment of a referenced nucleotide sequence; (ii) the complement of a referenced nucleotide sequence or portion thereof; (iii) a nucleic acid that is substantially identical to a referenced nucleic acid or the complement thereof; or (iv) a nucleic acid that hybridizes under stringent conditions to the referenced nucleic acid, complement thereof, or a sequence substantially identical thereto. A variant may be a nucleic acid sequence that is substantially identical over the full length of the full gene sequence or a fragment thereof. The nucleic acid sequence may be 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical over the full length of the gene sequence or a fragment thereof.

The term “variant,” as used with respect to a peptide or polypeptide, refers to a peptide or polypeptide that differs in amino acid sequence by the insertion, deletion, or conservative substitution of amino acids, but retain at least one biological activity. Variant may also refer to a protein with an amino acid sequence that is substantially

5 identical to a referenced protein with an amino acid sequence that retains at least one biological activity. A conservative substitution of an amino acid, i.e., replacing an amino acid with a different amino acid of similar properties (e.g., hydrophilicity, degree and distribution of charged regions) is recognized in the art as typically involving a minor change. These minor changes can be identified, in part, by considering the hydrophobic

10 index of amino acids, as understood in the art. (Kyte et al., 1982, J. Mol. Biol. 157:105-132). The hydrophobic index of an amino acid is based on a consideration of its hydrophobicity and charge. It is known in the art that amino acids of similar hydrophobic indexes can be substituted and still retain protein function. In one embodiment, amino acids having hydrophobic indexes of ± 2 are substituted. The hydrophilicity of amino acids

15 can also be used to reveal substitutions that would result in proteins retaining biological function. A consideration of the hydrophilicity of amino acids in the context of a peptide permits calculation of the greatest local average hydrophilicity of that peptide, a useful measure that has been reported to correlate well with antigenicity and immunogenicity. U.S. Patent No. 4,554,101, incorporated fully herein by reference. Substitution of amino

20 acids having similar hydrophilicity values can result in peptides retaining biological activity, for example immunogenicity, as is understood in the art. Substitutions may be performed with amino acids having hydrophilicity values within ± 2 of each other. Both the hydrophobicity index and the hydrophilicity value of amino acids are influenced by the particular side chain of that amino acid. Consistent with that observation, amino acid

25 substitutions that are compatible with biological function are understood to depend on the relative similarity of the amino acids, and particularly the side chains of those amino acids, as revealed by the hydrophobicity, hydrophilicity, charge, size, and other properties. A variant may be an amino acid sequence that is substantially identical over the full length of the amino acid sequence or fragment thereof. The amino acid sequence

30 may be 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical over the full length of the amino

acid sequence or a fragment thereof.

The terms “fragment” or “functional fragment,” as used herein, refer to a fragment of an antigen or a nucleic acid sequence encoding an antigen that, when administered to a subject, provides an increased immune response. Fragments are generally 10 or more amino acids or nucleic acids in length. “Fragment” may mean a polypeptide fragment of an antigen that is capable of eliciting an immune response in a subject. A fragment of an antigen may be 100% identical to the full length except missing at least one amino acid from the N and/or C terminal, in each case with or without signal peptides and/or a methionine at position 1. Fragments may comprise 20% or more, 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more percent of the length of the particular full-length antigen, excluding any heterologous signal peptide added. The fragment may comprise a fragment of a polypeptide that is 95% or more, 96% or more, 97% or more, 98% or more or 99% or more identical to the antigen and additionally comprise an N terminal methionine or heterologous signal peptide which is not included when calculating percent identity. A fragment of a nucleic acid sequence that encodes an antigen may be 100% identical to the full length except missing at least one nucleotide from the 5' and/or 3' end, in each case with or without sequences encoding signal peptides and/or a methionine at position 1. Fragments may comprise 20% or more, 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more percent of the length of the particular full length coding sequence, excluding any heterologous signal peptide added. The fragment may comprise a fragment that encode a polypeptide that is 95% or more, 96% or more, 97% or more, 98% or more or 99% or more identical to the antigen and additionally optionally comprise sequence encoding an N terminal methionine or heterologous signal peptide which is not included when calculating percent identity.

The term “isolated,” as used herein, means (1) altered or removed from the

natural state and/or (2) separated from at least some of the components with which it was associated when initially produced (whether in nature and/or in an experimental setting) and/or otherwise previously associated, and/or (3) designed, produced, prepared, and/or manufactured by the hand of man. In some embodiments, a nucleic acid or a peptide naturally present in a living subject is not “isolated,” but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is “isolated.” An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

In the context of the present disclosure, the following abbreviations for the commonly occurring nucleosides (nucleobase bound to ribose or deoxyribose sugar via N-glycosidic linkage) are used. “A” refers to adenosine, “C” refers to cytidine, “G” refers to guanosine, “T” refers to thymidine, and “U” refers to uridine.

Unless otherwise specified, a “nucleotide sequence encoding an amino acid sequence” includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns. The phrase nucleotide sequence that encodes a protein or an RNA may also include introns to the extent that the nucleotide sequence encoding the protein may in some version contain an intron(s). In addition, the nucleotide sequence may contain modified nucleosides that are capable of being translated by translational machinery in a cell. Exemplary modified nucleosides are described elsewhere herein. For example, an RNA (e.g., an IVT mRNA) where some or all of the uridines have been replaced with pseudouridine, 1-methyl pseudouridine, 5-methyl-uridine or another modified nucleoside, such as those described elsewhere herein. In some embodiments, the nucleotide sequence may contain a sequence where some or all cytosines are replaced with methylated cytidine, or another modified nucleoside, such as those described elsewhere herein.

The term “operably linked,” as used herein, refers to functional linkage between (1) a regulatory sequence and a heterologous nucleic acid sequence resulting in expression of the latter or (2) between two heterologous nucleic acid sequences resulting in expression of both. For example, a first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a

functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA or RNA sequences are contiguous and, where necessary to join two protein coding regions, in the same
5 reading frame.

The term “polynucleotide,” as used herein, is defined as a chain of nucleotides. Furthermore, nucleic acids are polymers of nucleotides. Thus, nucleic acids and polynucleotides as used herein are interchangeable. One skilled in the art has the general knowledge that nucleic acids are polynucleotides, which can be hydrolyzed into
10 the monomeric “nucleotides.” The monomeric nucleotides can be hydrolyzed into nucleosides. As used herein polynucleotides include, but are not limited to, all nucleic acid sequences which are obtained by any means available in the art, including, without limitation, recombinant means, i.e., the cloning of nucleic acid sequences from a recombinant library or a cell genome, using ordinary cloning technology and PCR™, and
15 the like, and by synthetic means.

The term “polyribonucleotide,” as used herein, is defined as a chain of ribonucleotides. Furthermore, nucleic acids are polymers of ribonucleotides. Thus, nucleic acids and polyribonucleotides as used herein are interchangeable. One skilled in the art has the general knowledge that nucleic acids are polyribonucleotides, which can
20 be hydrolyzed into the monomeric “nucleotides.” The monomeric nucleotides can be hydrolyzed into nucleosides. As used herein polynucleotides include, but are not limited to, all nucleic acid sequences which are obtained by any means available in the art, including, without limitation, recombinant means, i.e., the cloning of nucleic acid sequences from a recombinant library or a cell genome, using ordinary cloning
25 technology and PCR™, and the like, and by synthetic means.

The term “nucleoside-modified nucleic acid,” as used herein, refers to a nucleic acid comprising at least one modified nucleoside.

The term “modified nucleoside,” as used herein, refers to a nucleoside with a modification. For example, over one hundred different nucleoside modifications
30 have been identified in RNA (Rozenski, et al., 1999, The RNA Modification Database: 1999 update. Nucl Acids Res 27: 196-197).

The terms “peptide,” “polypeptide,” and “protein” are used interchangeably herein, and refer to a compound comprised of amino acid residues covalently linked by peptide bonds. A protein or peptide must contain at least two amino acids, and no limitation is placed on the maximum number of amino acids that can
5 comprise a protein’s or peptide’s sequence. Polypeptides include any peptide or protein comprising two or more amino acids joined to each other by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types.

10 “Polypeptides” include, for example, biologically active fragments, substantially homologous polypeptides, oligopeptides, homodimers, heterodimers, variants of polypeptides, modified polypeptides, derivatives, analogs, fusion proteins, among others. The polypeptides include natural peptides, recombinant peptides, synthetic peptides, or a combination thereof.

15 The term “promoter,” as used herein, refers to a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a polynucleotide sequence. By way of one non-limiting example, a promoter that is recognized by bacteriophage RNA polymerase and is used to generate the RNA by in vitro transcription.

20 The term “adjuvant,” as used herein, is defined as any molecule to enhance an antigen-specific adaptive immune response.

The term “pseudouridine” refers to $m^1acp^3\Psi$ (1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine), $m^1\Psi$ (1-methylpseudouridine), Ψm (2’-O-methylpseudouridine), m^5D (5-methyldihydrouridine), $m^3\Psi$ (3-methylpseudouridine), a
25 pseudouridine moiety that is not further modified, a monophosphate, diphosphate, or triphosphate of any of the above pseudouridines, or to any other pseudouridine known in the art.

The term “lipid nanoparticle” or “LNP,” as used herein, refers to a particle having at least one dimension on the order of nanometers (e.g., 1-1,000 nm), which
30 includes one or more lipids.

The term “lipid,” as used herein, refers to a group of organic compounds

that are derivatives of fatty acids (e.g., esters) and are generally characterized by being insoluble in water but soluble in many organic solvents. Lipids are usually divided in at least three classes: (1) “simple lipids” which include fats and oils as well as waxes; (2) “compound lipids” which include phospholipids and glycolipids; and (3) “derived lipids”
5 such as steroids.

The term “cationic lipid,” as used herein, refers to a lipid that is cationic or becomes cationic (protonated) as the pH is lowered below the pK of the ionizable group of the lipid, but is progressively more neutral at higher pH values. At pH values below the pK, the lipid is then able to associate with negatively charged nucleic acids. In some
10 embodiments, the cationic lipid comprises a zwitterionic lipid that assumes a positive charge on pH decrease.

The term “neutral lipid,” as used herein, refers to any one of a number of lipid species that exist in either an uncharged or neutral zwitterionic form at physiological pH. Representative neutral lipids include diacylphosphatidylcholines,
15 diacylphosphatidylethanolamines, ceramides, sphingomyelins, dihydro sphingomyelins, cephalins, and cerebroside.

The term “anionic lipid,” as used herein, refers to any lipid that is negatively charged at physiological pH.

The term “polymer conjugated lipid,” as used herein, refers to a molecule
20 comprising both a lipid portion and a polymer portion. An example of a polymer conjugated lipid is a pegylated lipid.

The term “pegylated lipid,” as used herein, refers to a molecule comprising both a lipid portion and a polyethylene glycol portion. Pegylated lipids are known in the art and include
25 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-s- DMG) and the like.

The term “liposome,” as used herein, is a generic term encompassing a variety of single and multilamellar lipid vehicles formed by the generation of enclosed lipid bilayers or aggregates. Liposomes can be characterized as having vesicular
30 structures with a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid layers separated by aqueous medium. They

form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers (Ghosh et al., 1991 Glycobiology 5: 505-10). However, compositions that have different structures in solution than the normal vesicular structure are also encompassed. For example, the lipids may assume a micellar structure or merely exist as nonuniform aggregates of lipid molecules. Also contemplated are lipofectamine-nucleic acid complexes.

The terms “subject,” “patient,” “individual,” and the like are used interchangeably herein, and refer to any animal, or cells thereof whether in vitro or in situ, amenable to the methods described herein. In some non-limiting embodiments, the patient, subject or individual is a mammal, bird, poultry, cattle, pig, horse, sheep, ferret, primate, dog, cat, guinea pig, rabbit, bat, or human.

The term “disease,” as used herein, is a state of health of a subject wherein the subject cannot maintain homeostasis, and wherein if the disease is not ameliorated then the subject’s health continues to deteriorate.

The term “disorder,” as used herein, refers to a state of health in a subject in which the subject is able to maintain homeostasis, but in which the subject’s state of health is less favorable than it would be in the absence of the disorder. Left untreated, a disorder does not necessarily cause a further decrease in the subject’s state of health.

The term “modulating,” as used herein, means mediating a detectable increase or decrease in the level of a response in a subject compared with the level of a response in the subject in the absence of a treatment or compound, and/or compared with the level of a response in an otherwise identical but untreated subject. The term encompasses perturbing and/or affecting a native signal or response thereby mediating a beneficial therapeutic response in a subject, such as a human.

The term “treat,” as used herein in reference to a disease, means to reduce the frequency or severity of at least one sign or symptom of a disease or disorder experienced by a subject.

The term “effective amount” as used herein, means an amount which provides a therapeutic or prophylactic benefit.

The term “therapeutic” as used herein means a treatment and/or

prophylaxis. A therapeutic effect is obtained by suppression, diminution, remission, prevention, or eradication of at least one sign or symptom of a disease or disorder.

The term “therapeutically effective amount” refers to the amount of the subject compound that will elicit the biological or medical response of a tissue, system, or subject that is being sought by the researcher, veterinarian, medical doctor or other clinician. The term “therapeutically effective amount” includes that amount of a compound that, when administered, is sufficient to prevent development of, or alleviate to some extent, one or more of the signs or symptoms of the disorder or disease being treated. The therapeutically effective amount will vary depending on the compound, the disease and its severity and the age, weight, etc., of the subject to be treated.

The term “transfected” or “transformed” or “transduced” as used herein refers to a process by which exogenous nucleic acid is transferred or introduced into the host cell. A “transfected” or “transformed” or “transduced” cell is one which has been transfected, transformed or transduced with exogenous nucleic acid. The cell includes the primary subject cell and its progeny.

The phrase “under transcriptional control” or “operatively linked” with reference to a promoter as used herein means that the promoter is in the correct location and orientation in relation to a polynucleotide to control the initiation of transcription by RNA polymerase and expression of the polynucleotide.

As used herein, “additional ingredients” include, but are not limited to, one or more of the following: excipients; surface active agents; dispersing agents; inert diluents; granulating and disintegrating agents; binding agents; lubricating agents; sweetening agents; flavoring agents; coloring agents; preservatives; physiologically degradable compositions such as gelatin; aqueous vehicles and solvents; oily vehicles and solvents; suspending agents; dispersing or wetting agents; emulsifying agents, demulcents; buffers; salts; thickening agents; fillers; emulsifying agents; antioxidants; antibiotics; antifungal agents; stabilizing agents; and pharmaceutically acceptable polymeric or hydrophobic materials. Other “additional ingredients” which may be included in the pharmaceutical compositions of the present disclosure are known in the art and described, for example in Remington’s Pharmaceutical Sciences (1985, Genaro, ed., Mack Publishing Co., Easton, PA), which is incorporated herein by reference.

Ranges: throughout this disclosure, various aspects of the present disclosure can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the present disclosure. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. This applies regardless of the breadth of the range.

Immunogenic Compositions

Among other things, in some embodiments, the provided compositions for inducing an immune response against *C. difficile* infection (CDI) in a subject. In some embodiments, the provided compositions are immunogenic compositions (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine). In some embodiments, the present disclosure provides an immunogenic composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine) for inducing an immune response against *C. difficile* infection (CDI) in a subject. In some embodiments, an immunogenic composition is a vaccine. In some embodiments, an immunogenic composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine) induces an immune response against *C. difficile* in a cell, tissue or subject. In some embodiments, an immunogenic composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine) induces a broad immune response against multiple strains of *C. difficile* in a cell, tissue or subject. In some embodiments, an immunogenic composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine) induces a protective immune response in a subject. In some embodiments, the immunogenic composition (e.g. an immunotherapeutic composition, pharmaceutical composition, or vaccine) is a composition for human medical use. In some embodiments, the immunogenic composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or

vaccine) is a composition for veterinary use.

In some embodiments, a composition is or comprises a pharmaceutical composition (e.g., an immunotherapeutic composition or vaccine). In some embodiments, a composition is, or comprises, an RNA-LNP composition (e.g., an RNA-LNP immunogenic composition or an RNA-LNP vaccine). In some embodiments, a composition is or comprises an mRNA-LNP composition (e.g., an mRNA-LNP immunogenic composition or an mRNA-LNP vaccine). In some embodiments, one or more RNA molecules encoding one or more *C. difficile* antigen (e.g., TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, fragments, or variants thereof, or any combination thereof) might also be formulated with an adjuvant. In some embodiments, a composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine) may comprise one or more adjuvants. In some embodiments, a composition (e.g., an immunotherapeutic composition, pharmaceutical composition, or vaccine) and its various components, may be prepared and/or administered by any method disclosed herein. In some embodiments, the induction of immunity by the expression of the *C. difficile* antigens are detected by observing an *in vivo* or *in vitro* response of all or any part of the immune system in a host against one or more of the antigens (e.g., TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, fragments, or variants thereof, or any combination thereof).

For example, a method for detecting the induction of cytotoxic T lymphocytes is well known. A foreign substance that enters the living body is presented to T cells and B cells by the action of antigen presenting cells (APCs). Some T cells that respond to the antigen presented by APC in an antigen specific manner differentiate into cytotoxic T cells (also referred to as cytotoxic T lymphocytes or CTLs) due to stimulation by the antigen. These antigen-stimulated cells then proliferate. This process is referred to herein as “activation” of T cells. Therefore, in some embodiments, CTL induction by an epitope of a polypeptide or peptide or combinations thereof is evaluated by presenting an epitope of a polypeptide or peptide or combinations thereof to a T cell by APC, and detecting the induction of CTL. Furthermore, APCs have the effect of activating B cells, CD4⁺ T cells, CD8⁺ T cells, macrophages, eosinophils and NK cells.

The present disclosure provides methods for evaluating the inducing

action of CTL using APCs, including but not limited to dendritic cells (DCs) and peripheral blood mononuclear cells (PBMCs). Detection of T cells having cytotoxic effects against the cells of interest after the contact with APC shows that the epitope of a polypeptide or peptide or combinations thereof has an activity of inducing the cytotoxic T cells. Furthermore, in some embodiments, the induced immune response is examined by measuring IFN-gamma produced and released by CTL in the presence of antigen-presenting cells that carry immobilized peptide or a combination of peptides by visualizing using anti-IFN-gamma antibodies, such as an ELISPOT assay.

In some embodiments, induction of immunity by expression of *C. difficile* antigens is confirmed by observing the induction of antibody production against *C. difficile* antigens. For example, when antibodies against an antigen are induced in a laboratory subject immunized with the composition encoding the antigens, and when antigen-associated pathology is suppressed by those antibodies, the composition is determined to induce immunity.

In some embodiments, specificity of an antibody response induced in a subject includes binding to many regions of a delivered antigen, as well as the induction of neutralization capable antibodies that that prevent infection or reduce disease severity.

In some embodiments, induction of immunity by expression of *C. difficile* antigens is confirmed by observing the induction of T cells, such as CD4+ T cells, CD8+ T cells, or a combination thereof. For example, CD4+ T cells can also lyse target cells, but mainly supply help in the induction of other types of immune responses, including CTL and antibody generation. The type of CD4+ T cell help can be characterized, as Th1, Th2, Th9, Th17, T regulatory (Treg), or T follicular helper (Tfh) cells. Each subtype of CD4+ T cell supplies help to certain types of immune responses. In some embodiments, a composition selectively induces T follicular helper cells, which drive potent antibody responses.

Nucleic Acids

Among other things, the present disclosure provides compositions. In some embodiments, a composition provided herein induces an immune response in a subject. In some embodiments, a composition comprises an RNA molecule. An RNA

molecule is also referred to herein as a polyribonucleotide.

In various embodiments the RNA molecule is an mRNA, a self-replicating RNA, self-amplifying RNA (saRNA), circular RNA, plasmids, siRNA, miRNA, antisense oligonucleotides, a gene editing component (for example, a guide RNA a tracr RNA, sgRNA, an mRNA encoding an RNA-guided nuclease, a gene or base editing protein, a zinc-finger nuclease, a Talen, a CRISPR nuclease, such as Cas9, a DNA molecule to be inserted or serve as a template for repair), and the like, or a combination thereof. In some embodiments, the composition comprises an mRNA molecule. In some embodiments, the composition comprises a self-amplifying RNA molecule. In some
5
10
embodiments, a composition comprises a circular RNA molecule.

In some embodiments, a composition comprises at least one RNA molecule encoding at least one *C. difficile* antigen. Exemplary *C. difficile* antigens that can be included in a composition of the present disclosure include, but are not limited to, TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, and CspC, and fragments or
15
variants thereof. In some embodiments, a composition comprises one or more RNA molecules encoding any combination of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, and CspC, or fragments or variants thereof. In some embodiments, a composition comprises one or more RNA molecules encoding a combination of TcdA, TcdB, PPEP-1, and CdeM, or fragments or variants thereof. In some embodiments, a composition
20
comprises one or more RNA molecules encoding a combination of TcdA and TcdB, or fragments or variants thereof. In some embodiments, a composition comprises one or more RNA molecules encoding a combination of TcdA, TcdB, and PPEP-1, or fragments or variants thereof.

In some embodiments, the composition comprises one or more mRNA
25
molecules encoding at least one *C. difficile* antigen. In some embodiments, one or more mRNA molecule encodes TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, or a fragment or variant thereof. In some embodiments, one or more mRNA molecule encodes a combination of one or more of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, or a fragment or variant thereof. In some
30
embodiments, a composition comprises one or more mRNA molecules encoding a combination of TcdA, TcdB, PPEP-1, and CdeM, or fragments or variants thereof. In

some embodiments, a composition comprises one or more mRNA molecules encoding a combination of TcdA and TcdB, or fragments or variants thereof. In some embodiments, a composition comprises one or more mRNA molecules encoding a combination of TcdA, TcdB, and PPEP-1, or fragments or variants thereof.

5 In some embodiments, the composition comprises one or more in vitro transcribed (IVT) RNA molecules encoding at least one *C. difficile* antigen. In some embodiments, the IVT RNA molecule encodes TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, or a fragment or variant thereof. In some embodiments, the composition comprises a combination of IVT RNA molecules encoding a combination of
10 two or more of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, or a fragments or variants thereof. In some embodiments, the composition comprises a combination of IVT RNA molecules encoding a combination of TcdA, TcdB, PPEP-1, and CdeM or fragments or variants thereof. In some embodiments, the composition comprises a combination of IVT RNA molecules encoding a combination of TcdA and
15 TcdB, or fragments or variants thereof. In some embodiments, the composition comprises a combination of IVT RNA molecules encoding a combination of TcdA, TcdB, and PPEP-1, or fragments or variants thereof.

In some embodiments, an RNA molecule encoding a *C. difficile* antigen further comprises a sequence encoding a tag or signal peptide. In some embodiments, an
20 RNA molecule encoding a *C. difficile* antigen further comprises a sequence encoding a 6XHis tag. In some embodiments, an RNA molecule includes additional sequences that encode a linker or tag sequences that are linked to an antigen by a peptide bond.

In some embodiments, an RNA molecule encoding a *C. difficile* antigen further comprises a sequence encoding at least one secretion signal. In some
25 embodiments, the at least one secretion signal is a modified IL-2 secretion signal MRMQLLLLIALLSLALVTNS (SEQ ID NO:33).

In some embodiments, the RNA molecule includes an additional sequence that encodes a ferritin domain. In some embodiments, the RNA molecule includes an additional sequence that encodes a protease cleavage site.

30 In some embodiments, an RNA molecule encodes a *C. difficile* antigen modified to disrupt at least one putative N-glycosylation site. In some embodiments,

putative N-glycosylation sites are disrupted by substituting an asparagine residue at a predicted N-glycosylation site with a glutamine (N to Q), lysine (N to K), alanine (N to A), or aspartic acid (N to D). Putative N-glycosylation sites that have been modified with an N to Q modification are represented by a “q” in the sequences provided in Table 1. In some embodiments, the “q” is modified to a lysine, alanine or aspartic acid.

In some embodiments, an RNA molecule encoding a *C. difficile* antigen comprises a nucleotide sequence encoding an amino acid sequence as set forth in Table 1.

Table 1: Amino acid sequences of *C. difficile* antigens.

Sequence Name	Sequence	SEQ ID NO:
TcdA antigen (IL-2 leader sequence underlined)	<u>MRMQLLLLI</u> ALSLALVTNSFNLVTGWQTINGKKYYFDINTGALT ALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFYFAPANTQ NNNIEGQAIVYQSKFLTLNGKKYYFDqNSKAVTGWRIINNEK YYFNPNNIAAAGVGLQVIDNNKYYFNPDTAISKGWQTVqGSR YYFDTDTAIAFNGYKTIDGKHFFYFSDCVVKIGVFSTSNNGFE YFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDqNSKAVT GLQTIDSKKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIG VFKGPNGFEYFAPANTDANNIEGQAAILYQNEFLTLNGKKYYF GSDSKAVTGWRIINNKYYFNPNNIAAIIHLCTINNDKYYFS YDGILQNGYITIERNNFYFDANqESKMVTGVFKGPNGFEYFA PANTHNNNIEGQAIVYQNKFLTLNGKKYYFDqDSKAVTGWQ TIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASSTGYTSINGKHFYFNTDGIMQIGVF KGPNGFEYFAPANTDANNIEGQAAILYQNKFLTLNGKKYYFGS DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT qTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFYFAPANT DANNIEGQAIRYQNRFLYLHDNIYFGqNSKAATGWVTIDGN RYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGF EYFAPANTDANNIEGQAIRYQNRFLHLLGKIYFGqNSKAVT GWQTINGKVYFMPDTAMAAAGGLFEIDGVIYFFGVDGVK APGIYG	SEQ ID NO:1
MutTcdA_H pylferitin with modified IL-2 signal	<u>MRMQLLLLI</u> ALSLALVTNSFNLVTGWQTINGKKYYFDINTGALT ALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFYFAPANTQ NNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVTGWRIINNEK YYFNPNNIAAAGVGLQVIDNNKYYFNPDTAISKGWQTVQGSR YYFDTDTAIAFNGYKTIDGKHFFYFSDCVVKIGVFSTSNNGFE YFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVT GLQTIDSKKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIG VFKGPNGFEYFAPANTDANNIEGQAAILYQNEFLTLNGKKYYF GSDSKAVTGWRIINNKYYFNPNNIAAIIHLCTINNDKYYFS YDGILQNGYITIERNNFYFDANQESKMVTGVFKGPNGFEYFA PANTHNNNIEGQAIVYQNKFLTLNGKKYYFDQDSKAVTGW QTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASSTGYTSINGKHFYFNTDGIMQIGVF KGPNGFEYFAPANTDANNIEGQAAILYQNKFLTLNGKKYYFGS	SEQ ID NO:2

	DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT QTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANT DANNIEGQAIRYQNRFLYLHDNIYYFGQNSKAATGWVTIDG NRYYPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGQNSKAV TGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVGDKV APGIYSGDIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFDKILDKIELIGNENHGLYLADQYVKGIASR KSLEVLFGPGS	
MutTcdA_H pylferitin with modified IL-2 signal, and 6XH tag	MRMQLLLLIALLALVTNSFNLVTGWQTINGKKYYFDINTGA ALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFEYFAPANTQ NNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVTGWRIINNEK YYFNPNNIAIAAVGLQVIDNNKYFNPDTAIIISKGWQTVQGSR YYFDTDTAIAFNGYKTIDGKHFYFSDCVVKIGVFSTSNNGFE YFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVT GLQTIDSKKKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIG VFKGPNNGFEYFAPANTDANNIEGQAILEYQNEFLTLNGKKYYF GSDSKAVTGWRIINNNKYFNPNNIAIAIHLCTINNDKYYFS YDGLQNGYITIERNNFYFDANQESKMVTGVFKGPNNGFEYFA PANTHNNNIEGQAIVYQNKFLTLNGKKYYFDQDSKAVTGW QTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDGIMQIGVF KGPNNGFEYFAPANTDANNIEGQAILEYQNKFLTLNGKKYYFGS DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT QTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANT DANNIEGQAIRYQNRFLYLHDNIYYFGQNSKAATGWVTIDG NRYYPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGQNSKAV TGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVGDKV APGIYSGDIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFDKILDKIELIGNENHGLYLADQYVKGIASR KSLEVLFGPGSHHHHH	SEQ ID NO:3
MutTcdA 2XLSKHpylf erritin with modified IL-2 signal	MRMQLLLLIALLALVTNSFNLVTGWQTINGKKYYFDINTGA ALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFEYFAPANTQ NNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVTGWRIINNEK YYFNPNNIAIAAVGLQVIDNNKYFNPDTAIIISKGWQTVQGSR YYFDTDTAIAFNGYKTIDGKHFYFSDCVVKIGVFSTSNNGFE YFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVT GLQTIDSKKKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIG VFKGPNNGFEYFAPANTDANNIEGQAILEYQNEFLTLNGKKYYF GSDSKAVTGWRIINNNKYFNPNNIAIAIHLCTINNDKYYFS YDGLQNGYITIERNNFYFDANQESKMVTGVFKGPNNGFEYFA PANTHNNNIEGQAIVYQNKFLTLNGKKYYFDQDSKAVTGW QTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDGIMQIGVF KGPNNGFEYFAPANTDANNIEGQAILEYQNKFLTLNGKKYYFGS DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT QTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANT DANNIEGQAIRYQNRFLYLHDNIYYFGQNSKAATGWVTIDG NRYYPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG	SEQ ID NO:4

	FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGQNSKAV TGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVGDKV APGIYGGGSGGGGSLSKDIIKLLNEQVNKEMNSSNLYMSMS SWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQL TSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDH ATFNFLQWYVAEQHEEEVLFDKILDKIELIGNENHGLYLADQ YVKGIKSRKSLEVLFGPGS	
MutTcdA 2XLSKHpylf erritin with modified IL-2 signal and 6xH tag	<u>MRMQLLLLIALLSLALVTNS</u> FNLVTGWQTINGKKYYFDINTGA ALTSYKIINGKHFFYNNDGVMQLGVFKGPDGFYFAPANTQ NNNIEGQAIIVYQSKFLTLNGKKYYFDQNSKAVTGWRIINNEK YYFNPNNIAIAAVGLQVIDNNKYFNPDTAIIISKGWQTVQGSR YYFDTDTAIAFNGYKTIDGKHFFYFSDCVVKIGVFSTSNCFE YFAPANTYNNNIEGQAIIVYQSKFLTLNGKKYYFDQNSKAVT GLQTIDSKKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFFYFNTDGIMQIG VFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLNGKKYYF GSDSKAVTGWRIINNKYYFNPNNIAIAIHLCTINNDKYYFS YDGILQNGYITIERNNFYFDANQESKMVTGVFKGPNGFEYFA PANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDQDSKAVTGW QTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASGTYSINGKHFFYFNTDGIMQIGVF KGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFVS DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT QTSIASTGYTIISGKHFFYFNTDGIMQIGVFKGPDGFYFAPANT DANNIEGQAIRYQNRFLYLHDNIYYFGQNSKAATGWVTIDG NRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGQNSKAV TGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVGDKV APGIYGGGSGGGGSLSKDIIKLLNEQVNKEMNSSNLYMSMS SWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQL TSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDH ATFNFLQWYVAEQHEEEVLFDKILDKIELIGNENHGLYLADQ YVKGIKSRKSLEVLFGPGSHHHHHH	SEQ ID NO:5
MutTcdA VRCHpyl ferritin with modified IL-2 signal	<u>MRMQLLLLIALLSLALVTNS</u> FNLVTGWQTINGKKYYFDINTGA ALTSYKIINGKHFFYNNDGVMQLGVFKGPDGFYFAPANTQ NNNIEGQAIIVYQSKFLTLNGKKYYFDQNSKAVTGWRIINNEK YYFNPNNIAIAAVGLQVIDNNKYFNPDTAIIISKGWQTVQGSR YYFDTDTAIAFNGYKTIDGKHFFYFSDCVVKIGVFSTSNCFE YFAPANTYNNNIEGQAIIVYQSKFLTLNGKKYYFDQNSKAVT GLQTIDSKKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFFYFNTDGIMQIG VFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLNGKKYYF GSDSKAVTGWRIINNKYYFNPNNIAIAIHLCTINNDKYYFS YDGILQNGYITIERNNFYFDANQESKMVTGVFKGPNGFEYFA PANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDQDSKAVTGW QTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASGTYSINGKHFFYFNTDGIMQIGVF KGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFVS DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT QTSIASTGYTIISGKHFFYFNTDGIMQIGVFKGPDGFYFAPANT DANNIEGQAIRYQNRFLYLHDNIYYFGQNSKAATGWVTIDG NRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGQNSKAV TGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVGDKV APGIYSGGDIKLLNEQVNKEMQSSNLYMSMSWYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF	SEQ ID NO:6

	EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFDKILDKIELIGNENHGLYLADQYVKGIKSR KSGSLEVLFGPGS	
MutTcdA_VRCHpyl ferritin with modified IL-2 signal and 6xH tag	MRMQLLLLLIALSLALVTNSFNLVTGWQTINGKKYYFDINTGA ALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFYFAPANTQ NNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVTGWRIINNEK YYFNPNNIAIAAVGLQVIDNNKYYFNPDTAISKGWQTVQGSR YYFDTDTAIAFNGYKTIDGKHFFYFSDCVVKIGVFSTSNQFE YFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDQNSKAVT GLQTIDSKYYFNTNTAEAAATGWQTIDGKKYYFNTNTAEAA TGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDQIMQIG VFKGPNQFEYFAPANTDANNIEGQAILYQNEFLTLNGKKYYF GSDSKAVTGWRIINNNKYYFNPNNIAIAIHLCTINNDKYYFS YDGILQNGYITIERNNFYFDANQESKMVTGVFKGPNQFEYFA PANTHNNNIEGQAIVYQNKFLTLNGKKYYFDQDSKAVTGW QTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNLNTAEAAATG WQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDQIMQIGVF KGNQFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFSGS DSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNT QTSIASTGYTIISGKHFFYFNTDQIMQIGVFKGPDGFYFAPANT DANNIEGQAIRYQNRFLYLHDNIYFYGQNSKAATGWVTIDG NRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGQNSKAV TGWQTINGKVYYFMPDTAMAAAGGLFEIDGVYFFGVGDKV APGIYSGGDIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFDKILDKIELIGNENHGLYLADQYVKGIKSR KSGSLEVLFGPGSHHHHHH	SEQ ID NO:7
TcdB antigen	MRMQLLLLLIALSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFqQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFATAVVGWCKDL EDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTI NDKVYFSDSGIIESGVQNIDNIFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIE TGWYDMEqESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIqYTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE	SEQ ID NO:8
MutTcdB_Hp ylferritin with modified IL-2 signal	MRMQLLLLLIALSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFQQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFATAVVGWCKDL EDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTI NDKVYFSDSGIIESGVQNIDNIFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIE TGWYDMEQESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIqYTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQLVISEGSGDIKLLNE QVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYE	SEQ ID NO:9

	HAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ HISESINNIVDHAISKSDHATFNFLQWYVAEQHEEEVLFKDIL DKIELIGNENHGLYLADQYVKGIASRKSLEVL FQGP GS	
MutTcdB_Hp ylferritin with modified IL-2 signal and 6xH tag	MRMQLLLLIALSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFQQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTA VVGWKDL EDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTI NDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIE TGWIYDMEQESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIQTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYFDPDTAQL VISEGSGDIKLLNE QVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYE HAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ HISESINNIVDHAISKSDHATFNFLQWYVAEQHEEEVLFKDIL DKIELIGNENHGLYLADQYVKGIASRKSLEVL FQGP GS HHH	SEQ ID NO:10
MutTcdB_2X LSKHpylferri tin with modified IL-2 signal	MRMQLLLLIALSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFQQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTA VVGWKDL EDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTI NDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIE TGWIYDMEQESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIQTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYFDPDTAQL VISEGGGSGGGGSG LSKDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQI FQKAYEHEQHISESINNIVDHAISKSDHATFNFLQWYVAEQH EEEVLFKDILDKIELIGNENHGLYLADQYVKGIASRKSLEVL FQGP GS	SEQ ID NO:11
MutTcdB_2X LSKHpylferri tin with modified IL-2 signal and 6xH tag	MRMQLLLLIALSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFQQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTA VVGWKDL EDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTI NDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIE TGWIYDMEQESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIQTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYFDPDTAQL VISEGGGSGGGGSG LSKDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQI FQKAYEHEQHISESINNIVDHAISKSDHATFNFLQWYVAEQH EEEVLFKDILDKIELIGNENHGLYLADQYVKGIASRKSLEVL	SEQ ID NO:12

	FQGPESHHHHHH	
MutTcdB_V RCHpylferri- n with modified IL-2 signal	MRMQLLLLIALLSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFQQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTA VVGWKDL EDGSKYYFDEDTAEA YIGLSLINDGQYYFNDDGIMQVGFVTI NDKVFFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYQGA VEYSGLVRVGEDVYYFGETYTIE TGWIYDMEQESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIQTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQL VISESGGDIKLLNE QVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYE HAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ HISESINNIVDHAISKSDHATFNFLQWYVAEQHEEEVLFDIL DKIELIGNENHGLYLADQYVKGIASRKSGSLEVLFGQPGS	SEQ ID NO:13
MutTcdB_V RCHpylferri- n with modified IL-2 signal and 6xH tag	MRMQLLLLIALLSLALVTNSNLITGFVTVGDDKYYFNPINGGA ASIGETIIDDKNYYFQQSGVLQTGVFSTEDGFKYFAPANTLDE NLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF SPETGKAFKGLNQIGDYKYYFNSDGVMMQKGFVSINDNKHYF DDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTA VVGWKDL EDGSKYYFDEDTAEA YIGLSLINDGQYYFNDDGIMQVGFVTI NDKVFFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDG YKYFAPANTVNDNIYQGA VEYSGLVRVGEDVYYFGETYTIE TGWIYDMEQESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVM QIGVFNTPDGFKYFAHQNTLDENFEGESIQTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQL VISESGGDIKLLNE QVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYE HAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ HISESINNIVDHAISKSDHATFNFLQWYVAEQHEEEVLFDIL DKIELIGNENHGLYLADQYVKGIASRKSGSLEVLFGQPGSH HHHHH	SEQ ID NO:14
PPEP-1 antigen	MRMQLLLLIALLSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPGKWEFTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL GNVNYLGVYPPEFFAESFAYYYLNQDTNSKLLSACPQTYSFL QNLAK	SEQ ID NO:15
PPEP-1 Hpylferri- tin with modified IL-2 signal	MRMQLLLLIALLSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPGKWEFTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL GNVNYLGVYPPEFFAESFAYYYLNQDTNSKLLSACPQTYSFL QNLAKGSGDIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAISKSDHATFNFLQWY VAEQHEEEVLFDILDKIELIGNENHGLYLADQYVKGIASR KSLEVLFGQPGS	SEQ ID NO:16
PPEP-1 Hpylferri- tin with modified	MRMQLLLLIALLSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPGKWEFTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL	SEQ ID NO:17

IL-2 signal and 6xH tag	GNVNYLGVPYEEFFAESFAYYYLNQDTNSKLKSACPQTYSFL QNLAKGSGDIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIKSR KSLEVLFGPGSHHHHHH	
PPEP-1_2XLSKHp ylferritin with modified IL-2 signal	<u>MRMQLLLLIALSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY</u> DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPKGWEGTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL GNVNYLGVPYEEFFAESFAYYYLNQDTNSKLKSACPQTYSFL QNLAKGGGSGGGSGLSKDIKLLNEQVNKEMNSSNLYMSM SSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQ LTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKD HATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLAD QYVKGIKSRKSLEVLFGPGS	SEQ ID NO:18
PPEP-1_2XLSKHp ylferritin with modified IL-2 signal and 6xH tag	<u>MRMQLLLLIALSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY</u> DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPKGWEGTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL GNVNYLGVPYEEFFAESFAYYYLNQDTNSKLKSACPQTYSFL QNLAKGGGSGGGSGLSKDIKLLNEQVNKEMNSSNLYMSM SSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQ LTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKD HATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLAD QYVKGIKSRKSLEVLFGPGSHHHHHH	SEQ ID NO:19
PPEP-1_VRCHpylfe ritin with modified IL-2 signal	<u>MRMQLLLLIALSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY</u> DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPKGWEGTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL GNVNYLGVPYEEFFAESFAYYYLNQDTNSKLKSACPQTYSFL QNLAKSGGDIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIKSR KSGSLEVLFGPGS	SEQ ID NO:20
PPEP-1_VRCHpylfe ritin with modified IL-2 signal with 6xH tag	<u>MRMQLLLLIALSLALVTNSDSTTIQQNKDTLSQIVVFPTGNY</u> DKNEANAMVNRLANIDGKYLNALKQNNLKIKLLSGKLTDEK EYAYLKGVVPKGWEGTGKTWDDVPGLGGSTVALRIGFSNK GKGHDAINLELHETAHAIDHIVLNDISKSAQFKQIFAKEGRSL GNVNYLGVPYEEFFAESFAYYYLNQDTNSKLKSACPQTYSFL QNLAKSGGDIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKF EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY VAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIKSR KSGSLEVLFGPGSHHHHHH	SEQ ID NO:21
Wild type (WT) CdeM	MENKKYANGGYSEDWYERGESTAKWFQNDREEYERAAYD EDRERRGSNCGCSDSGGNRPRNCERFRREAEIREREAREAFC ESSERKKEALAYECEARKLWEEAEKYWDEYSKYNYKGIEYL AEAARLFDEGMECEARRNGNNGGNNNNCCCHKCHKCNCNC CRK	SEQ ID NO:22
Modified CdeM (N to	MENKKYANGGYSEDWYERGESTAKWFQNDREEYERAAYD EDRERRGSqCGCSDSGGNRPRNCERFRREAEIREREAREAFCE SSERKKEALAYECEARKLWEEAEKYWDEYSKYqYKGIEYLA EAARLFDEGMECEARRNGNNGGNNNNCCCHKCHKCNCNCCK K	SEQ ID NO:23

Q at position 48 and 115)		
Wild type (WT) CdeM with the modified IL-2 secretion signal	MRMQLLLLLIALSLALVTNSMENKKYANGGYSEDWYERGES TAKWFQNDREEYERAAAYDEDRERRGSNCGCSDSGGNRPRN CERFRREAEIREREAREAFCESSERKKEALAYECEARKLWEE AEKYWDEYSKYNYKGYEYLAEAAARLFDEGMECEARRNGNN GGNNNNCCHKCHKCNCNCCRK	SEQ ID NO:24
Modified CdeM with the modified IL-2 secretion signal	MRMQLLLLLIALSLALVTNSMENKKYANGGYSEDWYERGES TAKWFQNDREEYERAAAYDEDRERRGSqCGCSDSGGNRPRNC ERFRREAEIREREAREAFCESSERKKEALAYECEARKLWEEA EKYWDEYSKYqYKGYEYLAEAAARLFDEGMECEARRNGNNG GNNNNCCHKCHKCNCNCCRK	SEQ ID NO:25
CWP84 or cell surface protein cwp84 [Clostridioide s difficile]	MRKYKSKKLSKLLALLTVCFLIVSTIPVSAENHKTLDGVETA EYSESYLQYLEDVKNGDTAKYNGVIPFPHEMEGTTLRNKGR SSLPSAYKSSVAYNPMDLGLTTPAKNQGSNLTCWSFSGMSTL EAYLKLKGYGTIDLSEEHLRWWATGGKYGWNLDDMSGSS NVTAGYLTAWAGPKLEKDIPYNLKSEAQGATKPSNMDTAP TQFNVTDVVRLNKDKETVKNAIMQYGSVTSGYAHYSTYFN KDETAYNCTNKRAPLNHAVAIVGWDDNYSKDNFASDVKPE SNGAWLVKSSWGEFNSMKGFFWISYEDKTLLTDTDNYAMK SVSKPDSKMYQLEYAGLSKIMSNKVTAANVFDFSRDSEK LDSVMFETDSVGAKYEVYYAPVVNGVPQNNMTKLASGT SYSGYINVPTNSYSLPKGKGAIVVVIDNTANPNREKSTLAYET NIDAYYLYEAKANLGESYILQNNKFEDINTYSEFSPCNFVIKA ITKTSSGQATSGESLTGADRYETAVKVSQKGWTSSQNAVLV NGDAIVDAL TATPFTA AIDSPILLTGKDNLDSKTKAELQRLGT KKVYLIGGENSLSKNVQTQLSNMGISVERISGSDRYKTSISLA QKLNSIKSVSQVAVANGVNLADAI SVGAAAADNNMPIILT NEKSELQGADEFLNSSKITKSYIIGGTATLSSNLESKLSN PTRLAGSNRNETNAKIIDKFYPSSDLKYAFVVKDGSKSQGD LIDGLAVGALGAKTDSPPVVLVGNKLDESQKNVLKSKKIETPIRV GGNGNESAFNELNTLLGK	SEQ ID NO:26
modified CWP84 [Clostridioide s difficile]	MRKYKSKKLSKLLALLTVCFLIVSTIPVSAENHKTLDGVETA EYSESYLQYLEDVKNGDTAKYNGVIPFPHEMEGTTLRNKGR SSLPSAYKSSVAYNPMDLGLTTPAKNQGSNLTCWSFSGMSTL EAYLKLKGYGTIDLSEEHLRWWATGGKYGWNLDDMSGSSq VTAIGYLTAWAGPKLEKDIPYNLKSEAQGATKPSNMDTAPT QFqVTDVVRLNKDKETVKNAIMQYGSVTSGYAHYSTYFNKD ETAYqCTNKRAPLNHAVAIVGWDDqYSKDNFASDVKPE SNGAWLVKSSWGEFNSMKGFFWISYEDKTLLTDTDNYAMKSVS KPDSDKMYQLEYAGLSKIMSNKVTAANVFDFSRDSEKLD SVMFETDSVGAKYEVYYAPVVNGVPQNNMTKLASGTVSYS GYINVPTNSYSLPKGKGAIVVVIDNTANPNREKSTLAYETNID AYYLYEAKANLGESYILQNNKFEDINTYSEFSPCNFVIKAITK TSSGQATSGESLTGADRYETAVKVSQKGWTSSQNAVLVNGD AIVDAL TATPFTA AIDSPILLTGKDNLDSKTKAELQRLGT KKVYLIGGENSLSKNVQTQLSNMGISVERISGSDRYKTSISLAQKL NSIKSVSQVAVANGVNLADAI SVGAAAADNNMPIILTNEKS ELQGADEFLNSSKITKSYIIGGTATLSSNLESKLSNPTRLAGSN RNETNAKIIDKFYPSSDLKYAFVVKDGSKSQGD LIDGLAVGA	SEQ ID NO:27

	LGAKTDSPPVVLVGNKLDESQKNVLKSKKIETPIRVGGNGqES AFNELNTLLGK	
Modified CWP66	MNISKQIVSLLTMTFLTVTLYGNTSNASTKDTLTGSGRWETA IKISQAGWTKSESAVLVNDNSIADALSATPFAKAKDAPILLTQ SNKLDSRTKAELKRLGVKNVYLIGGSIALSSEIEKQLNAENIN FERISGNSRYDTSKLAEKLDREKSISKIVVVNGEKGLADAVS VGAIAAQENMPIILSDSENGTEVADNFIDSKDIAKSYVIGGT SISNSVERSLPqATRIAGSSRSETNAKIIIEFYKDTDIKNIYVTK DGTKNKNDLIDSLAVGVLAANKSSPIVLAGNKLDTTQKDVL NTKIIDKVTQIGGLGNENVVEDILDIQEETKYTVETIDELNAAI KRADANDIIKFKPEKEKTINNSFSIETKKTVTIELDGRYRQTIT LDIPNGKFNNYAEIEGGVKLKNIKNESLVNKGSIQDLDIYDEN GCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSSTIIR NSGNIDTVTGKKEPAISGNKPKVNDTEKETKAARGLNPRVEA CSVPKKDYVMITIPNSPKDSRYKIYYRVVYNKPYAMDVGDKI NIGEWTVAPTDEEPFLEKAKNGCYVEAVEVNTSTKEVSRWG RTqATDDGF	SEQ ID NO:28
ZupT [Clostridioide s difficile]	MNNVIFAFLITLLAGLSTGIGSCIAFFAKKTNTKFLSISLGFS GVMIYVSMIEIFPKAQDALTKSMGEKLGSWSTVIAFFIGMAII AVIDKLIPQEENPHEIKKMEINIEEKNIKKNSLLRTGIFTAMAI AIHNFPEGLATFISALDDVTIAPIAIAIAIHNIPEGISVSVVYY ATGDKKKAFYYSFSLSGMSEPLGAIIGYVLLRNFLNDITLGIVF AIVGGIMVFISLDELLPSAREYGEHLSIYGLIAGMGVMAISL LLFK	SEQ ID NO:29
bile acid germinant receptor pseudoprotea se CspC	MEKSYCHYQGDIESALQENGINRYMVLNSQLAVIYVPLDFDE TILNNIIQVAWWESEPMSSLIEITNNVNNGETITTAAETDIY ENPYNDITGRGILLVIDSGIDYLHPDFINDDGTSLVLYLWDQ EANTNPPPEGFIGSEFTRSELNIAINRNDGSLSQDNIGTGTLV SGILVGNGRINSQYRGITTESDLIVVKLKSYPDITYYAGRINYS VSDFLAAITYVTNIARTENKPLIINLTIGVKSSAVATTSILDTFN ILSSAGVVVVSAGAGNQGNTDIHYSGRFSSVGEVQDVIIQDGD DYALDITLNTNGPDKVGAQIISPSGEVSHDIRYSPDFYIYRGKF NLENTTYAMRFIYPYITSGKENLEIRLRDIKPGVWILRLTSELII NGEYDIYLPNKNLIAPDTRFLDPDSVATITMYAASDDVITVGT FNNKTDSMWIGSSKGPIRGRGIKPDIVASGVDIISTYKNGTYN TGTGTGVSSSIVTGVLALLMEYLEKQDNVPRLSLFTQVLKTY LILGATKLEIYTYPNVSQGYGILNLKNTIQQIANTL	SEQ ID NO:30
Modified CspC	MEKSYCHYQGDIESALQENGINRYMVLNSQLAVIYVPLDFDE TILNNIIQVAWWESEPMSSLIEITNNVNNGETITTAAETDIY ENPYNDITGRGILLVIDSGIDYLHPDFINDDGTSLVLYLWDQ EANTNPPPEGFIGSEFTRSELNIAINRNDGSLSQDNIGTGTLVS GILVGNGRINSQYRGITTESDLIVVKLKSYPDITYYAGRIqYSVS DFLAAITYVTNIARTENKPLIIqLTIGVKSSAVATTSILDTFNILSS AGVVVVSAGAGNQGNTDIHYSGRFSSVGEVQDVIIQDGDYAL LDITLNTNGPDKVGAQIISPSGEVSHDIRYSPDFYIYRGKFNLE NTTYAMRFIYPYITSGKENLEIRLRDIKPGVWILRLTSELIIINGE YDIYLPNKNLIAPDTRFLDPDSVATITMYAASDDVITVGTFFNN KTDSMWIGSSKGPIRGRGIKPDIVASGVDIISTYKNGTYNTGT GTGVSSSIVTGVLALLMEYLEKQDNVPRLSLFTQVLKTYLIL GATKLEIYTYPNVSQGYGILNLKNTIQQIANTL	SEQ ID NO:31
Modified CspC with IL-2 signal	MRMQLLLIALSLALVTNSEKSYCHYQGDIESALQENGINRY MVLNSQLAVIYVPLDFDETLNIIQVAWWESEPMSSLIEITN NVNNGETITTAAETDIYENPYNDITGRGILLVIDSGIDYLHP DFINDDGTSLVLYLWDQEANTNPPPEGFIGSEFTRSELNIAIN RNDGSLSQDNIGTGTLVSGILVGNGRINSQYRGITTESDLIVV KLKSYPDITYYAGRIqYSVSDFLAAITYVTNIARTENKPLIIqLTI	SEQ ID NO:32

	GVKSSAVATTSLDTFNLSSAGVVVVSGAGNQGNTDIHYSGRFSSVGEVQDVIIQDGGDYALDITLNTNGPDKVGAQIISPSGEVSHDIRYSPDFYIYRGKFNLENTTYAMRFIYPYITSGKENLEIRLRDIKPGVWILRLTSELIINGEYDIYLPNKNLIAPDTRFLDPDSVATITMYAASDDVITVGTFFNNKTDMSWIGSSKGPIRGRGIKPDIVASGVDIISTYKNGTYNTGTGTGVSSSIVTGVLALLMEYLEKQDNVPRLSLFTQVLKTYLILGATKLEIYTPNVVSQGYGILNLKNTIQQIANTL	
--	---	--

In some embodiments, an RNA encoding a TcdA antigen encodes SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a TcdA antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7. In some embodiments, the variant of the TcdA antigen comprises a modification of at least one putative N-glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7. In various embodiments, the “q” can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a TcdA antigen encodes at least a fragment of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7.

In some embodiments, an RNA encoding a TcdB antigen encodes SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a TcdB antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14. In some embodiments, the variant of the TcdB antigen comprises a modification of at least one putative N-

glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14. In various embodiments, the “q” can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a TcdB antigen encodes at least a fragment of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14.

In some embodiments, an RNA encoding a PPEP-1 antigen encodes SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a PPEP-1 antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21. In some embodiments, the variant of the PPEP-1 antigen comprises a modification of at least one putative N-glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21. In various embodiments, the “q” can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a PPEP-1 antigen encodes at least a fragment of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21.

In some embodiments, an RNA encoding a CdeM antigen encodes SEQ

ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25 or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a CdeM antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25. In some
 5 embodiments, the variant of the CdeM antigen comprises a modification of at least one putative N-glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25. In various embodiments, the
 10 “q” can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a CdeM antigen encodes at least a fragment of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:22, SEQ
 15 ID NO:23, SEQ ID NO:24 or SEQ ID NO:25.

In some embodiments, an RNA encoding a CWP84 antigen encodes SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a CWP84 antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%,
 20 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:26 or SEQ ID NO:27. In some embodiments, the variant of the CWP84 antigen comprises a modification of at least one putative N-glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:26 or SEQ ID NO:27. In various embodiments, the “q”
 25 can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a CWP84 antigen encodes at least a fragment of SEQ ID NO:26 or SEQ ID NO:27 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:26 or SEQ ID NO:27. In some embodiments, the
 30 RNA encodes a CWP84 antigen operably linked to an IL-2 leader sequence (SEQ ID NO:33). Therefore, in some embodiments, the RNA encodes SEQ ID NO:26 or SEQ ID

NO:27 operably linked to SEQ ID NO:33.

In some embodiments, an RNA encoding a CWP66 antigen encodes SEQ ID NO:28, or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a CWP66 antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:28. In some embodiments, the variant of the CWP66 antigen comprises a modification of at least one putative N-glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:28. In various embodiments, the “q” can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a CWP66 antigen encodes at least a fragment of SEQ ID NO:28 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:28. In some embodiments, the RNA encodes a CWP66 antigen operably linked to an IL-2 leader sequence (SEQ ID NO:33). Therefore, in some embodiments, the RNA encodes SEQ ID NO:28 operably linked to SEQ ID NO:33.

In some embodiments, an RNA encoding a ZupT antigen encodes SEQ ID NO:29, or a fragment or variant thereof. In some embodiments, a nucleic acid molecule encoding a ZupT antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:29. In some embodiments, the variant of the ZupT antigen comprises a modification of at least one putative N-glycosylation site. In some embodiments, an RNA molecule encoding a ZupT antigen encodes at least a fragment of SEQ ID NO:29 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:29. In some embodiments, the RNA encodes a ZupT antigen operably linked to an IL-2 leader sequence (SEQ ID NO:33). Therefore, in some embodiments, the RNA encodes SEQ ID NO:29 operably linked to SEQ ID NO:33.

In some embodiments, an RNA encoding a CspC antigen encodes SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof. In some

embodiments, a nucleic acid molecule encoding a CspC antigen encodes an amino acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32. In some embodiments, the variant of the CspC antigen

5 comprises a modification of at least one putative N-glycosylation site. Exemplary putative N-glycosylation sites that can be modified include, but are not limited to, those indicated by a lower case “q” in SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32. In various embodiments, the “q” can be modified to an alternative amino acid residue such as “K”, “D” or “A.” In some embodiments, an RNA molecule encoding a CspC antigen

10 encodes at least a fragment of SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32 having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, of the full length of SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32.

In some embodiments, an RNA molecule comprises a 5' UTR, a 3'UTR, and/or a

15 polyA tail. Exemplary sequences of RNA molecules encoding the *C. difficile* antigens including 5' UTR, 3'UTR and a polyA tail are provided in Table 2, with the 5' UTR and 3'UTR sequences indicated with underlining.

In some embodiments, an RNA molecule encoding a *C. difficile* antigen comprises a nucleotide sequence as set forth in Table 2.

20

Table 2. Nucleotide sequences of *C. difficile* antigens

Sequence Name	Sequence	SEQ ID NO:
TcdA RNA coding sequence (IL-2 coding sequence in bold)	<u>AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCC</u> <u>UGGCCUGGUGACCAACUCCU</u> UUAACCGUGACCGGC UGGCAGACCAUCAACGGCAAGAAGUACUACUUCGACAUC AACACCGGCGCCGCCUGACCUCCUACAAGAUAUCAAC GGCAAGCACUUCUACUUAACAACGACGGCGUGAUGCAG CUGGGCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUC GCCCCGCAACACCCAGAACAACAACUACGAGGGCCAG GCCAUCGUGUACCAGUCCAAGUUCUGACCCUGAACGGC AAGAAGUACUACUUCGACCAGAACUCCAAGGCCGUGACC GGCUGGCGCAUCAUCAACAACGAGAAGUACUACUUAAC CCCAACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUC GACAACAACAAGUACUACUUAACCCCGACACCGCCAUC AUCUCCAAGGGCUGGCAGACCGUGCAGGGCUCGCGCUAC UACUUCGACACCGACACCGCCAUCGCCUUAACGGCUAC	SEQ ID NO:34

	AAGACCAUCGACGGCAAGCACUUCUACUUCGACUCCGAC UGCGUGGUGAAGAUCGGCGUGUUCUCCACCUCCAACGGC UUCGAGUACUUCGCCCCCGCCAACACCUACAACAACAAC AUCGAGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUG ACCCUGAACGGCAAGAAGUACUACUUCGACCAGAACUCC AAGGCCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAG UACUACUUAACACCAACACCGCCGAGGCCGCCACCGGC UGGCAGACCAUCGACGGCAAGAAGUACUACUUAACACC AACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGAC GGCAAGAAGUACUACUUAACACCAACACCGCCAUCGCC UCCACCGGCUACACCAUCAUCAACGGCAAGCACUUCUAC UUCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAG GGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACC GACGCCAACAAUCGAGGGCCAGGCCAUCCUGUACCAG AACGAGUUCUGACCCUGAACGGCAAGAAGUACUACUUC GGCUCCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUC AACAAACAAGAAGUACUACUUAACCCCAACAACGCCAUC GCCGCCAUCCACUGUGCACCAUCAACAACGACAAGUAC UACUUCUCCUACGACGGCAUCCUGCAGAACGGCUACAUC ACCAUCGAGCGCAACAACUUCUACUUCGACGCCAACAG GAGUCCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAAC GGCUUCGAGUACUUCGCCCCCGCCAACACCCACAACAAC AACAUUCGAGGGCCAGGCCAUCGUGUACCAGAACAAGUUC CUGACCCUGAACGGCAAGAAGUACUACUUCGACCAGGAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAG AAGUACUACUUAACCUGAACACCGCCGAGGCCGCCACC GGCUGGCAGACCAUCGACGGCAAGAAGUACUACUUAAC CUGAACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUC GACGGCAAGAAGUACUACUUAACACCAACACCUUCAUC GCCUCCACCGGCUACACCUCCAUCAACGGCAAGCACUUC UACUUAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUC AAGGGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAAC ACCGACGCCAACAAUCAUCGAGGGCCAGGCCAUCCUGUAC CAGAACAAGUUCUGACCCUGAACGGCAAGAAGUACUAC UUCGGCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACC AUCGACGGCAAGAAGUACUACUUAACACCAACACCGCC GUGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAG UACUACUUAACACCCAGACCUCCAUCGCCUCCACCGGC UACACCAUCAUCUCCGGCAAGCACUUCUACUUAACACC GACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGAC GGCUUCGAGUACUUCGCCCCCGCCAACACCGACGCCAAC AACAUUCGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUC CUGUACCUGCACGACAACAUCUACUACUUCGGCCAGAAC UCCAAGGCCGCCACCGGCUGGUGACCAUCGACGGCAAC CGCUACUACUUCGAGGCCAACACCGCCAUGGGCGCCAAC GGCUACAAGACCAUCGACAACAAGAACUUCUACUUCGCG AACGGCCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAAC GGCUUCGAGUACUUCGCCCCCGCCAACACCGACGCCAAC AACAUUCGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUC CUGCACCUGCUGGGCAAGAUCUACUACUUCGGCCAGAAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAG GUGUACUACUUAUGCCCGACACCGCCAUGGCCGCCGCC GGCGGCCUGUUCGAGAUCCGACGGCGUGAUCAUUCUUC GGCGUGGACGGCGUGAAGGCCCCCGGCAUCUACGGCUAA UAA	
--	---	--

TcdA RNA: (5'UTR, coding sequence. 3'UTR and polyA; IL-2 coding sequence in bold)	AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU UUAUUUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC CCUGGUGACCAACUCCUUCAACCUGGUGACCGGCUGGC AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA CCGGCGCCGCCUGACCUCUACAAGAUAUCAACGGCA AGCACUUCUACUUAACAACGACGGCGUGAUGCAGCUGG GCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUCGCCC CCGCCAACACCCAGAACAACAACAUUCGAGGGCCAGGCCA UCGUGUACCAGUCCAAGUUCUGACCCUGAACGGCAAGA AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU GGCGCAUCAUCAACAACGAGAAGUACUACUUAACCCCA ACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUCGACA ACAACAAGUACUACUUAACCCCGACACCGCCAUCAUCU CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG UGGUGAAGAUCGGCGUGUUCUCCACCUCACGGCUUCG AGUACUUCGCCCCCGCCAACACCUACAACAACAACUUCG AGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUGACCC UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAG CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU ACUUAACACCAACACCGCCGAGGCCGCCACCGGCUGGC AGACCAUCGACGGCAAGAAGUACUACUUAACACCAACA CCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACGGCA AGAAGUACUACUUAACACCAACACCGCCAUCGCCUCCA CCGGCUACACCAUCAUAACGGCAAGCACUUCUACUUA ACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCC CCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCGACG CCAACAACAUUCGAGGGCCAGGCCAUCCUGUACCAGAACG AGUUCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU CCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUCAACA ACAAGAAGUACUACUUAACCCCAACAACGCCAUCGCCG CCAUCCACCUGUGCACCACCAUCAACAACGACAAGUACU UCUCCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA UCGAGCGCAACAACUUCUACUUCGACGCCAACCAGGAGU CCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAACGGCU UCGAGUACUUCGCCCCCGCCAACACCCACAACAACAACA UCGAGGGCCAGGCCAUCGUGUACCAGAACAAGUUCUGA CCCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAGAAGU ACUACUUAACCUGAACACCGCCGAGGCCGCCACCGGCU GGCAGACCAUCGACGGCAAGAAGUACUACUUAACCUGA ACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACG GCAAGAAGUACUACUUAACACCAACACCUUAUCGCCU CCACCGGCUACACCUCCAUAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGG GCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAGA ACAAGUUCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACCAUCG ACGGCAAGAAGUACUACUUAACACCAACACCGCCGUGG CCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAGUACU ACUUAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC	SEQ ID NO: 35
---	--	---------------

	CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGUA CCUGCACGACAACUACUACUUCGGCCAGAACUCCAA GGCCGCCACCGGCUGGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGCA CCUGCUGGGCAAGAUACUACUUCGGCCAGAACUCCAA GGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGGUGUA CUACUUCUAGCCCCGACACCGCCAUGGGCGCCGCCGGCGG CCUGUUCGAGAUCCGACGGCGUGAUACUUCUUCGGCGU GGACGGCGUGAAGGGCCCCGGCAUCUACGGCUAAUAAAC <u>UAGUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAG</u> <u>CCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAUU</u> <u>ACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA</u> <u>GCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUC</u> <u>ACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdA_H pylferitin mRNA with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGAUCGCCCCUGUCCC UGGCCCCUGGUGACCAACUCCUUAACCGUGACCGGC UGGCAGACCAUCAACGGCAAGAAGUACUACUUCGACAUC AACACCGGCGCCGCCUGACCUCCUACAAGAUCAUCAAC GGCAAGCACUUCUACUUAACAACGACGGCGUGAUGCAG CUGGGCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUC GCCCCCGCCAACACCCAGAACAACAACAUCCGAGGGCCAG GCCAUCCGUGUACCAGUCCAAGUUCUGACCCUGAACGGC AAGAAGUACUACUUCGACCAGAACUCCAAGGCCGUGACC GGCUGGGCGCAUCAUCAACAACGAGAAGUACUACUUAAC CCCAACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUC GACAACAACAAGUACUACUUAACCCCGACACCGCCAUC AUCUCCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUAC UACUUCGACACCGACACCGCCAUCGCCUUAACGGCUAC AAGACCAUCGACGGCAAGCACUUCUACUUCGACUCCGAC UGCGUGGUGAAGAUCGGCGUGUUCUCCACCUCUCCAACGGC UUCGAGUACUUCGCCCCCGCCAACACCUACAACAACAAC AUCGAGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUG ACCCUGAACGGCAAGAAGUACUACUUCGACCAGAACUCC AAGGCCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAG UACUACUUAACACCAACACCGCCGAGGCCGCCACCGGC UGGCAGACCAUCGACGGCAAGAAGUACUACUUAACACC AACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGAC GGCAAGAAGUACUACUUAACACCAACACCGCCAUCGCC UCCACCGGCUACACCAUCAUAACGGCAAGCACUUCUAC UUAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAG GGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACC GACGCCAACAACAUCGAGGGGCCAGGCCAUCCUGUACCAG AACGAGUUCUGACCCUGAACGGCAAGAAGUACUACUUC GGCUCCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUC AACAACAAGAAGUACUACUUAACCCCAACAACGCCAUC GCCGCCAUCCACCUGUGCACCAUCAACAACGACAAGUAC	SEQ ID NO: 36

	UACUUCUCCUACGACGGCAUCCUGCAGAACGGCUACAUC ACCAUCGAGCGCAACAACUUCUACUUCGACGCCAACAG GAGUCCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAAC GGCUUCGAGUACUUCGCCCCGCCAACCCACAACAAC AACAUUCGAGGGCCAGGCCAUCGUGUACCAGAACAAGUUC CUGACCCUGAACGGCAAGAAGUACUACUUCGACCAGGAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAG AAGUACUACUUAACCUGAACACCGCCGAGGCCGCCACC GGCUGGCAGACCAUCGACGGCAAGAAGUACUACUUAAC CUGAACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUC GACGGCAAGAAGUACUACUUAACACCAACACCUUCAUC GCCUCCACCGGCUACACCUCCAUAACGGCAAGCACUUC UACUUAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUC AAGGGCCCCAACGGCUUCGAGUACUUCGCCCCGCCAAC ACCGACGCCAACACAUCGAGGGGCCAGGCCAUCCUGUAC CAGAACAAGUCCUGACCCUGAACGGCAAGAAGUACUAC UUCGGCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACC AUCGACGGCAAGAAGUACUACUUAACACCAACACCGCC GUGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAG UACUACUUAACACCCAGACCUCCAUCGCCUCCACCGGC UACACCAUCAUCUCCGGCAAGCACUUCUACUUAACACC GACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGAC GGCUUCGAGUACUUCGCCCCGCCAACACCGACGCCAAC AACAUUCGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUC CUGUACCUGCACGACAACAUCUACUACUUCGGCCAGAAC UCCAAGGCCGCCACCGGCUGGGUGACCAUCGACGGCAAC CGCUACUACUUCGAGCCCCAACACCGCCAUGGGCGCCAAC GGCUACAAGACCAUCGACAACAAGAACUUCUACUUCGCG AACGGCCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAAC GGCUUCGAGUACUUCGCCCCGCCAACACCGACGCCAAC AACAUUCGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUC CUGCACCUGCUGGGCAAGAUCUACUACUUCGGCCAGAAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAG GUGUACUACUUAUGCCCCGACACCGCCAUGGCCGCCGCC GGCGGCCUGUUCGAGAUCCGACGGCGUGAUCUACUUCUUC GGCGUGGACGGCGUGAAGGGCCCCGGCAUCUACGGCUCC GGCGACAUCAUCAAGCUGCUGAACGAGCAGGUGAACAAG GAGAUGAACUCCUCCAACCUGUACAUGUCCAUGUCCUCC UGGUGCUACACCCACUCCUGGACGGCGCCGGCCUGUUC CUGUUCGACCACGCCGCCGAGGAGUACGAGCACGCCAAG AAGCUGAUCAUCUCCUGAACGAGAACAACGUGCCCGUG CAGCUGACCUCCAUCUCCGCCCCGAGCACAAGUUCGAG GGCCUGACCCAGAUCCUCCAGAAGGCCUACGAGCACGAG CAGCACAUCCGAGUCCAUCAACAACAUCCGUGGACCAC GCCAUCAAGUCCAAGGACCACGCCACCUUCAAACUCCUG CAGUGGUACGUGGCCGAGCAGCACGAGGAGGAGGUGCUG UUCAAGGACAUCUGGACAAGAUCGAGCUGAUCGGCAAC GAGAACCACGGCCUGUACCUGGCCGACCAGUACGUGAAG GGCAUCGCCAAGUCCCGCAAGUCCUGGAGGUGCUGUUC CAGGGCCCCGGCUCCUAAUAA	
MutTcdA_H pyl ferritin mRNA with modified IL-2 signal	AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA AUCUCAAGCAAUCAAGCAUUCUACUUAUUGCAGCAAU UUAAAUCAUUUUAUUUAAAGCAAAAGCAAUUUUCUGAA AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC	SEQ ID NO: 37

	CCUGGUGACCAACUCCUUCAACCUGGUGACCGGCUGGC AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA CCGGCGCCGCCUGACCUCUACAAGAUAUCAACGGCA AGCACUUCUACUUCAACAACGACGGCGUGAUGCAGCUGG GCGUGUUCAAGGGCCCCGACGGCUUCGAGUACUUCGCCC CCGCCAACACCCAGAACAACAUAUCGAGGGCCAGGCCA UCGUGUACCAGUCCAAGUUCUGACCCUGAACGGCAAGA AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU GGCGCAUCAUCAACAACGAGAAGUACUACUUCAACCCCA ACAACGCCAU CGCCGCCGUGGGCCUGCAGGUGAUCGACA ACAACAAGUACUACUUCAACCCCGACACCGCCAUAUCU CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG UGGUGAAGAUCGGCGUGUUCUCCACCUGCAACGGCUUCG AGUACUUCGCCCCCGCCAACACCUACAACAACAACUUCG AGGGCCAGGCCAU CGUGUACCAGUCCAAGUUCUGACCC UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAGG CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU ACUUCAACACCAACACCGCCGAGGCGCCACCGGCUGGC AGACCAUCGACGGCAAGAAGUACUACUUCAACACCAACA CCGCCGAGGGCCGCCACCGGCUGGCAGACCAUCGACGGCA AGAAGUACUACUUCAACACCAACACCGCCAUCGCCUCCA CCGGCUACACCAUCAUCAACGGCAAGCACUUCUACUUA ACACCGACGGCAUCAUGCAGAU CGGCGUGUUAAGGGCC CCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCGACG CCAACAACAU CGAGGGCCAGGCCAUCCUGUACCAGAACG AGUUCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU CCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUCAACA ACAAGAAGUACUACUUCAACCCCAACAACGCCAUCGCCG CCAUCCACCUGUGCACCAUCAACAACGACAAGUACUACU UCUCCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA UCGAGCGCAACAACUUCUACUUCGACGCCAACCAGGAGU CCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAACGGCU UCGAGUACUUCGCCCCCGCCAACACCCACAACAACAACA UCGAGGGCCAGGCCAUCGUGUACCAGAACAAGUUCUGA CCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAGAAGU ACUACUUCAACCUGAACACCGCCGAGGCGCCACCGGCU GGCAGACCAUCGACGGCAAGAAGUACUACUUCAACCUGA ACACCGCCGAGGCGGCCACCGGCUGGCAGACCAUCGACG GCAAGAAGUACUACUUCAACACCAACACCUUCAUCGCCU CCACCGGCUACACCUCCAUAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAU CGGCGUGUUAAGG GCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAGA ACAAGUUCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACCAUCG ACGGCAAGAAGUACUACUUCAACACCAACACCGCCGUGG CCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAGUACU ACUUCAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAU CGGCGUGUUAAGGGCCCCGACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCUGUA CCUGCACGACAACAUCUACUACUUCGGCCAGAACUCCAA	
--	--	--

	GGCCGCCACCGGCUGGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGCA CCUGCUGGGCAAGAUACUACUUCGCGCCAGAACUCCAA GGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGGUGUA CUACUUCAUUGCCCGACACCGCCAUGGCCGCCGCCGGCGG CCUGUUCGAGAUCCGACGGCGUGAUACUUCUUCGCGGU GGACGGCGUGAAGGCCCGGCAUCUACGGCUCCGGCGA CAUCAUCAAGCUGCUGAACGAGCAGGUGAACAAAGGAGAU GAACUCCUCCAACCUGUACAUGUCCAUGUCCUCCUGGUG CUACACCCACUCCUGGACGGCGCCGGCCUGUUCUGUU CGACCACGCCGCCGAGGAGUACGAGCACGCCAAGAAGCU GAUCAUCUUCUGAACGAGAACAACGUGCCCGUGCAGCU GACCUCCAUCUCCGCCCGGAGCACAAAGUUCGAGGGCCU GACCCAGAUUCUCCAGAAGGCCUACGAGCACGAGCAGCA CAUCUCCGAGUCCAUCAACAACAUUGUGGACCACGCCAU CAAGUCCAAGGACCACGCCACCUUCAACUUCUGCAGUG GUACGUGGCCGAGCAGCACGAGGAGGAGGUGCUGUCAA GGACAUCUGGACAAGAUCGAGCUGAUCGGCAACGAGAA CCACGGCCUGUACCUGGCCGACCAGUACGUGAAGGGCAU CGCCAAGUCCCGCAAGUCCUGGAGGUGCUGUUCAGGG CCCCGGCUCCUAAUAAACUAGUAGUGACUGACUAGGAUC <u>UGGUUACCACUAAACCAGCCUCAAGAACCCCGAAUGGA</u> <u>GUCUCUAAGCUACAUAAUACCAACUACACUUAACAAAAU</u> <u>GUUGUCCCCCAAAUGUAGCCAUUCGUUUCUGCUCCUAA</u> <u>UAAAAAGAAAGUUUCUUCACAUUCUAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAA</u>	
MutTcdA_H pylferitin mRNA with modified IL-2 signal with 6xH tag	<u>AGCAUAAAAGUCUCAACACAACAUUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUUAUUGCAGCAAU</u> <u>UUAUAUCAAUUCUUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> <u>GCAUGCAGCUGCUGCUGCUGAUCGCCUUGUCCUUGGC</u> <u>CCUGGUGACCAACUCCUUAACCCUGGUGACCGGCUGGC</u> AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA CCGGCGCCGCCUGACCUCCUACAAGAUCAUCAACGGCA AGCACUUCUACUUAACAACGACGGCGUGAUGCAGCUGG GCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUCGCC CCGCCAACACCCAGAACAACAACAUUGAGGGCCAGGCCA UCGUGUACCAGUCCAAGUUCUGACCCUGAACGGCAAGA AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU GGCGCAUCAUCAACAACGAGAAGUACUACUUAACCCCA ACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUCGACA ACAACAAGUACUACUUAACCCCGACACCGCCAUCUACU CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG UGGUGAAGAUCGGCGUGUUCUCCACCUCACCGGCUUCG AGUACUUCGCCCCGCCAACACCUACAACAACAACUUCG AGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUGACCC UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAGG	SEQ ID NO: 38

	CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU ACUUCAACACCAACACCGCCGAGGCCGCCACCGGCUGGC AGACCAUCGACGGCAAGAAGUACUACUUAACACCAACA CCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACGGCA AGAAGUACUACUUAACACCAACACCGCCAUCGCCUCCA CCGGCUACACCAUCAUAACGGCAAGCACUUCUACUUA ACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCC CCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCGACG CCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAGAACG AGUUCCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU CCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUCAACA ACAAGAAGUACUACUUAACCCCAACAACGCCAUCGCCG CCAUCCACCUGUGCACCAUCAACAACGACAAGUACUACU UCUCCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA UCGAGCGCAACAACUUCUACUUCGACGCCAACCAGGAGU CCAAGAUGGUGACCGGGCGUGUUAAGGGCCCCAACGGCU UCGAGUACUUCGCCCCCGCCAACACCCACAACAACAACA UCGAGGGGCCAGGCCAUCCUGUACCAGAACAAGUCCUGA CCCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAGAAGU ACUACUUAACCUGAACACCGCCGAGGCCGCCACCGGC GGCAGACCAUCGACGGCAAGAAGUACUACUUAACCUGA ACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACG GCAAGAAGUACUACUUAACACCAACACCUUCAUCGCCU CCACCGGCUACACCUCCAUAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGG GCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAGA ACAAGUCCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACCAUCG ACGGCAAGAAGUACUACUUAACACCAACACCGCCGUGG CCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAGUACU ACUUCAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGACGGCU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGCA CCUGCUGGGCAAGAUUCUACUUCGCGCCAGAACUCCAA GGCCGCGCACCGGCUGGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGCA CCUGCUGGGCAAGAUUCUACUUCGCGCCAGAACUCCAA GGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGGUGUA CUACUUCUAGCCCCGACACCGCCAUGGCCGCCGCCGGCGG CCUGUUCGAGAUCCGACGGCGUGAUCUACUUCUUCGGCGU GGACGGCGUGAAGGCCCCCCGGCAUCUACGGCUCCGGCGA CAUCAUCAAGCUGCUGAACGAGCAGGUGAACAAGGAGAU GAACUCCUCCAACCUGUACAUGUCCAUGUCCUCCUGGUG CUACACCCACUCCUGGACGGCGCCGGCCUGUUCUGU CGACCACGCCGCCGAGGAGUACGAGCACGCCAAGAAGCU GAUCAUCUCCUGAACGAGAACAACGUGCCCGUGCAGCU GACCUCCAUCUCCGCCCCCGAGCACAAAGUUCGAGGGCCU GACCCAGAUUCUCCAGAAGGCCUACGAGCACGAGCAGCA	
--	---	--

	CAUCUCCGAGUCCAUCAACAACAUCGUGGACCACGCCAU CAAGUCCAAGGACCACGCCACCUUCAACUCCUGCAGUG GUACGUGGCCGAGCAGCAGGAGGAGGAGGUGCUGUCAA GGACAUCCUGGACAAGAUCGAGCUGAUCGGCAACGAGAA CCACGGCCUGUACCUGGCCGACCAGUACGUGAAGGGCAU CGCCAAGUCCC GCAAGUCCCUGGAGGUGCUGUCCAGGG CCCCGGCUCCC ACCACCACCACCACCACU AAUAAACUAG UAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCU CAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAUACC <u>AAUUACACUUAACAAAAUGUUGUCCCCCAAAUGUAGCC</u> <u>AUUCGUACUGCUCCUAAUAAAAAGAAAGUUUCUUCACA</u> <u>UUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdA 2XLSKHpylf erritin mRNA with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCU UCAACCUGGUGACCGGC UGGCAGACCAUCAACGGCAAGAAGUACUACUUCGACAUC AACACCGGCGCCGCCUGACCUCUACAAGAUCAUAAC GGCAAGCACUUCUACUUAACAACGACGGCGUGAUGCAG CUGGGCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUC GCCCCCGCCAACACCCAGAACAACAUCGAGGGGCCAUC GCCAUCGUGUACCAGUCCAAGUUCUGACCCUGAACGGC AAGAAGUACUACUUCGACCAGAACUCCAAGGCCGUGACC GGCUGGCGCAUCAUCAACAACGAGAAGUACUACUUAAC CCCAACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUC GACAACAACAAGUACUACUUAACCCCGACACCGCCAUC AUCUCCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUAC UACUUCGACACCGACACCGCCAUCGCCUUAACGGCUAC AAGACCAUCGACGGCAAGCACUUCUACUUCGACUCCGAC UGCGUGGUGAAGAUCGGCGUGUUCUCCACCUCCAACGGC UUCGAGUACUUCGCCCCCGCCAACACCUACAACAACAAC AUCGAGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUG ACCCUGAACGGCAAGAAGUACUACUUCGACCAGAACUCC AAGGCCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAG UACUACUUAACACCAACACCGCCGAGGCCGCCACCGGC UGGCAGACCAUCGACGGCAAGAAGUACUACUUAACACC AACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGAC GGCAAGAAGUACUACUUAACACCAACACCGCCAUCGCC UCCACCGGCUACACCAUCAUAACGGCAAGCACUUCUAC UUAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAG GGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACC GACGCCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAG AACGAGUUCUGACCCUGAACGGCAAGAAGUACUACUUC GGCUCCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUC AACAACAAGAAGUACUACUUAACCCCAACAACGCCAUC GCCGCCAUCCACUGUGCACCAUCAACAACGACAAGUAC UACUUCUCCUACGACGGCAUCCUGCAGAACGGCUACAUC ACCAUCGAGCGCAACAACUUCUACUUCGACGCCAACAG GAGUCCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAAC GGCUUCGAGUACUUCGCCCCCGCCAACACCCACAACAAC AACAUCCGAGGGCCAGGCCAUCGUGUACCAGAACAAGUUC CUGACCCUGAACGGCAAGAAGUACUACUUCGACCAGGAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAG AAGUACUACUUAACUGAACACCGCCGAGGCCGCCACC GGCUGGCAGACCAUCGACGGCAAGAAGUACUACUUAAC	SEQ ID NO:39

	CUGAACACCCGACCGCCGAGGCCGCGCCAGGCCGAGGCAUUC GACGGCAAGAAGUACUACUUAACACCAACACCUUCAUC GCCUCCACCGGCUACACCUCCAUAACGGCAAGCACUUC UACUUCAACACCGACGGCAUCAUGCAGAU CGGCGUGUUC AAGGGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAAC ACCGACGCCAACAAUCAUCGAGGGCCAGGCCAUCCUGUAC CAGAACAAGUUCUGACCCUGAACGGCAAGAAGUACUAC UUCGGCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACC AUCGACGGCAAGAAGUACUACUUAACACCAACACCGCC GUGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAG UACUACUUAACACCCAGACCUCCAUCGCCUCCACCGGC UACACCAUACUUCGCCGCAAGCACUUCUACUUAACACC GACGGCAUCAUCGAGAU CGGCGUGUUAAGGGCCCCGAC GGCUUCGAGUACUUCGCCCCCGCCAACACCGACGCCAAC AACAU CGAGGGCCAGGCCAUCCGCUACCAGAACC GCUUC CUGUACCUGCACGACAACAUCUACUACUUCGGCCAGAAC UCCAAGGCCGCCACCGGCUGGGUGACCAUCGACGGCAAC CGCUACUACUUCGAGCCCCAACACCGCCAUGGGCGCCAAC GGCUACAAGACCAUCGACAACAAGAACUUCUACUUCGCG AACGGCCUGCCCCAGAU CGGCGUGUUAAGGGCUCCAAC GGCUUCGAGUACUUCGCCCCCGCCAACACCGACGCCAAC AACAU CGAGGGCCAGGCCAUCCGCUACCAGAACC GCUUC CUGCACCUGCUGGGCAAGAU CUACUACUUCGGCCAGAAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAG GUGUACUACUUAUGCCCCGACACCGCCAUGGCCGCCGCC GGCGGCGUGUUCGAGAU CGGCGUGAU CUACUUCUUC GGCGUGGACGGCGUGAAGGGCCCCGGCAUCUACGGCGGC GGCUCCGGCGGGCGGGCGGCUCCGGCCUGUCCAAGGACAUC AUCAAGCUGCUGAACGAGCAGGUGAACAAGGAGAU GAAC UCCUCCAACCUGUACAUGUCCAUGUCCUCCUGGUGCUAC ACCCACUCCCGUGACGGCGCCGGCCUGUCCUGUUCGAC CACGCCGCCGAGGAGUACGAGCACGCCAAGAAGCUGAUC AUCUUCUGAACGAGAACAACGUGCCCGUGCAGCUGACC UCCAUCUCCGCCCCCGAGCACAAGUUCGAGGGCCUGACC CAGAU CUUCCAGAAGGCCUACGAGCACGAGCAGCACAUC UCCGAGUCCAUCAACAACAUCGUGGACCACGCCAUCAAG UCCAAGGACCACGCCACCUUCAACUCCUGCAGUGGUAC GUGGCCGAGCAGCACGAGGAGGAGGUGCUGUUAAGGAC AUCCUGGACAAGAU CGAGCUGAUCGGCAACGAGAACCAC GGCCUGUACCGUGGCCGACCAGUACGUGAAGGGCAUCGCC AAGUCCCGCAAGUCCCUUGGAGGUGCUGUCCAGGGCCCC GGCUCCUAAUAA	
MutTcdA 2XLSKHpylf erritin mRNA with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC CCUGGUGACCAACUCCUUAACCUGGUGACCGGCUGGC AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA CCGGCGCCGCCUGACCUCUACAAGAUAUCAACGGCA AGCACUUCUACUUAACAACGACGGCGUGAUGCAGCUGG GCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUCGCC CCGCCAACCCAGAACAAACAUCGAGGGCCAGGCCA UCGUGUACCAAGUCCAAGUCCUGACCCUGAACGGCAAGA AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU GGCGCAUCAUCAACAACGAGAAGUACUACUUAACCCCA	SEQ ID NO:40

ACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUCGACA ACAACAAGUACUACUUAACCCCGACACCGCCAUCUACU CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG UGGUGAAGAUCCGGCGUGUUCUCCACCUCACCGGCUUCG AGUACUUCGCCCCCGCCAACACCUACAACAACAACUACG AGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUGACCC UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAGG CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU ACUUAACACCAACACCGCCGAGGGCCGCCACCGGCUGGC AGACCAUCGACGGCAAGAAGUACUACUUAACACCAACA CCGCCGAGGGCCGCCACCGGCUGGCAGACCAUCGACGGCA AGAAGUACUACUUAACACCAACACCGCCAUCGCCUCCA CCGGCUACACCAUCAUAACGGCAAGCACUUCUACUUA ACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCC CCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCGACG CCAACAACAUCCGAGGGGCCAGGCCAUCCUGUACCAGAACG AGUUCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU CCGACUCCAAGGGCGUGACCGGCUGGCGCAUCAUAACA ACAAGAAGUACUACUUAACCCCAACAACGCCAUCGCCG CCAUCCACCGUGGCACCAUCAACAACGACAAGUACUACU UCUCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA UCGAGCGCAACAACUUCUACUUCGACGCCAACCAGGAGU CCAAGAUGGUGACCGGGCGUGUUAAGGGCCCCAACGGCU UCGAGUACUUCGCCCCCGCCAACACCCACAACAACAACA UCGAGGGGCCAGGCCAUCCUGUACCAGAACAAGUUCUGA CCCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAGAAGU ACUACUUAACCGUGAACACCGCCGAGGGCCGCCACCGGCU GGCAGACCAUCGACGGCAAGAAGUACUACUUAACCGUGA ACACCGCCGAGGGCCGCCACCGGCUGGCAGACCAUCGACG GCAAGAAGUACUACUUAACACCAACACCUUAUCGCCU CCACCGGCUACACCUCCAUCAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGG GCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAACAUCGAGGGGCCAGGCCAUCCUGUACCAGA ACAAGUUCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCCGACUCCAAGGGCCGUGACCGGGCUUGCGCACCAUCG ACGGCAAGAAGUACUACUUAACACCAACACCGCCGUGG CCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAGUACU ACUUAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCCUGUA CCUGCACGACAACAUCUACUACUUCGGCCAGAACUCCAA GGCCGCCACCGGCUGGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCCUGCA CCUGCUGGGCAAGAUCUACUACUUCGGCCAGAACUCCAA GGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGGUGUA CUACUUAUGCCCGACACCGCCAUGGCCGCCGCCGGCGG	
--	--

	CCUGUUCGAGAUCGACGGCGUGAUCUACUUCUUCGGCGU GGACGGCGUGAAGGCCCGGCAUCUACGGCGGCGGCUC CGGCGGCGGCGGCUCGCGGCCUGUCCAAGGACAUCAUCAA GCUGCUGAACGAGCAGGUGAACAAGGAGAUGAACUCCUC CAACCUGUACAUGUCCAUGUCCUCCUGGUGCUACACCCA CUCCUGGACGGCGCCGGCCUGUUCUGUUCGACCACGC CGCCGAGGAGUACGAGCACGCCAAGAAGCUGAUCAUUU CCUGAACGAGAACAACGUGCCCGUGCAGCUGACCUCU CUCCGCCCCGAGCACAAGUUCGAGGGCCUGACCCAGAU CUUCCAGAAGGCCUACGAGCACGAGCAGCACAUCUCCGA GUCCAUCAACAACUUGGACCACGCCAUCAAGUCCAA GGACCACGCCACCUUCAACUCCUGCAGUGGUACGUGGC CGAGCAGCACGAGGAGGAGGUGCUGUUAAGGACAUCU GGACAAGAUUCGAGCUGAUCGGCAACGAGAACCACGGCCU GUACCUGGCCGACCAGUACGUGAAGGGCAUCGCCAAGUC CCGCAAGUCCUGGAGGUGCUGUUCAGGGCCCCGGCUC CUAAUAAACUAGUAGUGACUGACUAGGAUCUGGUUACCA <u>CUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAG</u> <u>CUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCC</u> <u>CAAAUGUAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAA</u> <u>AGUUUCUUCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAA</u>	
MutTcdA 2XLSKHpylf erritin mRNA with modified IL-2 signal and 6xH tag	<u>AGCAUAAAAGUCUCAACACAACAUUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUUAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC CCUGGUGACCAACUCCUUAACCUGGUGACCGGCUGGC AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA CCGGCGCCGCCUGACCUCUACAAGAUAUCAACGGCA AGCACUUCUACUUAACAACGACGGCGUGAUGCAGCUGG GCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUCGCC CCGCCAACACCCAGAACAACAACUUCGAGGGCCAGGCCA UCGUGUACAGUCCAAGUCCUGACCCUGAACGGCAAGA AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU GGCGCAUCAUCAACAACGAGAAGUACUACUUAACCCCA ACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUCGACA ACAACAAGUACUACUUAACCCCGACACCGCCAUAUCU CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG UGGUGAAGAUCGGCGUGUUCUCCACCUCACCGGCUUCG AGUACUUCGCCCCGCCAACACCUACAACAACAACUUCG AGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUGACCC UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAGG CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU ACUUAACACCAACACCGCCGAGGCGGCCACCGGCUGGC AGACCAUCGACGGCAAGAAGUACUACUUAACACCAACA CCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACGGCA AGAAGUACUACUUAACACCAACACCGCCAUCGCCUCCA CCGGCUACACCAUCAUCAACGGCAAGCACUUCUACUUA ACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCC CCAACGGCUUCGAGUACUUCGCCCCGCCAACACCGACG	SEQ ID NO:41

CCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAGAACG AGUUCCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU CCGACUCCAAGGCCGUGACCGGCUUGGCGCAUCAUCAACA ACAAGAAGUACUACUUAACCCCAACAACGCCAUCCGG CCAUCCACCUGUGACCAUCAACAACGACAAGUACUACU UCUCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA UCGAGCGCAACAACUUCUACUUCGACGCCAACCGAGGU CCAAGAUGGUGACCGGGCGUGUUAAGGGCCCCAACGGCU UCGAGUACUUCGCCCCCGCCAACACCCACAACAACAACA UCGAGGGCCAGGCCAUCCUGUACCAGAACAAGUCCUGA CCCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUUGGACACCAUCGACGGCAAGAAGU ACUACUUAACCUGAACACCGCCGAGGCCGCCACCGGCU GGCAGACCAUCGACGGCAAGAAGUACUACUUAACCUGA ACACCGCCGAGGCCGCCACCGGCUUGGACACCAUCGACG GCAAGAAGUACUACUUAACACCAACACCUUCAUCGCCU CCACCGGCUACACCUCCAUAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGG GCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAACUUCGAGGGCCAGGCCAUCCUGUACCAGA ACAAGUCCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCGACUCCAAGGCCGUGACCGGCCUGCGCACCACUCC ACGGCAAGAAGUACUACUUAACACCAACACCGCCGUGG CCGUGACCGGCUUGGACACCAUAACGGCAAGAAGUACU ACUUAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCCUGUA CCUGCACGACAACAUCUACUACUUCGGCCAGAACUCCAA GGCCGCCACCGGCUUGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCCUGCA CCUGCUGGGCAAGAUCUACUACUUCGGCCAGAACUCCAA GGCCGUGACCGGCUUGGACACCAUAACGGCAAGGUGUA CUACUUAUGCCCGACACCGCCAUGGCCGCGCGCGGCGG CCUGUUCGAGAUCCGACGGCGUGAUCUACUUCUUCGGCGU GGACGGCGUGAAGGCCCGGCAUCUACGGCGGCGGCUU CGGCGGCGGCGGCUCCGGCCUGUCCAAGGACAUAUCAA GCUGCUGAACGAGCAGGUGAACAAGGAGAUGAACUCCUC CAACCUGUACAUGUCCAUGUCCUCCUGGUGCUACACCCA CUCCUGGACGGCGCCGGCCUGUCCUGUUCGACCACGC CGCCGAGGAGUACGAGCACGCCAAGAAGCUGAUCAUCU CCUGAACGAGAACAACGUGCCCGUGCAGCUGACCUCCA CUCCGCCCCCGAGCAACAAGUUCGAGGGCCUGACCCAGAU CUUCCAGAAGGCCUACGAGCACGAGCAGCACAUCCGA GUCCAUAACAACUUCGUGGACCACGCCAUCAAGUCCAA GGACCACGCCACCUUAACUUCGUGCAGUGGUACGUGGC CGAGCAGCACGAGGAGGAGGUGCUGUUAAGGACAUCU GGACAAGAUCGAGCUGAUCGGCAACGAGAACCACGGCCU GUACCGGGCGACAGUACGUGAAGGGCAUCGCCAAGUC CCGCAAGUCCUGGAGGUGCUGUUCAGGGCCCCGGCUC CCACCACCACCACCACCACUAAUAAACUAGUAGUGACU	
--	--

	<u>GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACA</u> <u>CCCGAAUGGAGUCUCUAAGCUACAUAUACCAACUUACA</u> <u>CUUACAAAAUGUUGUCCCCCAAAUGUAGCCAUUCGUAU</u> <u>CUGCUCUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdA_ VRCHpyl ferritin mRNA with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCUUCAACCUGGUGACCGGC UGGCAGACCAUCAACGGCAAGAAGUACUACUUCGACAUC AACACCGGCGCCGCCUGACCUCCUACAAGAUCAUAAC GGCAAGCACUUCUACUUAACAACGACGGCGUGAUGCAG CUGGGCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUC GCCCCCGCCAACACCCAGAACAACAACUAGGGGCCAG GCCAUCGUGUACCAGUCCAAGUUCUGACCCUGAACGGC AAGAAGUACUACUUCGACCAGAACUCCAAGGCCGUGACC GGCUGGGCGCAUCAUCAACAACGAGAAGUACUACUUAAC CCCAACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUC GACAACAACAAGUACUACUUAACCCCGACACCGCCAUC AUCUCCAAGGGCUGGCAGACCGUGCAGGGCUCGCGCUAC UACUUCGACACCGACACCGCCAUCGCCUUAACGGCUAC AAGACCAUCGACGGCAAGCACUUCUACUUCGACUCCGAC UGCGUGGUGAAGAUCGGCGUGUUCUCCACCUCCAACGGC UUCGAGUACUUCGCCCCCGCCAACACCUACAACAACAAC AUCGAGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUG ACCCUGAACGGCAAGAAGUACUACUUCGACCAGAACUCC AAGGCCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAG UACUACUUAACACCAACACCGCCGAGGCCGCCACCGGC UGGCAGACCAUCGACGGCAAGAAGUACUACUUAACACC AACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGAC GGCAAGAAGUACUACUUAACACCAACACCGCCAUCGCC UCCACCGGCUACACCAUCAUAACGGCAAGCACUUCUAC UUCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAG GGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACC GACGCCAACAAUCGAGGGGCCAGGCCAUCCUGUACCAG AACGAGUUCUGACCCUGAACGGCAAGAAGUACUACUUC GGCUCCGACUCCAAGGCCGUGACCGGCUGGCGCAUAC AACAAACAAGAAGUACUACUUAACCCCAACAACGCCAUC GCCGCCAUCCACUGUGCACCACCAACAACGACAAGUAC UACUUCUCCUACGACGGCAUCCUGCAGAACGGCUACAUC ACCAUCGAGCGCAACAACUUCUACUUCGACGCCAACAG GAGUCCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAAC GGCUUCGAGUACUUCGCCCCCGCCAACACCCACAACAAC AACAUUCGAGGGCCAGGCCAUCGUGUACCAGAACAGUUC CUGACCCUGAACGGCAAGAAGUACUACUUCGACCAGGAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAG AAGUACUACUUAACUGAACACCGCCGAGGCCGCCACC GGCUGGCAGACCAUCGACGGCAAGAAGUACUACUUAAC CUGAACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUC GACGGCAAGAAGUACUACUUAACACCAACACCUUCAUC GCCUCCACCGGCUACACCUCCAUAACGGCAAGCACUUC UACUUAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUC AAGGGCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAAC ACCGACGCCAACAAUCAUCGAGGGCCAGGCCAUCCUGUAC CAGAACAAGUUCUGACCCUGAACGGCAAGAAGUACUAC	SEQ ID NO:42

	UUCGGCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACC AUCGACGGCAAGAAGUACUACUUAACACCAACACCGCC GUGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAG UACUACUUAACACCCAGACCUCCAUCGCCUCCACCGGC UACACCAUCAUCUCCGGCAAGCACUUCUACUUAACACC GACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGAC GGCUUCGAGUACUUCGCCCCCGCCAACACCGACGCCAAC AACAUCCGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUC CUGUACCUGCACGACAACAUCUACUACUUCGGCCAGAAC UCCAAGGCCGCCACCGGCUGGGUGACCAUCGACGGCAAC CGCUACUACUUCGAGCCCAACACCGCCAUGGGCGCCAAC GGCUACAAGACCAUCGACAACAAGAACUUCUACUUCGCG AACGGCCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAAC GGCUUCGAGUACUUCGCCCCCGCCAACACCGACGCCAAC AACAUCCGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUC CUGCACCUGCUGGGCAAGAUCUACUACUUCGGCCAGAAC UCCAAGGCCGUGACCGGCUGGCAGACCAUCAACGGCAAG GUGUACUACUUAUGCCCCGACACCGCCAUGGCCGCCGCC GGCGGCCUGUUCGAGAUCCGACGGCGUGAUCUACUUCUUC GGCGUGGACGGCGUGAAGGGCCCCGGCAUCUACUCCGGC GGCGACAUCAUCAAGCUGCUGAACGAGCAGGUGAACAAAG GAGAUGCAGUCCUCCAACCUGUACAUGUCCAUGUCCUCC UGGUGCUACACCCACUCCUUGGACGGCGCCGGCCUGUUC CUGUUCGACCACGCCGCCGAGGAGUACGAGCACGCCAAG AAGCUGAUCAUCUUCUGAACGAGAACAACGUGCCCGUG CAGCUGACCUCCAUCUCCGCCCGGAGCACAAGUUCGAG GGCCUGACCCAGAUUCUCCAGAAGGCCUACGAGCACGAG CAGCACAUCUCCGAGUCCAUCAACAACAUCGUGGACCAC GCCAUCAAGUCCAAGGACCACGCCACCUUCAACUUCUG CAGUGGUACGUGGCCGAGCAGCACGAGGAGGAGGUGCUG UUCAAGGACAUCUGGACAAGAUCGAGCUGAUCGGCAAC GAGAACCACGGCCUGUACCUGGCCGACCAGUACGUGAAG GGCAUCGCCAAGUCCCGCAAGUCCGGCUCCUGGAGGUG CUGUCCAGGGCCCCGGCUCCUAAUAA	
MutTcdA_ VRCHpyl ferritin mRNA with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UAAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUUGUCCUGGC CCUGGUGACCAACUCCUUAACCUUGGUGACCGGCUGGC AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA CCGGCGCCGCCUGACCUCCUACAAGAUAUCAACGGCA AGCACUUCUACUUAACAACGACGGCGUGAUGCAGCUGG GCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUCGCCC CCGCCAACACCCAGAACAACAACAUCCGAGGGCCAGGCCA UCGUGUACCAGUCCAAGUUCUGACCCUGAACGGCAAGA AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU GGCGCAUCAUCAACAACGAGAAGUACUACUUAACCCCA ACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUCGACA ACAACAAGUACUACUUAACCCCGACACCGCCAUCAUCU CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG UGGUGAAGAUCGGCGUGUUCUCCACCUCCAACGGCUUCG AGUACUUCGCCCCCGCCAACACCUACAACAACAACUUCG AGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUGACCC	SEQ ID NO:43

UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAGG CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU ACUUCAACACCAACACCGCCGAGGCCGCCACCGGCUGGC AGACCAUCGACGGCAAGAAGUACUACUUAACACCAACA CCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACGGCA AGAAGUACUACUUAACACCAACACCGCCAUCGCCUCCA CCGGCUACACCAUCAUAACGGCAAGCACUUCUACUUA ACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCC CCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCGACG CCAACAACAUCCGAGGGCCAGGCCAUCCUGUACCAGAACG AGUUCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU CCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUCA ACAAGAAGUACUACUUAACCCCAACAACGCCAUCGCCG CCAUCCACCUGUGACCAUCAACAACGACAAGUACUACU UCUCCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA UCGAGCGCAACAACUUCUACUUCGACGCCAACCAGGAGU CCAAGAUGGUGACCGGCGUGUUAAGGGCCCCAACGGCU UCGAGUACUUCGCCCCCGCCAACACCCACAACAACA UCGAGGGCCAGGCCAUCCUGUACCAGAACAAGUCCUGA CCCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAGAAGU ACUACUUAACCUGAACACCGCCGAGGCCGCCACCGCU GGCAGACCAUCGACGGCAAGAAGUACUACUUAACCUGA ACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACG GCAAGAAGUACUACUUAACACCAACACCUUAUCGCCU CCACCGGCUACACCUCCAUAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGG GCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAACAUCGAGGGCCAGGCCAUCCUGUACCAGA ACAAGUUCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACCAUCG ACGGCAAGAAGUACUACUUAACACCAACACCGCCGUGG CCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAGUACU ACUUCAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGACGGCU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCCUGUA CCUGCACGACAACAUCUACUACUUCGGCCAGAACUCCAA GGCCGCCACCGGCUGGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAACAU CGAGGGCCAGGCCAUCCGCUACCAGAACCGCUUCCUGCA CCUGCUGGGCAAGAUUCUACUACUUCGGCCAGAACUCCAA GGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGGUGUA CUACUUAUGCCCCGACACCGCCAUGGCCGCCGCCGGCGG CCUGUUCGAGAUCCGACGGCGUGAUCUACUUCUUCGGCGU GGACGGCGUGAAGGCCCCCCGGCAUCUACUCCGGCGGCGA CAUCAUAAGCUGCUGAACGAGCAGGUGAACAAGGAGU GCAGUCCUCCAACCUGUACAUGUCCAUGUCCUCCUGGUG CUACACCCACUCCUGGACGGCGCCGGCCUGUUCUGUU CGACCACGGCCGCCGAGGAGUACGAGCACGCCAAGAAGCU GAUCAUCUCCUGAACGAGAACAACGUGCCCGUGCAGCU GACCUCCAUCUCCGCCCCCGAGCACAGUUCGAGGGCCU	
--	--

	<u>GACCCAGAUCUUC</u><u>CAGAAGGCCUACGAGCACGAGCAGCA</u> <u>CAUCUCCGAGUCCAUCAACAACAU</u><u>CGUGGACCACGCCAU</u> <u>CAAGUCCAAGGACCACGCCACCUUCAACU</u><u>UCCUGCAGUG</u> <u>GUACGUGGCCGAGCAGCACGAGGAGGAGGUGCUGU</u><u>UCAA</u> <u>GGACAUCCUGGACAAGAUCGAGCUGAUCGGCAACGAGAA</u> <u>CCACGGCCUGUACCUGGCCGACCAGUACGUGAAGGGCAU</u> <u>CGCCAAGUCCCGCAAGUCCGGCUCCCUGGAGGUGCUGUU</u> <u>CCAGGGCCCCGGCUCCUAAUAAACUAGUAGUGACUGACU</u> <u>AGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCG</u> <u>AAUGGAGUCUCUAAGCUACAUAUAAUACCAACUACACUUA</u> <u>CAAAAUGUUGUCCCCCAAAUAGUAGCCAUUCGUAUCUGC</u> <u>UCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdA_ VRCHpyl ferritin mRNA with modified IL-2 signal and 6xH tag	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAUUUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> <u>GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC</u> <u>CCUGGUGACCAACUCCUUAACCUGGUGACCGGCUGGC</u> <u>AGACCAUCAACGGCAAGAAGUACUACUUCGACAUCAACA</u> <u>CCGGCGCCGCCUGACCUCUACAAGAUAUCAACGGCA</u> <u>AGCACUUCUACUUAACAACGACGGCGUGAUGCAGCUGG</u> <u>GCGUGUUAAGGGCCCCGACGGCUUCGAGUACUUCGCCC</u> <u>CCGCCAACCCAGAACACAACAUCGAGGGCCAGGCCA</u> <u>UCGUGUACCAGUCCAAGUUCUGACCCUGAACGGCAAGA</u> <u>AGUACUACUUCGACCAGAACUCCAAGGCCGUGACCGGCU</u> <u>GGCGCAUCAUCAACAACGAGAAGUACUACUUAACCCCA</u> <u>ACAACGCCAUCGCCGCCGUGGGCCUGCAGGUGAUCGACA</u> <u>ACAACAAGUACUACUUAACCCCGACACCGCCAUAUCU</u> <u>CCAAGGGCUGGCAGACCGUGCAGGGCUCCCGCUACUACU</u> <u>UCGACACCGACACCGCCAUCGCCUUAACGGCUACAAGA</u> <u>CCAUCGACGGCAAGCACUUCUACUUCGACUCCGACUGCG</u> <u>UGGUGAAGAUCGGCGUGUUCUCCACCUCACCGGCUUCG</u> <u>AGUACUUCGCCCCCGCCAACACCUACAACAACAACUACG</u> <u>AGGGCCAGGCCAUCGUGUACCAGUCCAAGUUCUGACCC</u> <u>UGAACGGCAAGAAGUACUACUUCGACCAGAACUCCAAGG</u> <u>CCGUGACCGGCCUGCAGACCAUCGACUCCAAGAAGUACU</u> <u>ACUUAACACCAACACCGCCGAGGCCGCCACCGGCUGGC</u> <u>AGACCAUCGACGGCAAGAAGUACUACUUAACACCAACA</u> <u>CCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACGGCA</u> <u>AGAAGUACUACUUAACACCAACACCGCCAUCGCCUCCA</u> <u>CCGGCUACACCAUCAUAACGGCAAGCACUUCUACUUA</u> <u>ACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGGGCC</u> <u>CCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCGACG</u> <u>CCAACAACAU</u><u>CGAGGGGCCAGGCCAUCCUGUACCAGAACG</u> <u>AGUUCUGACCCUGAACGGCAAGAAGUACUACUUCGGCU</u> <u>CCGACUCCAAGGCCGUGACCGGCUGGCGCAUCAUCAACA</u> <u>ACAAGAAGUACUACUUAACCCCAACAACGCCAUCGCCG</u> <u>CCAUCCACCUGUGCACCAUCAACAACGACAAGUACUACU</u> <u>UCUCCUACGACGGCAUCCUGCAGAACGGCUACAUCACCA</u> <u>UCGAGCGCAACAACUUCUACUUCGACGCCAACCAGGAGU</u> <u>CCAAGAUGGUGACCGGGCGUGUUAAGGGCCCCAACGGCU</u> <u>UCGAGUACUUCGCCCCCGCCAACACCCACAACAACAACA</u>	SEQ ID NO:44

	UCGAGGGCCAGGCCAUCGUGUACCAGAACAAGUUCUGA CCCUGAACGGCAAGAAGUACUACUUCGACCAGGACUCCA AGGCCGUGACCGGCUGGCAGACCAUCGACGGCAAGAAGU ACUACUUAACCUGAACACCGCCGAGGCCGCCACCGGCU GGCAGACCAUCGACGGCAAGAAGUACUACUUAACCUGA ACACCGCCGAGGCCGCCACCGGCUGGCAGACCAUCGACG GCAAGAAGUACUACUUAACACCAACACCUUCAUCGCCU CCACCGGCUACACCUCCAUAACGGCAAGCACUUCUACU UCAACACCGACGGCAUCAUGCAGAUCCGGCGUGUUAAGG GCCCCAACGGCUUCGAGUACUUCGCCCCCGCCAACACCG ACGCCAACAAUCGAGGGCCAGGCCAUCCUGUACCAGA ACAAGUUCUGACCCUGAACGGCAAGAAGUACUACUUCG GCUCCGACUCCAAGGCCGUGACCGGCCUGCGCACCAUCG ACGGCAAGAAGUACUACUUAACACCAACACCGCCGUGG CCGUGACCGGCUGGCAGACCAUCAACGGCAAGAAGUACU ACUUAACACCCAGACCUCCAUCGCCUCCACCGGCUACAC CAUCAUCUCCGGCAAGCACUUCUACUUAACACCGACGG CAUCAUGCAGAUCCGGCGUGUUAAGGGCCCCGACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAAAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGUA CCUGCACGACAACAUCUACUACUUCGGCCAGAACUCCAA GGCCGCCACCGGCUGGGUGACCAUCGACGGCAACCGCUA CUACUUCGAGCCCAACACCGCCAUGGGCGCCAACGGCUA CAAGACCAUCGACAACAAGAACUUCUACUUCCGCAACGG CCUGCCCCAGAUCCGGCGUGUUAAGGGCUCCAACGGCUU CGAGUACUUCGCCCCCGCCAACACCGACGCCAACAAAU CGAGGGCCAGGCCAUCCGCUACCAGAACCAGCUUCCUGCA CCUGCUGGGCAAGAUCUACUACUUCGGCCAGAACUCCAA GGCCGUGACCGGCUGGCAGACCAUCAACGGCAAGGUGUA CUACUUCUAGCCCCGACACCGCCAUGGGCGCCGCCGGCGG CCUGUUCGAGAUCCGACGGCGUGAUUCUACUUCGCGGCU GGACGGCGUGAAGGCCCGGCAUCUACUCCGGCGGCGGA CAUCAUCAAGCUGCUGAACGAGCAGGUGAACAAAGGAGAU GCAGUCCUCCAACCUGUACAUGUCCAUGUCCUCCUGGUG CUACACCCACUCCUGGACGGCGCCGGCCUGUCCUGU CGACCACGCCGCCGAGGAGUACGAGCACGCCAAGAAGCU GAUCAUCUCCUGAACGAGAACAAACGUGCCCGUGCAGCU GACCUCCAUCUCCGCCCCGAGCACAAAGUUCGAGGGCCU GACCCAGAUUCUCCAGAAGGCCUACGAGCACGAGCAGCA CAUCUCCGAGUCCAUAACAACAUCGUGGACCACGCCAU CAAGUCCAAGGACCACGCCACCUUCAACUCCUGCAGUG GUACGUGGCCGAGCAGCACGAGGAGGAGGUGCUGUUA GGACAUCUGGACAAGAUCGAGCUGAUCGGCAACGAGAA CCACGGCCUGUACCUGGCCGACCAGUACGUGAAGGGCAU CGCCAAGUCCCGCAAGUCCGGCUCCUGGAGGUGCUGUU CCAGGGCCCCGGCUCCACCAACCAACCAACCAACUAAUA AACUAGUAGUGACUGACUAGGAUCUGGUUACCACUAAAC CAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACA AAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAU GUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUC UUCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
TcdB RNA	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCAACCUGAUCACCGGCUUC	SEQ ID NO:45

	GUGACCGUGGGCGACGACAAGUACUACUUCAACCCCAUC AACGGCGGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGAC GACAAGAACUACUACUCCAGCAGUCCGGCGUGCUGCAG ACCGGGGUGUUCUCCACCGAGGACGGCUUCAAGUACUUC GCCCCCGCCAACACCCUGGACGAGAACCUGGAGGGCGAG GCCAUCGACUUCACCGGCAAGCUGAUAUCGACGAGAAC AUCUACUACUUCGACGACAACUACCGCGGCGCCGUGGAG UGGAAGGAGCUGGACGGCGAGAUACUACUUCUCCCCC GAGACCGGCAAGGCCUUCAAGGGCCUGAACCAUAUCGGC GACUACAAGUACUACUUAACUCCGACGGCGUGAUGCAG AAGGGCUUCGUGUCCAUAACGACAACAAGCACUACUUC GACGACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUC GACGGCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUG CAGAUCCGGCGUGUUAACACCGAGGACGGCUUCAAGUAC UUCGCCCACCACAACGAGGACCUGGGCAACGAGGAGGGC GAGGAGAUUCUACUCCGGCAUCCUGAACUUAACAAC AAGAUACUACUUCGACGACUCCUACACCGCCGUGGUG GGCUGGAAGGACCUGGAGGACGGCUCCAAGUACUACUUC GACGAGGACACCGCCGAGGCCUACAUCGGCCUGUCCUG AUCAACGACGGCCAGUACUACUUAACGACGACGGCAUC AUGCAGGUGGGCUUCGUGACCAUAACGACAAGGUGUUC UACUUCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAG AACAUCCGACGACAACUACUUCUACAUCGACGACAACGGC AUCGUGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUAC AAGUACUUCGCCCCCGCCAACACCGUGAACGACAACAUC UACGGCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUG GGCGAGGACGUGUACUACUUCGGCGAGACCUACACCAUC GAGACCGGCUUGAUCUACGACAUGGAGCAGGAGUCCGAC AAGUACUACUUAACCCCGAGACCAAGAAGGCCUGCAAG GGCAUCAACCUGAUCGACGACAUAAGUACUACUUCGAC GAGAAGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAG AACAAACAACUACUACUUAACGAGAACGGCGAGAUGCAG UUCGGCUACAUAACAUCGAGGACAAGAUGUUCUACUUC GGCGAGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACC CCCGACGGCUUCAAGUACUUCGCCCCACCAGAACACCCUG GACGAGAACUUCGAGGGCGAGUCCAUCCAGUACACCGGC UGGCUGGACCUGGACGAGAGAAGCGCUACUACUUCACCGAC GAGUACAUCGCCGCCACCGGCUCCGUGAUAUCGACGGC GAGGAGUACUACUUCGACCCCGACACCGCCAGCUGGUG AUCUCCGAGUAAUAA	
TcdB RNA: (5'UTR, coding sequence, 3'UTR and polyA)	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUAUGCGCAUG</u> CAGCUGCUGCUGCUGAUCGCCUUGUCCUUGGCCUUGG UGACCAACUCCAACCUGAUCACCGGCUUCGUGACCGUGG GCGACGACAAGUACUACUUAACCCCAUCAACGGCGGGCG CCGCCUCCAUCGGCGAGACCAUAUCGACGACAAGAACU ACUACUUCAGCAGUCCGGCGUGCUGCAGACCGGGCGUGU UCUCCACCGAGGACGGCUUCAAGUACUUCGCCCCCGCCA ACACCCUGGACGAGAACCUGGAGGGCGAGGCCAUCGACU UCACCGGCAAGCUGAUAUCGACGAGAACAUACUACUACU UCGACGACAACUACCGCGCGCGCGUGGAGUGGAAGGAGC UGGACGGCGAGAUACUACUUCUCCCCCGAGACCGGCA AGGCCUUAAGGGCCUGAACCAAGAUCCGGCGACUACAAGU ACUACUUAACUCCGACGGCGUGAUGCAGAAGGGCUUCG	SEQ ID NO:46

	UGUCCAUCAACGACAACAAGCACUACUUCGACGACUCCG GCGUGAUGAAGGUGGGCUACACCGAGAUUCGACGGCAAGC ACUUCUACUUCGCCGAGAACGGCGAGAUUCGAGAUCCGCG UGUUCAACACCGAGGACGGCUUCAAGUACUUCGCCACC ACAACGAGGACCUGGGCAACGAGGAGGGCGAGGAGAUUCU CCUACUCCGGCAUCCUGAACUUAACAACAAGAUUCUACU ACUUCGACGACUCCUUCACCGCCGUGGUGGGCUUGGAAGG ACCUGGAGGACGGCUCCAAGUACUACUUCGACGAGGACA CCGCCGAGGCCUACAUCGGCCUGUCCCUGAUCAACGACG GCCAGUACUACUUAACGACGACGGCAUCAUGCAGGUGG GCUUCGUGACCAUCAACGACAAGGUGUUCUACUUCUCCG ACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACAUUCGACG ACAACUACUUCUACAUCGACGACAACGGCAUCGUGCAGA UCGGCGUGUUCGACACCUCCGACGGCUACAAGUACUUCG CCCCCGCCAACACCGUGAACGACAACAUCUACGGCCAGG CCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCGAGGACG UGUACUACUUCGGCGAGACCUACACCAUCGAGACCGGCU GGAUCUACGACAUGGAGCAGGAGUCCGACAAGUACUACU UCAACCCCGAGACCAAGAAGGCCUGCAAGGGCAUCAACC UGAUCGACGACAUCAAGUACUACUUCGACGAGAAGGGCA UCAUGCGCACCGGCCUGAUCUCCUUCGAGAACACAACU ACUACUUAACGAGAACGGCGAGAUAGCAGUUCGGCUACA UCAACAUCGAGGACAAGAUGUUCUACUUCGGCGAGGACG GCGUGAUGCAGAUCCGGCGUGUUAACACCCCCGACGGCU UCAAGUACUUCGCCCACCAGAACACCCUGGACGAGAACU UCGAGGGCGAGUCCAUCAGUACACCGGCGUGGUGGACC UGGACGAGAAGCGCUACUACUUCACCGACGAGUACAUCG CCGCCACCGGCUCCGUGAUCUACGACGGCGAGGAGUACU ACUUCGACCCCGACACCGCCAGCUGGUGAUCUCCGAGU AAUAAACUAGUAGUGACUGACUAGGAUCUGGUUACCACU <u>AAACCGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCU</u> <u>ACAUAAUACCAACUUAACUUAACAAAUGUUGUCCCCCA</u> <u>AAAUGUAGCCAUCGUACUGCUCCUAAUAAAAAGAAAG</u> <u>UUUCUUCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>A</u>	
MutTcdB_Hp ylferritin mRNA with modified IL-2 signal	AUGCGAUGCAGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCAACCUGAUCACCGGCUUC GUGACCGUGGGCGACGACAAGUACUACUUAACCCCAUC AACGGCGGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGAC GACAAGAACUACUACUUCAGCAGUCCGGCGUGCUGCAG ACCGGCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUC GCCCCCGCCAACACCCUGGACGAGAACCUGGAGGGCGAG GCCAUCGACUUCACCGGCAAGCUGAUCUACGACGAGAAC AUCUACUACUUCGACGACAACUACCGCGGCGCCGUGGAG UGGAAGGAGCUGGACGGCGAGAUACUACUUCUCCCCC GAGACCGGCAAGGCCUUAAGGGCCUGAACCAGAUCCGGC GACUACAAGUACUACUUAACUCCGACGGCGUGAUGCAG AAGGGCUUCGUGUCCAUAACGACAACAAGCACUACUUC GACGACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAU GACGGCAAGCACUUCUACUUCGCCGAGAACGGCGAGAU CAGAUCCGGCGUGUUAACACCGAGGACGGCUUCAAGUAC UUCGCCACCAACGAGGACCUGGGCAACGAGGAGGGC GAGGAGAUCCUACUCCGGCAUCCUGAACUUAACAAC	SEQ ID NO:47

	AAGAUCUACUACUUCGACGACUCCUUCACCGCCGUGGUG GGCUUGAAGGACCUGGAGGACGGCUCCAAGUACUACUUC GACGAGGACACCGCCGAGGCCUACAUCGGCCUGUCCCUG AUCAACGACGGCCAGUACUACUUC AACGACGACGGCAUC AUGCAGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUC UACUUCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAG AACAUUCGACGACAACUACUUCUACAUCGACGACAACGGC AUCGUGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUAC AAGUACUUCGCCCCGCAACACCGUGAACGACAACAUC UACGGCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUG GGCGAGGACGUGUACUACUUCGGCGAGACCUACACCAUC GAGACCGGCUGGAUCUACGACAUGGAGCAGGAGUCCGAC AAGUACUACUUC AACCCCGAGACCAAGAAGGCCUGCAAG GGCAUCAACCUGAUCGACGACAUCAAGUACUACUUCGAC GAGAAGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAG AACAAACAUCUACUUCUACACGAGAACGGCGAGAUGCAG UUCGGCUACAUCAACAUCGAGGACAAGAUGUUCUACUUC GGCGAGGACGGCGUGAUGCAGAUCCGGCGUGUACAACACC CCCGACGGCUUCAAGUACUUCGCCCACCAGAACACCCUG GACGAGAACUUCGAGGGCGAGUCCAUCCAGUACACCGGC UGGCUGGACCUGGACGAGAAGCGCUACUACUUCACCGAC GAGUACAUCGCCGCCACCGGCUCGUGAUCUACGACGGC GAGGAGUACUACUUCGACCCCGACACCGCCAGCUGGUG AUCUCCGAGGGCUCCGGCGACAUCAUCAAGCUGCUGAAC GAGCAGGUGAACAAAGGAGAUGAACUCCUCCAACCUGUAC AUGUCCAUGUCCUCCUGGUGCUACACCCACUCCUCCUGGAC GGCGCCGGCCUGUCCUGUUCGACCACGCCGCCGAGGAG UACGAGCACGCCAAGAAGCUGAUCUUCUCCUGAACGAG AACAAACGUGCCCGUGCAGCUGACCUCCAUCUCCGCCCCC GAGCACAAGUUCGAGGGCCUGACCCAGAUUCUCCAGAAG GCCUACGAGCACGAGCAGCACAUUCUCCGAGUCCAUCAAC AACAUUCGUGGACCACGCCAUCAAGUCCAAGGACCACGCC ACCUUCAACUUCUCCUGCAGUGGUACGUGGCCGAGCAGCAC GAGGAGGAGGUGCUGUUC AAGGACAUCUCCUGGACAAGAUC GAGCUGAUCGGCAACGAGAACCACGGCCUGUACCUCCGCC GACCAGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCC CUGGAGGUGCUGUUCACGGGCCCCGGCUCCUAAUAA	
MutTcdB_Hp ylferritin mRNA with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACACAACAUUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUUGUCCUUGGC CCUGGUGACCAACUCCAACCUGAUCACCGGCUUCGUGA CCGUGGGCGACGACAAGUACUACUUC AACCCCAUCAACG GCGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGACGACA AGAACUACUACUUC CAGCAGUCCGGCGUGCUGCAGACCG GCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUCGCCC CCGCCAACACCCUGGACGAGAACCUGGAGGGCGAGGCCA UCGACUUCACCGGCAAGCUGAUCUACGACGAGAACAUUCU ACUACUUCGACGACAACUACCGCGGCCGUGGAGUGGA AGGAGCUGGACGGCGAGAUACUACUUCUCCCCGAGA CCGGCAAGGCCUUAAGGGCCUGAACCAAGAUCCGGCGACU ACAAGUACUACUUC AACUCCGACGGCGUGAUGCAGAAGG GCUUCGUGUCCAUCAACGACAACAAGCACUACUUCGACG ACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUCCGACG	SEQ ID NO:48

	GCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUGCAGA UCGGCGUGUUAACACCGAGGACGGCUUCAAGUACUUCG CCCACCACAACGAGGACCUGGGCAACGAGGAGGGCGAGG AGAUCUCCUACUCCGGCAUCCUGAACUUAACAACAAGA UCUACUACUUCGACGACUCCUUCACCGCCGUGGUGGGCU GGAAGGACCUGGAGGACGGCUCCAAGUACUACUUCGACG AGGACACCGCCGAGGCCUACAUCGGCCUGUCCUGAUA ACGACGGCCAGUACUACUUAACGACGACGGCAUCAUGC AGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUCUACU UCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACA UCGACGACAACUACUUCUACAUCGACGACAACGGCAUCG UGCAGAUCGGCGUGUUCGACACCUCGACGGCUACAAGU ACUUCGCCCCCGCCAACACCGUGAACGACAACAUCUACG GCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCG AGGACGUGUACUACUUCGGCGAGACCUACACCAUCGAGA CCGGCUGGAUCUACGACAUGGAGCAGGAGUCCGACAAGU ACUACUUAACCCCGAGACCAAGAAGGCCUGCAAGGGCA UCAACCUGAUCGACGACAUAAGUACUACUUCGACGAGA AGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAGAACA ACAACUACUACUUAACGAGAACGGCGAGAUGCAGUUCG GCUACAUAACAUCGAGGACAAGAUGUUCUACUUCGGCG AGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACCCCGC ACGGCUUAAGUACUUCGCCACCAGAACACCCUGGACG AGAACUUCGAGGGCGAGUCCAUCAGUACACCGGCUGGC UGGACCUGGACGAGAAGCGCUACUACUUCACCGACGAGU ACAUCGCCGCCACCGGCUCCGUGAUAUCGACGGCGAGG AGUACUACUUCGACCCCGACACCGCCCAGCUGGUGAUCU CCGAGGGCUCCGGCGACAUAUCAAGCUGCUGAACGAGC AGGUGAACAAAGGAGAUGAACUCCUCCAACCGUACAUGU CCAUGUCCUCCUGGUGCUACACCCACUCCCUGGACGGCG CCGGCCUGUUCUGUUCGACCACGCCGCCGAGGAGUACG AGCACGCCAAGAAGCUGAUAUCUUCUGAACGAGAACA ACGUGCCCGUGCAGCUGACCUCCAUCUCCGCCCCCGAGC ACAAGUUCGAGGGCCUGACCCAGAUCUUCAGAAAGGCCU ACGAGCACGAGCAGCACAUCUCCGAGUCCAUAACAACA UCGUGGACCACGCCAUCAAGUCCAAGGACCACGCCACCU UCAACUCCUGCAGUGGUACGUGGCCGAGCAGCAGGAGG AGGAGGUGCUGUUAAGGACAUCUGGACAAGAUCGAGC UGAUCGGCAACGAGAACCACGGCCUGUACCUGGCCGACC AGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCCUGG AGGUGCUGUUCAGGGCCCCGGCUCCUAAUAAACUAGUA <u>GUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCA</u> <u>AGAACACCCGAAUGGAGUCUCUAAGCUACAUAUACCAA</u> <u>CUUACACUUACAAAUGUUGUCCCCCAAAUGUAGCCAU</u> <u>UCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUU</u> <u>CUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdB_Hp ylferritin mRNA with modified IL-2 signal with 6xH tag	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAA</u> <u>UUAUAUACUUCUUCUUAUUAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> <u>GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC</u> <u>CCUGGUGACCAACUCCAACCUGAUCACCGGCUUCGUGA</u> <u>CCGUGGGCGACGACAAGUACUACUUAACCCCAUCAACG</u>	SEQ ID NO:49

	GCGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGACGACA AGAACUACUACUCCAGCAGUCCGGCGUGCUGCAGACCG GCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUCGCCC CCGCCAACACCCUGGACGAGAACCUGGAGGGCGAGGCCA UCGACUUCACCGGCAAGCUGAUCAUCGACGAGAACAUUCU ACUACUUCGACGACAACUACCGCGGGCGCCGUGGAGUGGA AGGAGCUGGACGGCGAGAUACUACUUCUCCCCCGAGA CCGGCAAGGCCUUCAAGGGCCUGAACAGAUCCGGCGACU ACAAGUACUACUUAACUCCGACGGCGUGAUGCAGAAGG GCUUCGUGUCCAUCAACGACAACAAGCACUACUUCGACG ACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUCCGACG GCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUCCGAGA UCGGCGUGUUAACACCGAGGACGGCUUCAAGUACUUCG CCCACCACAACGAGGACCUGGGCAACGAGGAGGGCGAGG AGAUCUCCUACUCCGGCAUCCUGAACUUAACAACAAGA UCUACUACUUCGACGACUCCUUCACCGCCGUGGUGGGCU GGAAGGACCUGGAGGACGGCUCCAAGUACUACUUCGACG AGGACACCGCCGAGGCCUACAUCGGCCUGUCCCUGAUCA ACGACGGCCAGUACUACUUAACGACGACGGCAUCAUGC AGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUCUACU UCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACA UCGACGACAACUACUUCUACAUCGACGACAACGGCAUCG UGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUACAAGU ACUUCGCCCCCGCCAACACCGUGAACGACAACAUCUACG GCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCG AGGACGUGUACUACUUCGGCGAGACCUACACCAUCGAGA CCGGCUUGAUCAACGACAUGGAGCAGGAGUCCGACAAGU ACUACUUAACCCCGAGACCAAGAAGGCCUGCAAGGGCA UCAACCUGAUCGACGACAUAAGUACUACUUCGACGAGA AGGGCAUCAUGCGCACCGGCCUGAUUCUCCUUCGAGAACA ACAACUACUACUUAACGAGAACGGCGAGAUCCAGUUCG GCUACAUAACAUCGAGGACAAGAUGUUCUACUUCGGCG AGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACCCCCG ACGGCUUCAAGUACUUCGCCCACCAGAACACCCUGGACG AGAACUUCGAGGGCGAGUCCAUCCAGUACACCGGCUGGC UGGACCUGGACGAGAAGCGCUACUACUUCACCGACGAGU ACAUCGCCGCCACCGGCUCCGUGAUCAUCCGACGGCGAGG AGUACUACUUCGACCCCGACACCGCCCAGCUGGUGAUCU CCGAGGGCUCCGGCGACAUAUCAAGCUGCUGAACGAGC AGGUGAACAAAGGAGAUGAACUCCUCCAACCUGUACAUGU CCAUGUCCUCCUGGUGCUACACCCACUCCUGGACGGCG CCGGCCUGUUCUGUUCGACCACGCCGCCGAGGAGUACG AGCACGCCAAGAAGCUGAUCAUCUUCUGAACGAGAAC ACGUGCCCGUGCAGCUGACCUCAUCUCCGCCCCCGAGC ACAAGUUCGAGGGCCUGACCCAGAUUCUCCAGAAGGCCU ACGAGCACGAGCAGCAUCUCCGAGUCCAUAACAACA UCGUGGACCACGCCAUCAAGUCCAAGGACCACGCCACCU UCAACUUCUGCAGUGGUACGUGGCCGAGCAGCACGAGG AGGAGGUGCUGUUAAGGACAUCUGGACAAGAUCGAGC UGAUCGGCAACGAGAACCACGGCCUGUACCUGGCCGACC AGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCCUGG AGGUGCUGUUCAGGGCCCCGGCUCCCACCACCACCACC ACCACUAAUAAACUAGUAGUGACUGACUAGGAUCUGGU UACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU CUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUG UCCCCCAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAA	
--	---	--

	<u>AAGAAAGUUUCUUCACAUUCUAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAA</u>	
MutTcdB_2X LSKHpylferri tin with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCCUGUCCC UGGCCCCUGGUGACCAACUCCAACCUGAUCACCGGCUUC GUGACCGUGGGCGACGACAAGUACUACUUAACCCCAUC AACGGCGGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGAC GACAAGAACUACUACUUCAGCAGUCCGGCGUGCUGCAG ACCGGCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUC GCCCCCGCCAACACCCUGGACGAGAACCUGGAGGGCGAG GCCAUCGACUUCACCGGCAAGCUGAUCGACGAGAAC AUCUACUACUUCGACGACAACUACCGCGGGCGCCGUGGAG UGGAAGGAGCUGGACGGCGAGAUACUACUUCUCCCCC GAGACCGGCAAGGCCUUCAGGGCCUGAACCAAGAUCCGGC GACUACAAGUACUACUUAACUCCGACGGCGUGAUGCAG AAGGGCUUCGUGUCCAUAACGACAACAAGCACUACUUC GACGACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUC GACGGCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUG CAGAUCCGGCGUGUUAACACCGAGGACGGCUUCAAGUAC UUCGCCCCACCACAACGAGGACCUGGGCAACGAGGAGGGC GAGGAGAUUCUCCUACUCCGGCAUCCUGAACUUAACAAC AAGAUACUACUACUUCGACGACUCCUACACCGCCGUGGUG GGCUGGAAGGACCUGGAGGACGGCUCCAAGUACUACUUC GACGAGGACACCGCCGAGGCCUACAUCGGCCUGUCCUG AUCAACGACGGCCAGUACUACUUAACGACGACGGCAUC AUGCAGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUC UACUUCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAG AACAUCCGACGACAACUACUUCUACAUCGACGACAACGGC AUCGUGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUAC AAGUACUUCGCCCCCGCCAACACCGUGAACGACAACAUC UACGGCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUG GGCGAGGACGUGUACUACUUCGGCGAGACCUACACCAUC GAGACCGGCUUGAUCUACGACAUGGAGCAGGAGUCCGAC AAGUACUACUUAACCCCGAGACCAAGAAGGCCUGCAAG GGCAUCAACCUGAUCGACGACAUAAGUACUACUUCGAC GAGAAGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAG AACAAACAACUACUACUUAACGAGAACGGCGAGAUGCAG UUCGGCUACAUAACAUCGAGGACAAGAUGUUCUACUUC GGCGAGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACC CCCGACGGCUUAAGUACUUCGCCCCACCAGAACACCCUG GACGAGAACUUCGAGGGCGAGUCCAUCCAGUACACCGGC UGGCUUGGACCUGGACGAGAAGCGCUACUACUUCACCGAC GAGUACAUCGCCGCCACCGGCUCCGUGAUCGACGGC GAGGAGUACUACUUCGACCCCGACACCGCCAGCUGGUG AUCUCCGAGGGCGGGCGGCUCCGGCGGGCGGGCUCCGGC CUGUCCAAGGACAUAUCAAGCUGCUGAACGAGCAGGUG AACAAAGGAGAUGAACUCCUCCAACCUGUACAUGUCCAUG UCCUCCUGGUGCUACACCCACUCCUUGGACGGCGCCGGC CUGUCCUGUUCGACCACGCCGCCGAGGAGUACGAGCAC GCCAAGAAGCUGAUAUCUCCUGAACGAGAACAACGUG CCCGUGCAGCUGACCUCUACUCCGCCCCGAGCACAAG UUCGAGGGCCUGACCCAGAUCUCCAGAAGGCCUACGAG CACGAGCAGCACAUCUCCGAGUCCAUAACAACAUCGUG GACCACGCCAUCAAGUCCAAGGACCACGCCACCUUCAAC	SEQ ID NO:50

	UUCCUGCAGUGGUACGUGGCCGAGCAGCACGAGGAGGAG GUGCUGUUCAAGGACAUCUUGGACAAGAUCGAGCUGAUC GGCAACGAGAACCACGGCCUGUACCUGGCCGACCAGUAC GUGAAGGGCAUCGCCAAGUCCCGCAAGUCCUGGAGGUG CUGUUCAGGGCCCCGGCUCCUAAUAA	
MutTcdB_2X LSKHpylferri tin with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> <u>GCAUGCAGCUGCUGCUGCUGAUCGCCCUUGUCCCUGGC</u> <u>CCUGGUGACCAACUCCAACCUGAUCACCGGCUUCGUGA</u> CCGUGGGCGACGACAAGUACUACUUAACCCCAUCAACG GCGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGACGACA AGAAUACUACUUCAGCAGUCCGGCGUGCUGCAGACCG GCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUCGCC CCGCCAACACCCUGGACGAGAACCUGGAGGGCGAGGCCA UCGACUUCACCGGCAAGCUGAUAUCGACGAGAAACUUCU ACUACUUCGACGACAACUACCGCGGCGCCGUGGAGUGGA AGGAGCUGGACGGCGAGAUACUACUUCUCCCCCGAGA CCGGCAAGGCCUUAAGGGCCUGAACCAAGAUCCGGCGACU ACAAGUACUACUUAACUCCGACGGCGUGAUGCAGAAGG GCUUCGUGUCCAUAACGACAACAAGCACUACUUCGACG ACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUCCGACG GCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUCCAGA UCGGCGUGUUAACACCGAGGACGGCUUCAAGUACUUCG CCCACCACAACGAGGACCUGGGCAACGAGGAGGGCGAGG AGAUCUCCUACUCCGGCAUCCUGAACUUAACAACAAGA UCUACUACUUCGACGACUCCUACACCGCCGUGGUGGGCU GGAAGGACCUGGAGGACGGCUCCAAGUACUACUUCGACG AGGACACCGCCGAGGCCUACAUCGGCCUGUCCCUGAUCA ACGACGGCCAGUACUACUUAACGACGACGGCAUCAUGC AGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUCUACU UCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACA UCGACGACAACUACUUCUACAUCGACGACAACGGCAUCG UGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUACAAGU ACUUCGCCCCCGCCAACACCGUGAACGACAACAUCUACG GCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCG AGGACGUGUACUACUUCGGCGAGACCUACACCAUCGAGA CCGGCUGGAUCUACGACAUGGAGCAGGAGUCCGACAAGU ACUACUUAACCCCGAGACCAAGAAGGCCUGCAAGGGCA UCAACCUGAUCGACGACAUAAGUACUACUUCGACGAGA AGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAGAACA ACAACUACUACUUAACGAGAACGGCGAGAUCCAGUUCG GCUACAUAACAUCGAGGACAAGAUGUUCUACUUCGGCG AGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACCCCCG ACGGCUUCAAGUACUUCGCCACCCAGAACACCCUGGACG AGAACUUCGAGGGCGAGUCCAUCAGUACACCGGCUUGG UGGACCUGGACGAGAAGCGCUACUACUUCACCGACGAGU ACAUCGCCGCCACCGGCUCCGUGAUAUCGACGGCGAGG AGUACUACUUCGACCCCGACACCGCCAGCUGGUGAUCU CCGAGGGCGGGCUCCGGCGGGCGGGCUCCGGCCUGU CCAAGGACAUAUCAAGCUGCUGAACGAGCAGGUGAACA AGGAGAUGAACUCCUCCAACCUGUACAUGUCCAUGUCCU CCUGGUGCUACACCCACUCCUGGACGGCGCCGGCCUGU UCCUGUUCGACCACGCCGCCGAGGAGUACGAGCACGCCA	SEQ ID NO:51

	AGAAGCUGAUCAUUCUCCUGAACGAGAACAACGUGCCCG UGCAGCUGACCUCCAUCUCCGCCCCGAGCACAAGUUCG AGGGCCUGACCCAGAUCUCCAGAAGGCCUACGAGCACG AGCAGCACAUCUCCGAGUCCAUCAACAACAUUGUGGACC ACGCCAUCAAGUCCAAGGACCACGCCACCUUCAACUUC UGCAGUGGUACGUGGCCGAGCAGCACGAGGAGGAGGUGC UGUUCAAGGACAUCUCCUGGACAAGAUCGAGCUGAUCGGCA ACGAGAACCACGGCCUGUACCUGGCCGACCAGUACGUGA AGGGCAUCGCCAAGUCCCGCAAGUCCUGGAGGUGCUGU UCCAGGGCCCCGGCUCCUAAUAAACUAGUAGUGACUGAC <u>UAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCC</u> <u>GAAUGGAGUCUCUAAGCUACAUAUACCAACUUAACACU</u> <u>ACAAAAUGUUGUCCCCCAAAUGUAGCCAUUCGUAUCUG</u> <u>CUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdB_2X LSKHpylferri tin with modified IL-2 signal with 6xH tag	AGCAUAAAAGUCUCAACACAACAUUAUACAAAACAAACGA <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACC</u> AUGC GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC CCUGGUGACCAACUCCAACCUGAUCACCGGCUUCGUGA CCGUGGGCGACGACAAGUACUACUUAACCCCAUCAACG GCGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGACGACA AGAACUACUACUCCAGCAGUCCGGCGUGCUGCAGACCG GCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUCGCCC CCGCCAACACCCUGGACGAGAACCUGGAGGGCGAGGCCA UCGACUUCACCGGCAAGCUGAUCUACGACGAGAACAUUCU ACUACUUCGACGACAACUACCGCGGGCGCCGUGGAGUGGA AGGAGCUGGACGGCGAGAUGCACUACUUCUCCCCCGAGA CCGGCAAGGCCUUAAGGGCCUGAACCAAGAUCCGGCGACU ACAAGUACUACUUAACUCCGACGGCGUGAUGCAGAAGG GCUUCGUGUCCAUCAACGACAACAAGCACUACUUCGACG ACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUUCGACG GCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUCCAG UCGGCGUGUUAACACCGAGGACGGCUUCAAGUACUUCG CCCACCACAACGAGGACCUUGGGCAACGAGGAGGGCGAGG AGAUCUCCUACUCCGGCAUCCUGAACUUAACAACAAGA UCUACUACUUCGACGACUCCUACACCGCCGUGGUGGGCU GGAAGGACCUGGAGGACGGCUCCAAGUACUACUUCGACG AGGACACCGCCGAGGCCUACAUCGGCCUGUCCUGAUCU ACGACGGCCAGUACUACUUAACGACGACGGCAUCAUGC AGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUCUACU UCUCGACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACA UCGACGACAACUACUUCUACAUCGACGACAACGGCAUCG UGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUACAAGU ACUUCGCCCCCGCCAACACCGUGAACGACAACAUCUACG GCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCG AGGACGUGUACUACUUCGGCGAGACCUACACCAUCGAGA CCGGCUUGAUCUACGACAUGGAGCAGGAGUCCGACAAGU ACUACUUAACCCGAGACCAAGAAGGCCUGCAAGGGCA UCAACCUGAUCGACGACAUAAGUACUACUUCGACGAGA AGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAGAACA ACAACUACUACUUAACGAGAACGGCGAGAUGCAGUUCG	SEQ ID NO:52

	GCUACAUCAACAUCGAGGACAAGAUGUUCUACUUCGGCG AGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACCCCCG ACGGCUUCAAGUACUUCGCCACCAGAACACCCUGGACG AGAACUUCGAGGGCGAGUCCAUCCAGUACACCGGCUGG UGGACCUGGACGAGAAGCGCUACUACUUCACCGACGAGU ACAUCGCCGCCACCGGCUCCGUGAUCUACGACGGCGAGG AGUACUACUUCGACCCCCGACACCGCCCAGCUGGUGAUCU CCGAGGGCGGGCGGCUCCGGCGGGCGGCGGCUCCGGCCUGU CCAAGGACAUCAUCAAGCUGCUGAACGAGCAGGUGAACAA AGGAGAUGAACUCCUCCAACCUGUACAUGUCCAUGUCCU CCUGGUGCUACACCCACUCCUGGACGGCGCCGGCCUGU UCCUGUUCGACCACGCCGCCGAGGAGUACGAGCACGCCA AGAAGCUGAUCAUUCUCCUGAACGAGAACAACGUGCCCG UGCAGCUGACCUCUCCUCCGCCCCGAGCACAAGUUCG AGGGCCUGACCCAGAUUCUCCAGAAGGCCUACGAGCACG AGCAGCACAUCUCCGAGUCCAUAACAACAUCGUGGACC ACGCCAUCAAGUCCAAGGACCACGCCACCUUCAAACUUC UGCAGUGGUACGUGGCCGAGCAGCACGAGGAGGAGGUGC UGUUAAGGACAUCUCCUGGACAAGAUCGAGCUGAUCGGCA ACGAGAACCACGGCCUGUACCUGGCCGACCAGUACGUGA AGGGCAUCGCCAAGUCCCGCAAGUCCUGGAGGUGCUGU UCCAGGGCCCCGGCUCCACCACCACCACCACCACUAAU AAACUAGUAGUGACUGACUAGGAUCUGGUUACCACUAAA <u>CCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACA</u> <u>UAAUACCAACUUAACAUUACAAAAUGUUGUCCCCCAAAA</u> <u>UGUAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUU</u> <u>CUUCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdB_V RCHpylferri n with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCAACCUGAUCACCGGCUUC GUGACCGUGGGCGACGACAAGUACUACUUAACCCCAUC AACGGCGGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGAC GACAAGAACUACUACUUCAGCAGUCCGGCGUGCUGCAG ACCGGCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUC GCCCCGCCAACACCCUGGACGAGAACCUGGAGGGCGAG GCCAUCGACUUCACCGGCAAGCUGAUCUACGACGAGAAC AUCUACUACUUCGACGACAACUACCGCGGGCGCCGUGGAG UGGAAGGAGCUGGACGGCGAGAUACUACUUCUCCCCC GAGACCGGCAAGGCCUUAAGGGCCUGAACCAAGUACGGC GACUACAAGUACUACUUAACUCCGACGGCGUGAUGCAG AAGGGCUUCGUGUCCAUAACGACAACAAGCACUACUUC GACGACUCCGGCGUGAUGAAGGUGGGCUACACCGGAGUAC GACGGCAAGCACUUCUACUUCGCCGAGAACGGCGAGAU CAGAUCCGGCGUGUUAACACCGAGGACGGCUUCAAGUAC UUCGCCCACCACAACGAGGACCUGGGCAACGAGGAGGGC GAGGAGAUUCUCCUACUCCGGCAUCCUGAACUUAACAAC AAGAUCUACUACUUCGACGACUCCUACACCGCCGUGGUG GGCUGGAAGGACCUGGAGGACGGCUCCAAGUACUACUUC GACGAGGACACCGCCGAGGCCUACAUCGGCCUGUCCUG AUAACGACGGCGAGUACUACUUAACGACGACGGCAUC AUGCAGGUGGGCUUCGUGACCAUAACGACAAGGUGUUC UACUUCUCCGACUCCGGCAUACUACGAGUCCGGCGUGCAG AACAUACGACGACAACUACUUCUACAUCGACGACAACGGC AUCGUGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUAC	SEQ ID NO:53

	AAGUACUUCGCCCCGCAACACCGUGAACGACAACAUC UACGGCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUG GGCGAGGACGUGUACUACUUCGGCGAGACCUACACCAUC GAGACCGGCUGGAUCUACGACAUGGAGCAGGAGUCCGAC AAGUACUACUUAACCCCGAGACCAAGAAGGCCUGCAAG GGCAUCAACCUGAUCGACGACAUCAAGUACUACUUCGAC GAGAAGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAG AACAACAACUACUACUUAACGAGAACGGCGGAGAUAGCAG UUCGGCUACAUCAACAUCGAGGACAAGAUGUUCUACUUC GGCGAGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACC CCCGACGGCUUAAGUACUUCGCCCACCAGAACACCCUG GACGAGAACUUCGAGGGCGAGUCCAUCAGUACACCGGC UGGCUUGGACCUGGACGAGAAGCGCUACUACUUCACCGAC GAGUACAUCGCCGCCACCGGCUCGUGAUCUACGACGGC GAGGAGUACUACUUCGACCCCGACACCGCCAGCUGGUG AUCUCCGAGUCCGGCGGCGACAUCAUCAAGCUGCUGAAC GAGCAGGUGAACAAAGGAGAUAGCAGUCCUCCAACCUGUAC AUGUCCAUGUCCUCCUGGUGCUACACCCACUCCUGGAC GGCGCCGGCCUGUCCUGUUCGACCACGCCGCCGAGGAG UACGAGCACGCCAAGAAGCUGAUCUACUUCUGAACGAG AACAACGUGCCCGUGCAGCUGACCUCUCCUCCGCCGCC GAGCACAAGUUCGAGGGCCUGACCCAGAUUCUCCAGAAG GCCUACGAGCACGAGCAGCACAUCUCCGAGUCCAUAAC AACAUCGUGGACCACGCCAUCAAGUCCAAGGACCACGCC ACCUCAACUUCUGCAGUGGUACGUGGCCGAGCAGCAC GAGGAGGAGGUGCUGUUAAGGACAUCUGGACAAGAUC GAGCUGAUCGGCAACGAGAACCACGGCCUGUACCUGGCC GACCAGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCC GGCUCCUGGAGGUGCUGUCCAGGGCCCCGGCUCCUAA UAA	
MutTcdB_V RCHpylferri- n with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAUUUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUUGUCCUGGC CCUGGUGACCAACUCCAACCUGAUCACCGGCUUCGUGA CCGUGGGCGACGACAAGUACUACUUAACCCCAUCAACG CGGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGACGACA AGAACUACUACUUCAGCAGUCCGGCGUGCUGCAGACCG GCGUGUUCUCCACCAGGACGGCUUAAGUACUUCGCCC CCGCCAACACCCUGGACGAGAACCUGGAGGGCGAGGCCA UCGACUUCACCGGCAAGCUGAUCUACGACGAGAACAUCU ACUACUUCGACGACAACUACCGCGGCGCCGUGGAGUGGA AGGAGCUGGACGGCGAGAUACUACUUCUCCCCCGAGA CCGGCAAGGCCUUAAGGGCCUGAACCAAGAUCCGCGACU ACAAGUACUACUUAACUCCGACGGCGUGAUGCAGAAGG GCUUCGUGUCCAUCAACGACAACAAGCACUACUUCGACG ACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUCCGACG GCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUCCAGA UCGGCGUGUUAACACCGAGGACGGCUUAAGUACUUCG CCCACCACAACGAGGACCUGGGCAACGAGGAGGGCGAGG AGAUCUCCUACUCCGGCAUCCUGAACUUAACAACAAGA UCUACUACUUCGACGACUCCUUAACCGCCGUGGUGGGCU GGAAGGACCUGGAGGACGGCUCCAAGUACUACUUCGACG AGGACACCGCCGAGGCCUACAUCGGCCUGUCCUGAUCU ACGACGGCCAGUACUACUUAACGACGACGGCAUCAUGC	SEQ ID NO:54

	AGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUCUACU UCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACA UCGACGACAACUACUUCUACAUCGACGACAACGGCAUCG UGCAGAUCCGGCGUGUUCGACACCUCGACGGCUACAAGU ACUUCGCCCCCGCCAACACCGUGAACGACAACAUCUACG GCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCG AGGACGUGUACUACUUCGGCGAGACCUACACCAUCGAGA CCGGCUGGAUCUACGACAUGGAGCAGGAGUCCGACAAGU ACUACUUCAACCCCGAGACCAAGAAGGCCUGCAAGGGCA UCAACCUGAUCGACGACAUAAGUACUACUUCGACGAGA AGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAGAACA ACAACUACUACUUAACGAGAACGGCGAGAUAGCAGUUCG GCUACAUAACAUCGAGGACAAGAUGUUCUACUUCGGCG AGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACCCCCG ACGGCUUCAAGUACUUCGCCACCAGAACACCCUGGACG AGAACUUCGAGGGCGAGUCCAUCAGUACACCGGCUUGG UGGACCUGGACGAGAAGCGCUACUACUUCACCGACGAGU ACAUCGCCGCCACCGGCUCCGUGAUCUACGACGGCGAGG AGUACUACUUCGACCCCCGACACCGCCCAGCUGGUGAUCU CCGAGUCCGGCGGCGACAUAUCAAGCUGCUGAACGAGC AGGUGAACAAAGGAGAUAGCAGUCCUCCAACCUGUACAUG CCAUGUCCUCCUGGUGCUACACCCACUCCUGGACGGCG CCGGCCUGUUCUUGUUCGACCACGCCGCCGAGGAGUACG AGCACGCCAAGAAGCUGAUCUUCUCCUGAACGAGAACA ACGUGCCCGUGCAGCUGACCUCCAUCUCCGCCCCCGAGC ACAAGUUCGAGGGCCUGACCCAGAUUCUCCAGAAGGCCU ACGAGCACGAGCAGCACAUCUCCGAGUCCAUAACAACA UCGUGGACCACGCCAUCAAGUCCAAGGACCACGCCACCU UCAACUUCUGCAGUGGUACGUGGCCGAGCAGCAGCAGG AGGAGGUGCUGUUAAGGACAUCUGGACAAGAUCGAGC UGAUCGGCAACGAGAACCACGGCCUGUACCUGGCCGACC AGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCCGGCU CCCUGGAGGUGCUGUUCUCCAGGGCCCCGGCUCCUAAUAAA <u>CUAGUAGUGACUGACUAGGAUCUGGUUACCACUAAACCA</u> <u>GCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUA</u> <u>UACCAACUUAACUUAACAAAAUGUUGUCCCCCAAAUUGU</u> <u>AGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCU</u> <u>CACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
MutTcdB_V RCHpylferri n with modified IL-2 signal with 6xH tag	<u>AGCAUAAAAGUCUCAACACAACAUUAACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCCUUGUCCCUGGC CCUGGUGACCAACUCCAACCUGAUCACCGGCUUCGUGA CCGUGGGCGACGACAAGUACUACUUAACCCCAUCAACG GCGGGCGCCGCCUCCAUCGGCGAGACCAUCAUCGACGACA AGAACUACUACUUCAGCAGUCCGGCGUGCUGCAGACCG GCGUGUUCUCCACCGAGGACGGCUUCAAGUACUUCGCCC CCGCCAACACCCUGGACGAGAACCUGGAGGGCGAGGCCA UCGACUUCACCGGCAAGCUGAUCUACGACGAGAACAUUCU ACUACUUCGACGACAACUACCGCGGGCGCCGUGGAGUGGA AGGAGCUGGACGGCGAGAUACUACUUCUCCCCCGAGA CCGGCAAGGCCUUAAGGGCCUGAACAGAUCCGGCGACU ACAAGUACUACUUAACUCCGACGGCGUGAUGCAGAAGG	SEQ ID NO:55

	GCUUCGUGUCCAUCAACGACAACAAGCACUACUUCGACG ACUCCGGCGUGAUGAAGGUGGGCUACACCGAGAUCGACG GCAAGCACUUCUACUUCGCCGAGAACGGCGAGAUGCAGA UCGGCGUGUUAACACCGAGGACGGCUUCAAGUACUUCG CCCACCACAACGAGGACCUUGGGCAACGAGGAGGGCGAGG AGAUCUCCUACUCCGGCAUCCUGAACUUAACAACAAGA UCUACUACUUCGACGACUCCUUCACCGCCGUGGUGGGCU GGAAGGACCUGGAGGACGGCUCCAAGUACUACUUCGACG AGGACACCGCCGAGGCCUACAUCGGCCUGUCCCUGAUC ACGACGGCCAGUACUACUUAACGACGACGGCAUCAUGC AGGUGGGCUUCGUGACCAUCAACGACAAGGUGUUCUACU UCUCCGACUCCGGCAUCAUCGAGUCCGGCGUGCAGAACA UCGACGACAACUACUUCUACAUCGACGACAACGGCAUCG UGCAGAUCCGGCGUGUUCGACACCUCCGACGGCUACAAGU ACUUCGCCCCCGCCAACACCGUGAACGACAACAUCUACG GCCAGGCCGUGGAGUACUCCGGCCUGGUGCGCGUGGGCG AGGACGUGUACUACUUCGGCGAGACCUACACCAUCGAGA CCGGCUGGAUCUACGACAUGGAGCAGGAGUCCGACAAGU ACUACUUAACCCCGAGACCAAGAAGGCCUGCAAGGGCA UCAACCUGAUCGACGACAUAAGUACUACUUCGACGAGA AGGGCAUCAUGCGCACCGGCCUGAUCUCCUUCGAGAACA ACAACUACUACUUAACGAGAACGGCGAGAUAGCAGUUCG GCUACAUCAACAUCGAGGACAAGAUGUUCUACUUCGCG AGGACGGCGUGAUGCAGAUCCGGCGUGUUAACACCCCCG ACGGCUUCAAGUACUUCGCCCACCAGAACACCCUGGACG AGAACUUCGAGGGCGAGUCCAUCAGUACACCGGCUGGC UGGACCUGGACGAGAAGCGCUACUACUUCACCGACGAGU ACAUCGCCGCCACCGGCUCCGUGAUCUACGACGGCGAGG AGUACUACUUCGACCCCGACACCGCCCAGCUGGUGAUCU CCGAGUCCGGCGGCGACAUCAUAAGCUGCUGAACGAGC AGGUGAACAAAGGAGAUGCAGUCCUCCAACCUGUACAUGU CCAUGUCCUCCUGGUGCUACACCCACUCCUUGGACGGCG CCGGCCUGUCCUGUUCGACCACGCCGCCGAGGAGUACG AGCACGCCAAGAAGCUGAUAUCUUCUGAACGAGAACA ACGUGCCCGUGCAGCUGACCUCCAUCUCCGCCCCCGAGC ACAAGUUCGAGGGCCUGACCCAGAUCUUCAGAAAGGCCU ACGAGCACGAGCAGCACAUUCGAGUCCAUCAACAACA UCGUGGACCACGCCAUCAAGUCCAAGGACCACGCCACCU UCAACUUCUGCAGUGGUACGUGGCCGAGCAGCAGCAGG AGGAGGUGCUGUUAAGGACAUCUGGACAAGAUCGAGC UGAUCGGCAACGAGAACCACGGCCUGUACCUGGCCGACC AGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCCGGCU CCCUGGAGGUGCUGUUCAGGGCCCCGGCUCCACCACC ACCACCACCACUAAUAAACUAGUAGUGACUGACUAGGA <u>UCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUG</u> <u>GAGUCUCUAAGCUACAUAUACCAACUUAACAUUACAAA</u> <u>AUGUUGUCCCCCAAAUGUAGCCAUUCGUAUCUGCUCCU</u> <u>AAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAA</u>	
PPEP-1 RNA; (5'UTR, coding sequence,	AUGCGCAUGCAGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCGACUCCACCACCAUCCAG CAGAACAAGGACACCCUGUCCAGAUCCGUGGUGUUC ACCGGCAACUACGACAAGAACGAGGCCAACGCCAUGGUG AACCGCCUGGCCAACAUCCGACGGCAAGUACCUGAACGCC	SEQ ID NO:56

3'UTR and polyA)	CUGAAGCAGAACAACCUGAAGAUCUAGCUGCUGUCCGGC AAGCUGACCGACGAGAAGGAGUACGCCUACCUGAAGGGC GUGGUGCCCAAGGGCUGGGAGGGCACC GGCAAGACCUGG GACGACGUGCCCGGCCUGGGCGGCUCACCUGGGCCUG CGCAUCGGCUUCUCCAACAAGGGCAAGGGCCACGACGCC AUCAACCUGGAGCUGCAGACGAGACCGCCACGCCAUCGAC CACAU CGUGCUGAACGACAUCUCCAAGUCCGCCCAGUUC AAGCAGAUCUUCGCCAAGGAGGGCCGCUCCUGGGCAAC GUGAACUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCC GAGUCCUUCGCCUACUACUACCUGAACCAGGACACCAAC UCCAAGCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUC CUGCAGAACCUGGCCAAGUAAUAA	
PPEP-1 RNA; (5'UTR, coding sequence, 3'UTR and polyA)	<u>AGCAUAAAAGUCUCAACACAACAUUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCU AUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUAUGCGCAUG</u> CAGCUGCUGCUGCUGAUCGCCCUGUCCUGGGCCUGG UGACCAACUCCGACUCCACCACCAUCCAGCAGAACAAGG ACACCCUGUCCCGAUCGUGGUGUUC CCGGCAACU ACGACAAGAACGAGGCCAACGCCAUGGUGAACCGCCUGG CCAACAUCGACGGCAAGUACCUGAACGCCUGAAGCAGA ACAACCUGAAGAUCAAGCUGCUGUCCGGCAAGCUGACCG ACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGUGCCCA AGGGCUGGGAGGGCACCGGCAAGACCUGGGACGACGUGC CCGGCCUGGGCGGCUCACCUGGGCCUGCGCAUCGGCU UCUCCAACAAGGGCAAGGGCCACGACGCCAUCAACCUGG AGCUGCACGAGACCGCCCACGCCAUCGACCACAUCGUGC UGAACGACAUCUCCAAGUCCGCCCAGUUCAAGCAGAUUC UCGCCAAGGAGGGCCGCUCCUGGGCAACGUGAACUACC UGGGCGUGUACCCCGAGGAGUUCUUCGCCGAGUCCUUCG CCUACUACUACCUGAACCAGGACACCAACUCCAAGCUGA AGUCCGCCUGCCCCCAGACCUACUCCUUCUGCAGAACC UGGCCAAGUAAUAAACUAGUAGUGACUGACUAGGAUCU <u>GGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAG</u> <u>UCUCUAAGCUACAUAUACCAACUUAACAUUACAAAAAUG</u> <u>UUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAU</u> <u>AAAAAGAAAGUUUCUUCACAUUCUAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAA</u>	SEQ ID NO:57
PPEP-1_ Hpylferritin mRNA with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGGCCUGGUGACCAACUCCGACUCCACCACCAUCCAG CAGAACAAGGACACCCUGUCCCGAUCGUGGUGUUC CCG ACCGGCAACUACGACAAGAACGAGGCCAACGCCAUGGUG AACC GCCUGGCCAACAU CGACGGCAAGUACCUGAACGCC CUGAAGCAGAACAACCUGAAGAUCUAGCUGCUGUCCGGC AAGCUGACCGACGAGAAGGAGUACGCCUACCUGAAGGGC GUGGUGCCCAAGGGCUGGGAGGGCACC GGCAAGACCUGG GACGACGUGCCCGGCCUGGGCGGCUCACCUGGGCCUG CGCAUCGGCUUCUCCAACAAGGGCAAGGGCCACGACGCC AUCAACCUGGAGCUGCAGACGAGACCGCCACGCCAUCGAC CACAU CGUGCUGAACGACAUCUCCAAGUCCGCCCAGUUC AAGCAGAUCUUCGCCAAGGAGGGCCGCUCCUGGGCAAC	SEQ ID NO:58

	GUGAACUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCC GAGUCCUUCGCCUACUACUACCUGAACCAGGACACCAAC UCCAAGCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUC CUGCAGAACCUGGCCAAGGGCUCCGGCGACAUCAUCAAG CUGCUGAACGAGCAGGUGAACAAGGAGAUGAACUCCUCC AACCUGUACAUGUCCAUGUCCUCCUGGUGCUACACCCAC UCCCUGGACGGCGCCGGCCUGUUCUCCUGUUCGACCACGCC GCCGAGGAGUACGAGCACGCCAAGAAGCUGAUCUUCUUC CUGAACGAGAACAACGUGCCCGUGCAGCUGACCUCCAUC UCCGCCCCCGAGCACAAGUUCGAGGGCCUGACCCAGAUC UCCAGAAGGCCUACGAGCACGAGCAGCACAUCUCCGAG UCCAUCAACAACAUCGUGGACCACGCCAUCAAGUCCAAG GACCACGCCACCUUCAACUUCUGCAGUGGUACGUGGCC GAGCAGCACGAGGAGGAGGUGCUGUUCAAGGACAUCUCCUG GACAAGAUCGAGCUGAUCGGCAACGAGAACCACGGCCUG UACCUGGCCGACCAGUACGUGAAGGGCAUCGCCAAGUCC CGCAAGUCCUGGAGGUGCUGUUCCAGGGCCCCGGCUCC UAAUAA	
PPEP-1_ HpyIferritin mRNA with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUUGUCCUGGC CCUGGUGACCAACUCCGACUCCACCACCAUCCAGCAGAA CAAGGACACCCUGUCCCAGAUCGUGGUGUUCUCCACCCGG CAACUACGACAAGAACGAGGCCAACGCCAUGGUGAACCG CCUGGCCAACAUACGACGGCAAGUACCUGAACGCCUGAA GCAGAACAAACCUGAAGAUCAAGCUGCUGUCCGGCAAGCU GACCGACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGU GCCCAAGGGCUGGGAGGGCACC GGCAAGACCUGGGACGA CGUGCCCGGCCUGGGCGGCCUCCACCUGGGCCUUGCGCAU CGGCUUCUCCAACAAGGGCAAGGGCCACGACGCCAUCAA CCUGGAGCUGCACGAGACCGCCACGCCAUCGACCACAU CGUGCUGAACGACAUCUCCAAGUCCGCCAGUUCAAGCA GAUCUUCGCCAAGGAGGGCCGCUCCUGGGCAACGUGAA CUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCCGAGUC CUUCGCCUACUACUACCUGAACCAGGACACCAACUCCAA GCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUCUGCA GAACCUGGCCAAGGGCUCCGGCGACAUCAUAAGCUGCU GAACGAGCAGGUGAACAAGGAGAUGAACUCCUCCAACCU GUACAUGUCCAUGUCCUCCUGGUGCUACACCCACUCCCU GGACGGCGCCGGCCUGUUCUUGUUCGACCACGCCGCCGA GGAGUACGAGCACGCCAAGAAGCUGAUCUUCUCCUGAA CGAGAACAAACGUGCCCGUGCAGCUGACCUCCAUCUCCGC CCCCGAGCACAAGUUCGAGGGCCUGACCCAGAUCUCCA GAAGGCCUACGAGCACGAGCAGCACAUCUCCGAGUCCA CAACAACAUCGUGGACCACGCCAUCAAGUCCAAGGACCA CGCCACCUUCAACUCCUGCAGUGGUACGUGGCCGAGCA GCACGAGGAGGAGGUGCUGUUCAAGGACAUCUCCUGGACAA GAUCGAGCUGAUCGGCAACGAGAACCACGGCCUGUACCU	SEQ ID NO:59

	<u>GGCCGACCAGUACGUGAAGGGCAUCGCCAAGUCCCGCAA</u> <u>GUCCCUUGGAGGUGCUGUUC CAGGGCCCCGGCUCCUAAUA</u> <u>AACUAGUAGUGACUGACUAGGAUCUGGUUACCACUAAAC</u> <u>CAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAU</u> <u>AAUACCAACUUACACUUACAAAAUGUUGUCCCCCAAAU</u> <u>GUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUC</u> <u>UUCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
PPEP-1_ HpyI ferritin mRNA with modified IL-2 signal with 6xH tag	<u>AGCAUAAAAGUCUCAACACAACAUUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCU AUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> <u>GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC</u> <u>CCUGGUGACCAACUCCGACUCCACCACCAUCCAGCAGAA</u> <u>CAAGGACACCCUGUCCCAGAUCGUGGUGUUC CCCCACCGG</u> <u>CAACUACGACAAGAACGAGGCCAACGCCAUGGUGAACCG</u> <u>CCUGGCCAACAU CGACGGCAAGUACCUGAACGCCUGAA</u> <u>GCAGAACAAACCUGAAGAUCAAGCUGCUGUCCGGCAAGCU</u> <u>GACCGACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGU</u> <u>GCCCAAGGGCUGGGAGGGCACCGGCAAGACCUGGGACGA</u> <u>CGUGCCCGGCCUGGGCGGCUC CACCGUGGCCUGCGCAU</u> <u>CGGCUUCUCCAACAAGGGCAAGGGCCACGACGCCAUCAA</u> <u>CCUGGAGCUGCAGACGAGACCGCCACGCCAUCGACCACAU</u> <u>CGUGCUGAACGACAUUC CCAAGUCCGCCAGUUCAAGCA</u> <u>GAUCUUCGCCAAGGAGGGCCGCUCCUGGGCAACGUGAA</u> <u>CUACCUGGGCGUGUAC CCGGAGGAGUUCUUCGCCGAGUC</u> <u>CUUCGCCUACUACUACCUGAACCAGGACACCAACUCCAA</u> <u>GCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUCUGCA</u> <u>GAACCUGGCCAAGGGCUCCGGCGACAUCAUCAAGCUGCU</u> <u>GAACGAGCAGGUGAACAAAGGAGAUGAACUCCUCCAACCU</u> <u>GUACAUGUCCAUGUCCUCCUGGUGCUACACCCACUCCCU</u> <u>GGACGGCGCCGGCCUGUUC CUGUUCGACCACGCCGCCGA</u> <u>GGAGUACGAGCACGCCAAGAAGCUGAUCAUCUUCUGAA</u> <u>CGAGAACAAACGUGCCCGUGCAGCUGACCUCCAUCUCCGC</u> <u>CCCCGAGCACAAAGUUCGAGGGCCUGACCCAGAUCUCCA</u> <u>GAAGGCCUACGAGCACGAGCAGCACAUUC CGAGUCCA</u> <u>CAACAACAUUCGUGGACCACGCCAUCAAGUCCAAGGACCA</u> <u>CGCCACCUUCAACUCCUGCAGUGGUACGUGGCCGAGCA</u> <u>GCACGAGGAGGAGGUGCUGUUCAAGGACAUC CUGGACAA</u> <u>GAUCGAGCUGAUCGGCAACGAGAACACGGCCUGUACCU</u> <u>GGCCGACCAGUACGUGAAGGGCAUCGCCAAGUCCCGCAA</u> <u>GUCCCUUGGAGGUGCUGUUC CAGGGCCCCGGCUCCACCA</u> <u>CCACCACCACCACUAAUAAACUAGUAGUGACUGACUAG</u> <u>GAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA</u> <u>UGGAGUCUCUAAGCUACAUAUACCAACUUAACAUACA</u> <u>AAUUGUUGUCCCCCAAAUUGUAGCCAUUCGUAUCUGCUC</u> <u>CUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	SEQ ID NO:60

	<u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAA</u>	
PPEP-1_2XLSKHp ylferritin mRNA with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCGACUCCACCACCAUCCAG CAGAACAAGGACACCCUGUCCCAGAUCGUGGUGUCCCC ACCGGCAACUACGACAAGAACGAGGCCAACGCCAUGGUG AACCGCCUGGCCAACAUCCGACGGCAAGUACCUGAACGCC CUGAAGCAGAACAACCUGAAGAUAAGCUGCUGUCCGGC AAGCUGACCGACGAGAAGGAGUACGCCUACCUGAAGGGC GUGGUGCCCAAGGGCUGGGAGGGCACCGGCAAGACCUGG GACGACGUGCCCGGCCUGGGCGGCUCACCUGGGCCCUG CGCAUCGGCUUCUCCAACAAGGGCAAGGGCCACGACGCC AUCAACCUGGAGCUGCACGAGACCGCCCACGCCAUCGAC CACAUCGUGCUGAACGACAUCUCCAAGUCCGCCCAGUUC AAGCAGAUUCUCCGAAGGAGGGCCGCUCCUGGGCAAC GUGAACUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCC GAGUCCUUCGCCUACUACUACCUGAACCAGGACACCAAC UCCAAGCUGAAGUCCGCCUGCCCCAGACCUACUCCUUC CUGCAGAACCUGGCCAAGGGCGGGCGGCUCGCGCGGGC GGCUCCGGCCUGUCCAAGGACAUCAUCAAGCUGCUGAAC GAGCAGGUGAACAAAGGAGAUGAACUCCUCCAACCUGUAC AUGUCCAUGUCCUCCUGGUGCUACACCCACUCCUGGAC GGCGCCGGCCUGUUCUGUUCGACCACGCCGCCGAGGAG UACGAGCACGCCAAGAAGCUGAUAUCUUCUGAACGAG AACAACGUGCCCGUGCAGCUGACCUCCAUCUCCGCCCCC GAGCACAAGUUCGAGGGCCUGACCCAGAUCUCCAGAAG GCCUACGAGCACGAGCAGCACAUCCGAGUCCAUAAC AACAUCGUGGACCACGCCAUCAAGUCCAAGGACCACGCC ACCUUCAACUUCUGCAGUGGUACGUGGGCCGAGCAGCAC GAGGAGGAGGUGCUGUUAAGGACAUCUGGACAAGAUC GAGCUGAUCGGCAACGAGAACCACGGCCUGUACCUGGCC GACCAGUACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCC CUGGAGGUGCUGUCCAGGGCCCCGCUCCUAUAUA	SEQ ID NO:61
PPEP-1_2XLSKHp ylferritin mRNA with modified IL-2 signal	AGCAUAAAAGUCUACAACACAACAUAAUACAAAACAAACGA AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU UUAUAUCAAUUUCUUUUAAGCAAAAGCAAUUUCUGAA AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC GCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCCUGGC CCUGGUGACCAACUCCGACUCCACCACCAUCCAGCAGAA CAAGGACACCCUGUCCCAGAUCGUGGUGUCCCCACCGG CAACUACGACAAGAACGAGGCCAACGCCAUGGUGAACCG CCUGGCCAACAUCCGACGGCAAGUACCUGAACGCCUGAA GCAGAACAAACCUGAAGAUAAGCUGCUGUCCGGCAAGCU GACCGACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGU GCCCAAGGGCUGGGAGGGCACCGGCAAGACCUGGGACGA CGUGCCCGGCCUGGGCGGCUCACCUGGGCCCUGCGCAU CGGCUUCUCCAACAAGGGCAAGGGCCACGACGCCAUCAA CCUGGAGCUGACGAGACCGCCCACGCCAUCGACCAU CGUGCUGAACGACAUCUCCAAGUCCGCCCAGUUAAGCA GAUCUUCGCCAAGGAGGGCCGCUCCUGGGCAACGUGAA CUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCCGAGUC CUUCGCCUACUACUACCUGAACCAGGACACCAACUCCAA GCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUCUGCA GAACCUGGCCAAGGGCGGGCGGCUCGCGCGGGCGGCUC	SEQ ID NO:62

	CGGCCUGUCCAAGGACAUCAUCAAGCUGCUGAACGAGCA GGUGAACAAGGAGAUGAACUCCUCCAACCUGUACAUGUC CAUGUCCUCCUGGUGCUACACCCACUCCUGGACGGCGC CGGCCUGUCCUGUUCGACCACGCCGCCGAGGAGUACGA GCACGCCAAGAAGCUGAUCUUCUCCUGAACGAGAACAA CGUGCCCGUGCAGCUGACCUCUCCUCCGCCCCCGAGCAC AAGUUCGAGGGCCUGACCCAGAUUCCAGAAGGCCUAC GAGCACGAGCAGCACAUUCUCCGAGUCCAUAACAACAUC GUGGACCACGCCAUCAAGUCCAAGGACCACGCCACCUUC AACUUCUGCAGUGGUACGUGGCCGAGCAGCAGGAG GAGGUGCUGUUAAGGACAUCUGGACAAGAUCGAGCUG AUCGGCAACGAGAACCACGGCCUGUACCUGGCCGACCAG UACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCCUGGAG GUGCUGUCCAGGGCCCCGGCUCCUAAUAAACUAGUAGU <u>GACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAG</u> <u>AACACCCGAAUGGAGUCUCUAAGCUACAUAUACCAACU</u> <u>UACACUUAACAAAUGUUGUCCCCCAAAUAGGCCAUUC</u> <u>GUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
PPEP- 1_2XLKHp ylferritin mRNA with modified IL-2 signal with 6xH tag	AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAUAUCAAUUUUAUUAAGCAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC CCUGGUGACCAACUCCGACUCCACCACCAUCCAGCAGAA CAAGGACACCCUGUCCAGAUUCGUGGUGUUCUCCACCCGG CAACUACGACAAGAACGAGGCCAACGCCAUGGUGAACCG CCUGGCCAACAUUCGACGGCAAGUACCUGAACGCCUGAA GCAGAACAAACCUGAAGAUAAGCUGCUGUCCGGCAAGCU GACCGACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGU GCCAAGGGCUGGGAGGGCACCGGCAAGACCUGGGACGA CGUGCCCGGCCUGGGCGGCUCACCUGGGCCUGCGCAU CGGCUUCUCCAACAAGGGCAAGGGCCACGACGCCAUCAA CCUGGAGCUGCAGACGACCGCCACGCCAUCGACCAU CGUGCUGAACGACAUCUCCAAGUCCGCCAGUUAAGCA GAUCUUCGCCAAGGAGGGCGCUCCUGGGCAACGUGAA CUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCCGAGUC CUUCGCCUACUACUACCUGAACCAGGACACCAACUCCAA GCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUCUGCA GAACCUGGCCAAGGGCGGGCGGCUCCGGCGGGCGGCGCUC CGGCCUGUCCAAGGACAUCAUCAAGCUGCUGAACGAGCA GGUGAACAAGGAGAUGAACUCCUCCAACCUGUACAUGUC CAUGUCCUCCUGGUGCUACACCCACUCCUGGACGGCGC CGGCCUGUCCUGUUCGACCACGCCGCCGAGGAGUACGA GCACGCCAAGAAGCUGAUCUUCUCCUGAACGAGAACAA CGUGCCCGUGCAGCUGACCUCUCCGCCCCCGAGCAC AAGUUCGAGGGCCUGACCCAGAUUCCAGAAGGCCUAC GAGCACGAGCAGCACAUUCUCCGAGUCCAUAACAACAUC GUGGACCACGCCAUCAAGUCCAAGGACCACGCCACCUUC AACUUCUGCAGUGGUACGUGGCCGAGCAGCAGGAG GAGGUGCUGUUAAGGACAUCUGGACAAGAUCGAGCUG AUCGGCAACGAGAACCACGGCCUGUACCUGGCCGACCAG UACGUGAAGGGCAUCGCCAAGUCCCGCAAGUCCUGGAG	SEQ ID NO:63

	GUGCUGUCCAGGGCCCCGGCUCACCACCACCACCAC CACUAAUAAACUAGUAGUGACUGACUAGGAUCUGGUUA CCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCU AAGCUACAUAUACCAACUACACUUAACAAAUGUUGUC CCCCAAAUGUAGCCAUUCGUUUCUGCUCCUAAUAAAAA GAAAGUUUCUACAUUCUAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAA	
PPEP-1 _VRCHpylfe ritin mRNA with modified IL-2 signal	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCGACUCCACCACCAUCCAG CAGAACAAGGACACCCUGUCCCAGAUUCGUGGUGUUC ACCGGCAACUACGACAAGAACGAGGCCAACGCCAUGGUG AACCGCCUGGCCAACAUUCGACGGCAAGUACCUGAACGCC CUGAAGCAGAACACCUGAAGAUCAAGCUGCUGUCCGGC AAGCUGACCGACGAGAAGGAGUACGCCUACCUGAAGGGC GUGGUGCCCCAAGGGCUGGGAGGGCACCGGCAAGACCUGG GACGACGUGCCCGGCCUGGGCGGCUCACCUGGGCCCUG CGCAUCGGCUUCUCCAACAAGGGCAAGGGCCACGACGCC AUAACCUGGAGCUGCACGAGACCGCCACGCCAUCGAC CACAUCGUGCUGAACGACAUCUCCAAGUCCGCCAGUUC AAGCAGAUUCUUCGCCAAGGAGGGCCGCUCCUGGGCAAC GUGAACUACCUGGGCGUGUACCCCGAGGAGUUCUUCGCC GAGUCCUUCGCCUACUACUACCUGAACCCAGGACACCAAC UCCAAGCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUC CUGCAGAACCUGGCCAAGUCCGGCGGCGACAUAUCAAG CUGCUGAACGAGCAGGUGAACAAAGGAGAUGCAGUCCUCC AACCGUACAUGUCCAUGUCCUCCUGGUGCUACACCCAC UCCUGGACGGCGCCGGCCUGUCCUGUUCGACCACGCC GCCGAGGAGUACGAGCACGCCAAGAAGCUGAUCAUCUUC CUGAACGAGAACAACGUGCCCGUGCAGCUGACCUCCAUC UCCGCCCCCGAGCACAAGUUCGAGGGCCUGACCCAGAUC UCCAGAAGGCCUACGAGCACGAGCAGCACAUCUCCGAG UCCAUCAACAACAUCGUGGACCACGCCAUAAGUCCAAG GACCACGCCACCUCAACUCCUGCAGUGGUACGUGGCC GAGCAGCACGAGGAGGAGGUGCUGUUAAGGACAUCUG GACAAGAUCGAGCUGAUCGGCAACGAGAACCACGGCCUG UACCUGGCCGACCAGUACGUGAAGGGCAUCGCCAAGUCC CGCAAGUCCGGCUCUCCUGGAGGUGCUGUCCAGGGCCCC GGCUCCUAAUAA	SEQ ID NO:64
PPEP-1 _VRCHpylfe ritin mRNA with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAUAUCAUUUCUUUUAAAGCAAAAGCAAUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCCUGGC CCUGGUGACCAACUCCGACUCCACCACCAUCCAGCAGAA CAAGGACACCCUGUCCCAGAUUCGUGGUGUUCUCCACCGG CAACUACGACAAGAACGAGGCCAACGCCAUGGUGAACCG CCUGGCCAACAUUCGACGGCAAGUACCUGAACGCCUGAA GCAGAACAAACCUGAAGAUCAAGCUGCUGUCCGGCAAGCU GACCGACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGU GCCCAAGGGCUGGGAGGGCACCGGCAAGACCUGGGGACGA CGUGCCCGGCCUGGGCGGCUCACCUGGGCCCUGCGCAU CGGCUUCUCCAACAAGGGCAAGGGCCACGACGCCAUCAA CCUGGAGCUGCACGAGACCGCCACGCCAUCGACCACAU	SEQ ID NO:65

	CGUGCUGAACGACAUCUCCAAGUCCGCCAGUUCAAGCA GAUCUUCGCCAAGGAGGGCCGCUCCCUGGGCAACGUGAA CUACCUUGGGCGUGUACCCCGAGGAGUUCUUCGCCGAGUC CUUCGCCUACUACUACCUGAACCAGGACACCAACUCCAA GCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUCUGCA GAACCUUGGCAAGUCCGGCGGCGACAUCAUAAGCUGCU GAACGAGCAGGUGAACAAGGAGAUGCAGUCCUCCAACCU GUACAUGUCCAUGUCCUCCUGGUGCUACACCCACUCCCU GGACGGCGCCGCCUGUUCUGUUCGACCACGCCGCCGA GGAGUACGAGCACGCCAAGAAGCUGAUCAUCUUCUGAA CGAGAACAACGUGCCCGUGCAGCUGACCUCCAUCUCCGC CCCCGAGCACAAGUUCGAGGGCCUGACCCAGAUCUCCA GAAGGCCUACGAGCACGAGCAGCACAUCUCCGAGUCCA CAACAACAUCGUGGACCACGCCAUCAAGUCCAAGGACCA CGCCACCUUCAACUCCUGCAGUGGUACGUGGCCGAGCA GCACGAGGAGGAGGUGCUGUUCAAGGACAUCUGGACAA GAUCGAGCUGAUCGGCAACGAGAACCACGGCCUGUACCU GGCCGACCAGUACGUGAAGGGCAUCGCCAAGUCCCGCAA GUCCGGCUCCCUUGGAGGUGCUGUUCAGGGCCCCGGCUC CUAAUAAACUAGUAGUGACUGACUAGGAUCUGGUUACCA <u>CUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAA</u> <u>GUACAUAUACCAACUACACUUAACAAAUGUUGUCCCC</u> <u>CAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAA</u> <u>AGUUUCUUCACAUUCUAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAA</u>	
PPEP-1 _VRCHpylfe ritin mRNA with modified IL-2 signal	<u>AGCAUAAAAGUCUCAACACAACAUUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAUAUCAAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUUGC</u> <u>GCAUGCAGCUGCUGCUGCUGAUCGCCCUUGUCCCUGGC</u> <u>CCUGGUGACCAACUCCGACUCCACCACCAUCCAGCAGAA</u> CAAGGACACCCUGUCCCGAUCGUGGUGUUCCCCACCGG CAACUACGACAAGAACGAGGCCAACGCCAUGGUGAACCG CCUGGCCAACAUACGACGGCAAGUACCUGAACGCCCGAA GCAGAACAACCUGAAGAUCAAGCUGCUGUCCGGCAAGCU GACCGACGAGAAGGAGUACGCCUACCUGAAGGGCGUGGU GCCCAAGGGCUGGGAGGGCACCAGCAAGACCUGGGACGA CGUGCCCGGCCUGGGCGGCUCACCUGUGGCCUUGCGCAU CGGCUUCUCCAACAAGGGCAAGGGCCACGACGCCAUCAA CCUGGAGCUGCACGAGACCGCCACGCCAUCGACCACAU CGUGCUGAACGACAUCUCCAAGUCCGCCAGUUCAAGCA GAUCUUCGCCAAGGAGGGCCGCUCCCUGGGCAACGUGAA CUACCUUGGGCGUGUACCCCGAGGAGUUCUUCGCCGAGUC CUUCGCCUACUACUACCUGAACCAGGACACCAACUCCAA GCUGAAGUCCGCCUGCCCCCAGACCUACUCCUUCUGCA GAACCUUGGCAAGUCCGGCGGCGACAUCAUAAGCUGCU GAACGAGCAGGUGAACAAGGAGAUGCAGUCCUCCAACCU GUACAUGUCCAUGUCCUCCUGGUGCUACACCCACUCCCU GGACGGCGCCGCCUGUUCUGUUCGACCACGCCGCCGA GGAGUACGAGCACGCCAAGAAGCUGAUCAUCUUCUGAA CGAGAACAACGUGCCCGUGCAGCUGACCUCCAUCUCCGC CCCCGAGCACAAGUUCGAGGGCCUGACCCAGAUCUCCA GAAGGCCUACGAGCACGAGCAGCACAUCUCCGAGUCCA CAACAACAUCGUGGACCACGCCAUCAAGUCCAAGGACCA	SEQ ID NO:66

	CGCCACCUUCAACUCCUGCAGUGGUACGUGGCCGAGCA GCACGAGGAGGAGGUGCUGUUAAGGACAUCCUGGACAA GAUCGAGCUGAUCGGCAACGAGAACCACGGCCUGUACCU GGCCGACCAGUACGUGAAGGGCAUCGCCAAGUCCCCGAA GUCCGGCUCCCUGGAGGUGCUGUUCAGGGCCCCGGCUC <u>CCACCACCACCACCACCACUAAUAAACUAGUAGUGACU</u> <u>GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACA</u> <u>CCCGAAUGGAGUCUCUAAGCUACAUAUACCAACUUACA</u> <u>CUUACAAAAUGUUGUCCCCCAAAUAGCCAUUCGUAU</u> <u>CUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
Wild type (WT) CdeM mRNA sequence without secretion signal (exosporium morphogenic protein CdeM)	AUGGAGAACAAGAAGUACGCCAACGGCGGCUACUCCGAG GACUGGUACGAGCGCGGCGAGUCCACCGCCAAGUGGUUC CAGAACGACCGCGAGGAGUACGAGCGCGCCGCCUACGAC GAGGACCGCGAGCGCCGCGGCUCCAACUGCGGCGUCUCC GACUCCGGCGGCAACCGCCCCCGCAACUGCGAGAGGUUC CGCCGCGAGGCGGAGAUCCGCGAGCGCGAGGCCCCGCGAG GCCUUCUGCGAGUCCUCCGAGCGCAAGAAGGAGGCCUG GCCUACGAGUGCGAGGCCCGCAAGCUGUGGGAGGAGGCC GAGAAGUACUGGGACGAGUACUCCAAGUACAACUACAAG GGCAUCGAGUACCUGGCCGAGGCCGCCGCCUGUUCGAC GAGGGCAUGGAGUGCGAGGCCCGCCGCAACGGCAACAAC GGCGGCAACAACAACAACUGCUGCCACAAGUGCCACAAG UGCAACUGCAACUGCUGCCGCAAGUAAUAA	SEQ ID NO:67
Wild type (WT) CdeM mRNA sequence without secretion signal (exosporium morphogenic protein CdeM)	AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUAUGGAGAACA</u> AGAAGUACGCCAACGGCGGCUACUCCGAGGACUGGUACG AGCGCGGCGAGUCCACCGCCAAGUGGUUCCAGAACGACC GCGAGGAGUACGAGCGCGCCGCCUACGACGAGGACCGCG AGCGCCGCGGCUCCAACUGCGGCGUCUCCGACUCCGGCG GCAACCGCCCCCGCAACUGCGAGAGGUUCCGCCGCGAGG CCGAGAUCCGCGAGCGCGAGGCCCGCGAGGCCUUCUGCG AGUCCUCCGAGCGCAAGAAGGAGGCCUUGGCCUACGAGU GCGAGGCCCCGCAAGCUGUGGGAGGAGGCCGAGAAGUACU GGGACGAGUACUCCAAGUACAACUACAAGGGCAUCGAGU ACCUGGCCGAGGCCGCCGCCUGUUCGACGAGGGCAUGG AGUGCGAGGCCCGCCGCAACGGCAACAACGGCGGCAACA ACAACAACUGCUGCCACAAGUGCCACAAGUGCAACUGCA ACUGCUGCCGCAAGUAAUAAACUAGUAGUGACUGACUAG <u>GAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA</u> <u>UGGAGUCUCUAAGCUACAUAUACCAACUUACACUUACA</u> <u>AAUUGUUGUCCCCCAAAUAGCCAUUCGUUUCUGCUC</u> <u>CUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAA</u>	SEQ ID NO:68

Wild type (WT) CdeM mRNA sequence with the modified IL-2 secretion signal	AUGCGCAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCGAGAACAAGAAGUACGCC AACGGCGGCUACUCCGAGGACUGGUACGAGCGCGGCGAG UCCACCGCCAAGUGGUUCCAGAACGACCGCGAGGAGUAC GAGCGCGCCGCCUACGACGAGGACCGCGAGCGCCGCGGC UCCAACUGCGGCGUCUCCGACUCCGGCGGCAACCGCCCCC GCAACUGCGAGAGGUUCCGCCGCGAGGCCGAGAUCCGCG AGCGCGAGGCCCCGCGAGGCCUUCUGCGAGUCCUCCGAGC GCAAGAAGGAGGCCCUAGGCUACGAGUGCGAGGCCCGCA AGCUGUGGGAGGAGGCCGAGAAGUACUGGGACGAGUAC UCCAAGUACAACUACAAGGGCAUCGAGUACCUUGGCCGAG GCCGCCCGCCUGUUCGACGAGGGCAUGGAGUGCGAGGCC CGCCGCAACGGCAACAACGGCGGCAACAACAACACUGC UGCCACAAGUGCCACAAGUGCAACUGCAACUGCUGCCGC AAGUAAUAA	SEQ ID NO:69
Wild type (WT) CdeM mRNA sequence with the modified IL-2 secretion signal	AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAUAUCAAUUUCUUUUAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUAUGCGCAUG</u> CAGCUGCUGCUGCUGAUCGCCCUGUCCCUGGCCUGG UGACCAACUCCGAGAACAAGAAGUACGCCAACGGCGGC UACUCCGAGGACUGGUACGAGCGCGGCGAGUCCACCGCC AAGUGGUUCCAGAACGACCGCGAGGAGUACGAGCGCGCC GCCUACGACGAGGACCGCGAGCGCCGCGGCUCCAACUGC GGCUGCUCCGACUCCGGCGGCAACCGCCCCCGCAACUGC GAGAGGUUCCGCCGCGAGGCCGAGAUCCGCGAGCGCGAG GCCCCGAGGGCCUUCUGCGAGUCCUCCGAGCGCAAGAAG GAGGCCCUGGCCUACGAGUGCGAGGCCCGCAAGCUGUGG GAGGAGGCCGAGAAGUACUGGGACGAGUACUCCAAGUAC AACUACAAGGGCAUCGAGUACCUUGGCCGAGGCCGCCCGC CUGUUCGACGAGGGCAUGGAGUGCGAGGCCCGCCGCAAC GGCAACAACGGCGGCAACAACAACAACUGCUGCCACAAG UGCCACAAGUGCAACUGCAACUGCUGCCGCAAGUAAUAA <u>ACUAGUAGUGACUGACUAGGAUCUGGUUACCACUAAACC</u> <u>AGCCUCAAGAACACCCGAAUGGAGUCUCUAAAGCUACAUA</u> <u>AUACCAACUUACACUUACAAAAUGUUGUCCCCCAAAUUG</u> <u>UAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCU</u> <u>UCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	SEQ ID NO:70
Modified CdeM mRNA sequence without secretion signal	AUGGAGAACAAGAAGUACGCCAACGGCGGCUACUCCGAG GACUGGUACGAGCGCGGCGAGUCCACCGCCAAGUGGUUC CAGAACGACCGCGAGGAGUACGAGCGCGCCGCCUACGAC GAGGACCGCGAGCGCCGCGGCUCCAGUGCGGCUGCUCC GACUCCGGCGGCAACCGCCCCCGCAACUGCGAGAGGUUC CGCCGCGAGGCCGAGAUCCGCGAGCGCGAGGCCCGCGAG GCCUUCUGCGAGUCCUCCGAGCGCAAGAAGGAGGCCUG	SEQ ID NO:71

	GCCUACGAGUGCGAGGCCCGCAAGCUGUGGGAGGAGGCC GAGAAGUACUGGGACGAGUACUCCAAGUACCAGUACAAG GGCAUCGAGUACCUGGCCGAGGCCGCCGCCUGUUCGAC GAGGGCAUGGAGUGCGAGGCCCGCCGCAACGGCAACAAC GGCGGCAACAACAACAACUGCUGGCCACAAGUGCCACAAG UGCAACUGCAACUGCUGGCCGCAAGUAAUAA	
Modified CdeM mRNA sequence without secretion signal	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAUUUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUAUGGAGAACA</u> AGAAGUACGCCAACGGCGGCUACUCCGAGGACUGGUACG AGCGCGGCGAGUCCACCGCCAAGUGGUUCCAGAACGACC GCGAGGAGUACGAGCGCGCCGCCUACGACGAGGACCGCG AGCGCCGCGGCUCCAGUGCGGCUUGCUCCGACUCCGGCG GCAACCGCCCCCGCAACUGCGAGAGGUUCCGCCGCGAGG CCGAGAUCCGCGAGCGCGAGGCCCGCGAGGCCUUCUGCG AGUCCUCCGAGCGCAAGAAGGAGGCCCUUGGCCUACGAGU GCGAGGCCCGCAAGCUGUGGGAGGAGGCCGAGAAGUACU GGGACGAGUACUCCAAGUACCAGUACAAGGGCAUCGAGU ACCUGGCCGAGGCCGCCGCCUGUUCGACGAGGGCAUGG AGUGCGAGGCCCGCCGCAACGGCAACAACGGCGGCAACA ACAACAACUGCUGGCCACAAGUGCCACAAGUGCAACUGCA ACUGCUGCCGCAAGUAAUAAACUAGUAGUGACUGACUAG <u>GAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA</u> <u>UGGAGUCUCUAAGCUACAUAUACCAACUUAACACUACA</u> <u>AAAUUGUUGUCCCCCAAAAUAGCCAUUCGUAUCUGCUC</u> <u>CUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAA</u>	SEQ ID NO:72
Modified CdeM mRNA sequence with the modified IL-2 secretion signal	AUGCGCAUGCAGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCGAGAACAAGAAGUACGCC AACGGCGGCUACUCCGAGGACUGGUACGAGCGCGGCGAG UCCACCGCCAAGUGGUUCCAGAACGACCGCGAGGAGUAC GAGCGCGCCGCCUACGACGAGGACCGCGAGCGCCGCGGC UCCCAGUGCGGCUUGCUCCGACUCCGGCGGCAACCGCCCC GCAACUGCGAGAGGUUCCGCCGCGAGGCCGAGAUCGCG AGCGCGAGGCCCGCGAGGCCUUCUGCGAGUCCUCCGAGC GCAAGAAGGAGGCCCUUGGCCUACGAGUGCGAGGCCCGCA AGCUGUGGGAGGAGGCCGAGAAGUACUGGGACGAGUAC UCCAAGUACCAGUACAAGGGCAUCGAGUACCUGGCCGAG GCCGCCCGCCUGUUCGACGAGGGCAUGGAGUGCGAGGCC CGCCGCAACGGCAACAACGGCGGCAACAACAACAACUGC UGCCACAAGUGCCACAAGUGCAACUGCAACUGCUGCCGC AAGUAAUAA	SEQ ID NO:73
Modified CdeM mRNA sequence with the modified IL-2	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAUUUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUAUGCGCAUG</u> CAGCUGCUGCUGCUGAUCGCCCUGUCCCUGGCCUGG UGACCAACUCCGAGAACAAGAAGUACGCCAACGGCGGC	SEQ ID NO:74

secretion signal	UACUCCGAGGACUGGUACGAGCGCGGCGAGUCCACCGCC AAGUGGUUCCAGAACGACCGCGAGGAGUACGAGCGCGCC GCCUACGACGAGGACCGCGAGCGCCGCGGCUCCAGUGC GGCUGCUCCGACUCCGGCGGCAACCGCCCCGCAACUGC GAGAGGUUCCGCCGCGAGGCCGAGAUCCGCGAGCGCGAG GCCCCGCGAGGCCUUCUGCGAGUCCUCCGAGCGCAAGAAG GAGGCCUGGCCUACGAGUGCGAGGCCCGCAAGCUGUGG GAGGAGGCCGAGAAGUACUGGGACGAGUACUCCAAGUAC CAGUACAAGGGCAUCGAGUACCUGGCCGAGGCCGCCCGC CUGUUCGACGAGGGCAUGGAGUGCGAGGCCCGCCGCAAC GGCAACAACGGCGGCAACAACAACUGCUGCCACAAG UGCCACAAGUGCAACUGCAACUGCUGCCGCAAGUAAUAA <u>ACUAGUAGUGACUGACUAGGAUCUGGUUACCACUAAACC</u> <u>AGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUA</u> <u>AUACCAACUUACACUUACAAAAUGUUGUCCCCCAAAAUG</u> <u>UAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCU</u> <u>UCACAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
CWP84 mRNA or cell surface protein cwp84 mRNA	AUGCGGAAGUACAAGUCCAAGAAGCUGUCCAAGCUGCUG GCCCUGCUGACCGUGUGCUUCCUGAUCGUGUCCACCAUC CCCGUGUCCGCCGAGAACCACAAGACCCUGGACGGCGUG GAGACCGCCGAGUACUCCGAGUCCUACCUGCAGUACCUG GAGGACGUGAAGAACGGCGACACCGCCAAGUACAACGGC GUGAUCCCCUUCUCCCCACGAGAUUGGAGGGCACCACCCUG CGGAACAAGGGCCGGUCCUCCUGGCCUCCGCCUACAAG UCCUCCGUGGCCUACAACCCCAUGGACCUGGGCCUGACC ACCCCCGCCAAGAACCAGGGCUCCUGAACACCUGCUGG UCCUUCUCCGGCAUGUCCACCCUGGAGGCCUACCUGAAG CUGAAGGGCUACGGCACCUACGACCUGUCCGAGGAGCAC CUGCGGUGGUGGGCCACCGGCGGCAAGUACGGCUGGAAC CUGGACGACAUGUCCGGCUCCUCCAACGUGACCGCCAUC GGCUACCUGACCGCCUGGGCCGGCCCCAAGCUGGAGAAG GACAUCCCCUACAACCUGAAGUCCGAGGCCCAGGGCGCC ACCAAGCCCUCCAACAUGGACACCGCCCCACCCAGUUCA ACGUGACCGACGUGGUGCGGCUGAACAAGGACAAGGAGA CCGUGAAGAACGCCAUCAUGCAGUACGGCUCCGUGACCU CCGGCUACGCCACUACUCCACCUACUUAACAAGGACG AGACCGCCUACAACUGCACCAACAAGCGGGCCCCCUGA ACCACGCCGUGGCCAUCGUGGGCUGGGACGACAACUACU CCAAGGACAACUUCGCCUCCGACGUGAAGCCCCGAGUCCA ACGGCGCCUGGCUGGUGAAGUCCUCCUGGGGCGAGUUA ACUCCAUGAAGGGCUUCUUCUGGAUCUCCUACGAGGACA AGACCCUGCUGACCGACACCGACAACUACGCCAUGAAGU CCGUGUCCAAGCCCGACUCCGACAAGAAGAUGUACCAGC UGGAGUACGCCGGCCUGUCCAAGAUAUGUCCAACAAGG UGACCGCCGCCAACGUGUUCGACUUCUCCGGGACUCCG AGAAGCUGGACUCCGUGAUGUUCGAGACCGACUCCGUGG GCGCCAAGUACGAGGUGUACUACGCCCCGUGGUGAACG GCGUGCCCCAGAACACUCCAUGACCAAGCUGGCCUCCG GCACCGUGUCCUACUCCGGCUACAUCAACGUGCCCACCA	SEQ ID NO:75

	ACUCCUACUCCUGCCCAAGGGCAAGGGCGCCAUCGUGG UGGUGAUCGACAACACCGCCAACCCCAACCGGGAGAAGU CCACCCUGGCCUACGAGACCAACAUCGACGCCUACUACC UGUACGAGGCCAAGGCCAACCUGGGCGAGUCCUACAUC UGCAGAACAACAAGUUCGAGGACAUCAACACCUACUCCG AGUUCUCCCCUGCAACUUCGUGAUAAGGCCAUCACCA AGACCUCCUCCGGCCAGGCCACCUCCGGCGAGUCCUGA CCGGCGCCGACCGGUACGAGACCGCCGUGAAGGUGUCCC AGAAGGGCUGGACCUCCUCCAGAACGCCGUGCUGGUGA ACGGCGACGCCAUCGUGGACGCCUGACCGCCACCCCU UCACCGCCGCCAUCGACUCCCCAUCCUGCUGACCGGCAA GGACAACCUGGACUCCAAGACCAAGGCCGAGCUGCAGCG GCUGGGCACCAAGAAGGUGUACCUGAUCGGCGGCGAGAA CUCCUGUCCAAGAAGCUGCAGACCCAGCUGUCCAACAU GGGCAUCUCCGUGGAGCGGAUCUCCGGCUCCGACCGGUA CAAGACCUCCAUCUCCUGGCCCAGAAGCUGAACUCCA CAAGUCCGUGUCCAGGUGGCCGUGGCCAACGGCGUGAA CGGCCUGGCCGACGCCAUCUCCGUGGGCGCCGCCGCCGCC GACAACAACAUGCCCAUCAUCCUGACCAACGAGAAGUCC GAGCUGCAGGGCGCCGACGAGUCCUGAACUCCUCCAAG AUCACCAAGUCCUACAUAUCGGCGGCACCGCCACCCUG UCCUCCAACCUGGAGUCCAAGCUGUCCAACCCACCCGG CUGGCCGGCUCCAACCGGAACGAGACCAACGCCAAGAUC AUCGACAAGUUCUACCCUCCUCCGACCUGAAGUACGCC UUCGUGGUGAAGGACGGCUCCAAGUCCAGGGCGACCUG AUCGACGGCCUGGCCGUGGGCGCCUGGGCGCCAAGACC GACUCCCCCGUGGUGCUGGUGGGCAACAAGCUGGACGAG UCCCAGAAGAACGUGCUGAAGUCCAAGAAGAUCCGAGACC CCCAUCCGGGUGGGCGGCAACGGCAACGAGUCCGCCUUC AACGAGCUGAACACCCUGCUGGGCAAGUAAUAA	
CWP84 mRNA or cell surface protein cwp84 mRNA	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAUAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GGAAGUACAAGUCCAAGAAGCUGUCCAAGCUGCUGGGCC UGCUGACCGUGUGCUUCCUGAUCGUGUCCACCAUCCCCG UGUCCGCCGAGAACCACAAGACCCUGGACGGCGUGGAGA CCGCCGAGUACUCCGAGUCCUACCUGCAGUACCUGGAGG ACGUGAAGAACGGCGACACCGCCAAGUACAACGGCGUGA UCCCCUCCCCCACGAGAUGGAGGGCACCAACCCUGCGGA ACAAGGGCCGGUCCUCCUCCUGCCUCCGCCUACAAGUCCU CCGUGGCCUACAACCCCAUGGACCUGGGCCUGACCACCC CCGCCAAGAACCAGGGCUCCUGAACACCUGCUGGUCCU UCUCCGGCAUGUCCACCCUGGAGGCCUACCUGAAGCUGA AGGGCUACGGCACCUACGACCUGUCCGAGGAGCACCUGC GGUGGUGGGCCACCGGCGGCAAGUACGGCUGGAACCUGG ACGACAUGUCCGGCUCCUCCAACGUGACCGCCAUCGGCU ACCUGACCGCCUGGGCCGGCCCAAGCUGGAGAAGGACA UCCCCUACAACCUGAAGUCCGAGGCCACGGCGCCACCA	SEQ ID NO:76

AGCCCUCCAACAUGGACACCGCCCCACCCAGUUCAACG UGACCGACGUGGUGCGGCUGAACAAGGACAAGGAGACCG UGAAGAACGCCAUCAUGCAGUACGGCUCCGUGACCUCCG GCUACGCCCACUACUCCACCUACUUCAACAAGGACGAGA CCGCCUACAACUGCACCAACAAGCGGGCCCCCUGAACC ACGCCGUGGCCAUUCGUGGGCUGGGACGACAACUACUCCA AGGACAACUUCGCCUCCGACGUGAAGCCCGAGUCCAACG GCGCCUGGCUGGUGAAGUCCUCCUGGGGCGAGUUCAACU CCAUGAAGGGCUUCUUCUGGAUCUCCUACGAGGACAAGA CCCGUGUGACCGACACCGACAACUACGCCAUGAAGUCCG UGUCCAAGCCCGACUCCGACAAGAAGAUGUACCAGCUGG AGUACGCCGGCCUGUCCAAGAUCAUGUCCAACAAGGUGA CCGCCGCCAACGUGUUCGACUUCUCCCGGGACUCCGAGA AGCUGGACUCCGUGAUGUUCGAGACCGACUCCGUGGGCG CCAAGUACGAGGUGUACUACGCCCCCGUGGUGAACGGCG UGCCCCAGAACAACUCCAUGACCAAGCUGGCCUCCGGCA CCGUGUCCUACUCCGGCUACAUCAACGUGGCCACCAACU CCUACUCCUGCCCAAGGGCAAGGGCGCCAUCGUGGUGG UGAUCGACAACACCGCCAACCCCAACCGGGAGAAGUCCA CCCGUGGCCUACGAGACCAACAUCGACGCCUACUACCUGU ACGAGGCCAAGGCCAACCUGGGCGAGUCCUACAUCUGC AGAACACAAGUUCGAGGACAUAACACCUACUCCGAGU UCUCCCCUGCAACUUCGUGAUCAAGGCCAUCACCAAGA CCUCCUCCGGCCAGGCCACCUCGGCGAGUCCUGACCGG CGCCGACCGGUACGAGACCGCCGUGAAGGUGUCCAGAA GGGUGGACCUCCUCCAGAACGCCGUGCUGGUGAACGG CGACGCCAUCGUGGACGCCUGACCGCCACCCCUUACCC GCCGCCAUCGACUCCCCCAUCCUGCUGACCGGCAAGGAC AACCUGGACUCCAAGACCAAGGCCGAGCUGCAGCGGCUG GGCACCAAGAAGGUGUACCUGAUCGGCGGCGAGAACUCC CUGUCCAAGAACGUGCAGACCCAGCUGUCCAACAUGGGC AUCUCCGUGGAGCGGAUCUCCGGCUCCGACCGGUACAAG ACCUCCAUCUCCUGGGCCAGAAGCUGAACUCCAUAAG UCCGUGUCCAGGUGGGCCGUGGCCAACGGCGUGAACGGC CUGGGCCGACGCCAUCUCCGUGGGCGCCGCCGCCGCCGAC AACAACAUGCCCAUCAUCCUGACCAACGAGAAGUCCGAG CUGCAGGGCGCCGACGAGUUCUGAACUCCUCCAAGAUC ACCAAGUCCUACAUCAUCGGCGGCACCGCCACCCUGUCC UCCAACCUGGAGUCCAAGCUGUCCAACCCCAACCGGCUG GCCGGCUCCAACCGGAACGAGACCAACGCCAAGAUAUC GACAAGUUCUACCCCUCCUCCGACCUGAAGUACGCCUUC GUGGUGAAGGACGGCUCCAAGUCCCAGGGCGACCUGAUC GACGGCCUGGCCGUGGGCGCCCUGGGCGCCAAGACCGAC UCCCCCGUGGUGCUGGUGGGCAACAAGCUGGACGAGUCC CAGAAGAACGUGCUGAAGUCCAAGAAGAUUCGAGACCC AUCCGGGUGGGCGGCAACGGCAACGAGUCCGCCUUAAC GAGCUGAACACCCUGCUGGGCAAGUAAUAAACUAGUAGU <u>GACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAG</u> <u>AACACCCGAAUGGAGUCUCUAAGCUACAUAUACCAACU</u> <u>UACACUUAACAAAUGUUGUCCCCCAAAUGUAGCCAUUC</u> <u>GUUUCGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
--	--

	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
modified CWP84 mRNA [Clostridioide s difficile]	AUGCGGAAGUACAAGUCCAAGAAGCUGUCCAAGCUGCUG GCCCUGCUGACCGUGUGCUUCCUGAUCGUGUCCACCAUC CCCGUGUCCGCCGAGAACCACAAGACCCUGGACGGCGUG GAGACCGCCGAGUACUCCGAGUCCUACCUGCAGUACCUG GAGGACGUGAAGAACGGCGACACCGCCAAGUACAACGGC GUGAUCCCUUCCCCCACGAGAUGGAGGGCACCACCCUG CGAACAAGGGCCGGUCCUCCUGGCCUCCGCCUACAAG UCCUCCGUGGCCUACAACCCAUGGACCUGGGCCUGACC ACCCCCGCCAAGAACCAGGGCUCCUGAACACCUGCUGG UCCUUCUCCGGCAUGUCCACCCUGGAGGCCUACCUGAAG CUGAAGGGCUACGGCACCUACGACCUGUCCGAGGAGCAC CUGCGUGGUGGGCCACCGGCGGCAAGUACGGCUGGAAC CUGGACGACAUGUCCGGCUCCUCCAGGUGACCGCCAUC GGCUACCUGACCGCCUGGGCCGGCCCCAAGCUGGAGAAG GACAUCCCUACAACCUGAAGUCCGAGGCCCAGGGCGCC ACCAAGCCCUCCAACAUGGACACCGCCCCACCCAGUCC AGGUGACCGACGUGGUGCGGCUGAACAAAGGACAAGGAGA CCGUGAAGAACGCCAUCAUGCAGUACGGCUCCGUGACCU CCGGCUACGCCACUACUCCACCUACUUAACAAGGACG AGACCGCCUACCAGUGCACCAACAAGCGGGCCCCCUGA ACCACGCCGUGGCCAUCGUGGGCUGGGACGACCAGUACU CCAAGGACAACUUCGCCUCCGACGUGAAGCCCCGAGUCCA ACGGCGCCUGGCUGGUGAAGUCCUCCUGGGGCGAGUUA ACUCCAUGAAGGGCUUCUUCUGGAUCUCCUACGAGGACA AGACCCUGCUGACCGACACCGACAACUACGCCAUGAAGU CCGUGUCCAAGCCCCGACUCCGACAAGAAGAUUACCAGC UGGAGUACGCCGGCCUGUCCAAGAUAUGUCCAACAAGG UGACCGCCGCCAACGUGUUCGACUUCUCCGGGACUCCG AGAAGCUGGACUCCGUGAUGUUCGAGACCGACUCCGUGG GCGCCAAGUACGAGGUGUACUACGCCCCCGUGGUGAACG GCGUGCCCCAGAACAACUCCAUGACCAAGCUGGCCUCCG GCACCGUGUCCUACUCCGGCUACAUAACGUGCCCACCA ACUCCUACUCCUGCCCCAAGGGCAAGGGCGCCAUCGUGG UGGUGAUCGACAACACCGCCAACCCCAACCGGGAGAAGU CCACCCUGGCCUACGAGACCAACAUCGACGCCUACUACC UGUACGAGGCCAAGGCCAACCUGGGCGAGUCCUACAUC UGCAGAACAACAAGUUCGAGGACAUAACACCUACUCCG AGUUCUCCCCUGCAACUUCGUGAUAAGGCCAUACCA AGACCUCCUCCGGCCAGGCCACCUCCGGCGAGUCCUGA CCGGCGCCGACCGGUACGAGACCGCCGUGAAGGUGUCCC AGAAGGGCUGGACCUCCUCCAGAACGCCGUGCUGGUGA ACGGCGACGCCAUCGUGGACGCCUGACCGCCACCCCU UCACCGCCGCCAUCGACUCCCCCAUCCUGCUGACCGGCAA GGACAACCGGACUCCAAGACCAAGGCCGAGCUGCAGCG GCUGGGCACCAAGAAGGUGUACCUGAUCGGCGGCGAGAA CUCCUGUCCAAGAACGUGCAGACCCAGCUGUCCAACAU GGGCAUCUCCGUGGAGCGGAUCUCCGGCUCCGACCGGUA CAAGACCUCCAUCUCCUGGCCCAGAAGCUGAACUCCA CAAGUCCGUGUCCAGGUGGCCGUGGCCAACGGCGUGAA CGGCCUGGCCGACGCCAUCUCCGUGGGCGCCGCCGCCG GACAACAACAUGCCCAUAUCCUGACCAACGAGAAGUCC	SEQ ID NO:77

	GAGCUGCAGGGCGCCGACGAGUUCCUGAACUCCUCCAAG AUCACCAAGUCCUACAUAUCGGCGGCACCGCCACCCUG UCCUCCAACCUGGAGUCCAAGCUGUCCAACCCACCCGG CUGGCCGGCUCCAACCGGAACGAGACCAACGCCAAGAUC AUCGACAAGUUCUACCCUCCUCCGACCUGAAGUACGCC UUCGUGGUGAAGGACGGCUCCAAGUCCCAGGGCGACCUG AUCGACGGCCUGGCCGUGGGCGCCUUGGGCGCCAAGACC GACUCCCCCGUGGUGCUGGUGGGCAACAAGCUGGACGAG UCCCAGAAGAACGUGCUGAAGUCCAAGAAGAUCCGAGACC CCCAUCCGGGUGGGCGGCAACGGCCAGGAGUCCGCCUUC AACGAGCUGAACACCCUGCUGGGCAAGUAAUAA	
modified CWP84 mRNA [Clostridioide s difficile]	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> GGAAGUACAAGUCCAAGAAGCUGUCCAAGCUGCUGGGCC UGCUGACCGUGUGCUUCCUGAUCGUGUCCACCAUCCCCG UGUCCGCCGAGAACCACAAGACCCUGGACGGCGUGGAGA CCGCCGAGUACUCCGAGUCCUACCUGCAGUACCUGGAGG ACGUGAAGAACGGCGACACCGCCAAGUACAACGGCGUGA UCCCCUCCCCACGAGAUGGAGGGCACCACCCUGCGGA ACAAGGGCCGGUCCUCCUCCCCUCCGCCUACAAGUCCU CCGUGGCCUACAACCCCAUGGACCUGGGCCUGACCACC CCGCCAAGAACCAGGGCUCCUGAACACCUGCUGGUCCU UCUCCGGCAUGUCCACCCUGGAGGCCUACCUGAAGCUGA AGGGCUACGGCACCUACGACCUGUCCGAGGAGCACCUGC GGUGGUGGGCCACCGGCGGCAAGUACGGCUGGAACCUGG ACGACAUGUCCGGCUCCUCCAGGUGACCGCCAUCGGCU ACCUGACCGCCUGGGCCGGCCCCAAGCUGGAGAAGGACA UCCCCUACAACCUGAAGUCCGAGGCCCAGGGCGCCACCA AGCCCUCCAACAUGGACACCGCCCCCACCAGUUCAGG UGACCGACGUGGUGCGGCUGAACAAAGGACAAGGAGACCG UGAAGAACGCCAUCAUGCAGUACGGCUCCGUGACCUCCG GCUACGCCACUACUCCACCUACUUCACAACAAGGACGAGA CCGCCUACCAGUGCACCAACAAGCGGGCCCCCUGAACC ACGCCGUGGCCAUUCGUGGGCUGGGACGACCAGUACUCCA AGGACAACUUCGCCUCCGACGUGAAGCCCAGUCCAACG GCGCCUGGCUUGUGAAGUCCUCCUGGGCGAGUUCAACU CCAUGAAGGGCUUCUUCUGGAUCUCCUACGAGGACAAGA CCCUGCUGACCGACACCGACAACUACGCCAUGAAGUCCG UGUCCAAGCCCGACUCCGACAAGAAGAUUACCAGCUGG AGUACGCCGGCCUGUCCAAGAUCAUGUCCAACAAGGUGA CCGCCGCCAACGUGUUCGACUUCUCCCGGGACUCCGAGA AGCUGGACUCCGUGAUGUUCGAGACCGACUCCGUGGGCG CCAAGUACGAGGUGUACUACGCCCCCGUGGUGAACGGCG UGCCCCAGAACACUCCAUGACCAAGCUGGGCCUCCGGCA CCGUGUCCUACUCCGGCUACAUAACGUGGCCACCAACU CCUACUCCUGCCCAAGGGCAAGGGCGCCAUCGUGGUGG UGAUCGACAACACCGCCAACCCCAACCGGGAGAAGUCCA CCCUGGCCUACGAGACCAACAUCGACGCCUACUACCUGU ACGAGGCCAAGGCCAACCUGGGCGAGUCCUACAUCUGC AGAACACAAGUUCGAGGACAUAACACCUACUCCGAGU UCUCCCCUGCAACUUCGUGAUAAGGCCAUACCAAGA CCUCCUCCGGCCAGGCCACCUCGGCGAGUCCUGACCGG CGCCGACCGGUACGAGACCGCCGUGAAGGUGUCCAGAA GGGCUGGACCUCUCCAGAACGCCGUGCUGGUGAACGG	SEQ ID NO:78

	CGACGCCAUCGUGGACGCCUGACCGCCACCCCUUCACC GCCGCCAUCGACUCCCCCAUCCUGCUGACCGGCAAGGAC AACCUGGACUCCAAGACCAAGGCCGAGCUGCAGCGGCUG GGCACCAAGAAGGUGUACCUGAUCGGCGGCAGAACUCC CUGUCCAAGAACGUGCAGACCCAGCUGUCCAACAUGGGC AUCUCCGUGGAGCGGAUCUCCGGCUCCGACCGGUACAAG ACCUCCAUCUCCCUUGGCCCAGAAGCUGAACUCCAUAAG UCCGUGUCCAGGUGGGCCGUGGCCAACGGCGUGAACGGC CUGGCCGACGCCAUCUCCGUGGGCGCCGCCGCCGCCGAC AACAAACAUGCCCAUCAUCCUGACCAACGAGAAGUCCGAG CUGCAGGGCGCCGACGAGUUCUGAACUCCUCCAAGAUC ACCAAGUCCUACAUCUCCGCGGCACCGCCACCCUGUCC UCCAACCUGGAGUCCAAGCUGUCCAACCCCAACCGGCUG GCCGGCUCCAACCGGAACGAGACCAACGCCAAGAUAUC GACAAGUUCUACCCUCCUCCGACCUGAAGUACGCCUUC GUGGUGAAGGACGGCUCCAAGUCCCAGGGCGACCUGAUC GACGGCCUGGCCGUGGGCGCCUGGGCGCCAAGACCGAC UCCCCCGUGGUGCUGGUGGGCAACAAGCUGGACGAGUCC CAGAAGAACGUGCUGAAGUCCAAGAAGAUCCGAGACCC AUCCGGGUGGGCGGCAACGGCCAGGAGUCCGCCUUAAC GAGCUGAACACCCUGCUGGGCAAGUAAUAA<u>ACUAGUAGU</u> <u>GACUGACUAGGAUCUGGUUACCACUAAACCAGCCUAAG</u> <u>AACACCCGAAUGGAGUCUCUAAGCUACAUAUACCAACU</u> <u>UACACUUACAAAAUGUUGUCCCCCAAAUGUAGCCAUUC</u> <u>GUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
Modified CWP66 mRNA	AUGAACAUCCUCCAAGCAGAUCCGUGUCCCUGCUGACCAUG ACCUUCCUGACCGUGACCCUGUACGGCAACACCUCCAAC GCCUCCACCAAGGACACCCUGACCGGCUCGCGCCGGUGG GAGACCGCCAUCAAGAUCCUCCAGGCCGGCUGGACCAAG UCCGAGUCCGCCGUGCUGGUGAACGACAACUCCAUCGCC GACGCCCUGUCCGCCACCCCUUCGCCAAGGCCAAGGAC GCCCCCAUCCUGCUGACCCAGUCCAACAAGCUGGACUCC CGGACCAAGGCCGAGCUGAAGCGGCUGGGCGUGAAGAAC GUGUACCUGAUCGGCGGCUCUACGCCUUGUCCUCCGAG AUCGAGAAGCAGCUGAACGCCGAGAACAUCAACUUCGAG CGGAUCUCCGGCAACUCCCGGUACGACACCUCCUGAAG CUGGCCGAGAAGCUGGACCGGGAGAAGUCCAUCUCCAAG AUCGUGGUGGUGAACGGCGAGAAGGGCCUGGCCGACGCC GUGUCCGUGGGCGCCAUCGCCGCCAGGAGAACAUGCCC AUCAUCCUGUCCGACUCCGAGAACGGCACCGAGGUGGCC GACAACUUCUACGACUCCAAGGACAUCGCCAAGUCCUAC GUGAUCGGCGGCACCUACUCCAUCUCCAACUCCGUGGAG CGGUCCUGCCCCAGGCCACCCGGAUCGCCGGCUCCUCC GGUCCGAGACCAACGCCAAGAUCAUCGAGGAGUUCUACA AGGACACCGACAUCAAGAACAUCUACGUGACCAAGGACG GCACCAAGAACAAGAACGACCUGAUCGACUCCUGGCCG UGGGCGUGCUGGCCGCCAAGAACUCCUCCCCCAUCGUGC UGGCCGGCAACAAGCUGGACACCACCCAGAAGGACGUGC UGAACACCAAGAUCAUCGACAAGGUGACCCAGAUCGGCG	SEQ ID NO:79

	GCCUGGGCAACGAGAACGUGGUGGAGGACAUCCUGGACA UCCAGGAGGAGACCAAGUACACCGUGGAGACCAUCGACG AGCUGAACGCCGCCAUCAAGCGGGCCGACGCCAACGACA UCAUCAAGUUCAAGCCCGAGAAGGAGAAGACCAUCAACA ACUCCUUCUCCAUCGAGACCAAGAAGACCGUGACCAUCG AGCUGGACGGCCGGUACCGGCAGACCAUACCCUGGACA UCCCCAACGGCAAGUUCAACAACUACGCCGAGAUCGAGG GCGGCGUGAAGCUGAAGAACAUCAAGAACGAGUCCCUUG UGAACAAGGGCUCCAUCCAGGACCUGGACAUCUACGACG AGAACGGCUGCAAGAU CGAGAACGAGUCCUCCGGCGAGA UCUGGUUCGUGACCAUCGUGGAGGAGGCCAACGACGUGU ACAUCGUGAACUCCGGCGACAUCACCAAGAUCUCCAACA ACUCCUCCUCCACCAUCAUCCGGAACUCCGGCAACAUCG ACACCGUGACCGGCAAGAAGGAGCCCGCCAUCUCCGGCA ACAAGCCCAAGGUGAACGACACCGAGAAGGAGACCAAGG CCGCCCCGGGGCCUGAACCCCCGGGUGGAGGCCUGCUCCG UGCCCAAGAAGGACUACGUGAUGAUCACCAUCCCCAACU CCCCCAAGGACUCCCGGUACAAGAUCUACUACCGGGUGG UGUACAACAAGCCUACGCCAUGGACGUGGGCGACAAGA UCAACAUCGGCGAGUGGACCGUGGCCCCCACCAGCAGG AGCCCUUCCUGGAGAAGGCCAAGAACGGCUGCUACGUGG AGGCCGUGGAGGUGAACACCUCCACCAAGGAGGUGUCCC GGUGGGGCCGGACCCAGGCCACCGACGACGGCUUCUAAU AA	
Modified CWP66 mRNA	<u>AGCAUAAAAGUCUCAACACAACAUUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u> <u>UUAUUUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGA</u> ACAUCUCCAAGCAGAUCGUGUCCUGCUGACCAUGACCU UCCUGACCGUGACCCUGUACGGCAACACCUCCAACGCCU CCACCAAGGACACCCUGACCGGCUCGCGCCGGUGGGAGA CCGCCAUCAAGAUUCCCAAGGCCGGCUGGACCAAGUCCG AGUCCGCCGUGCUGGUGAACGACAACUCCAUCGCCGACG CCCUGUCCGCCACCCCUUCGCCAAGGCCAAGGACGCCCC CAUCCUGCUGACCCAGUCCAACAAGCUGGACUCCCGGAC CAAGGCCGAGCUGAAGCGGCUGGGCGUGAAGAACGUGUA CCUGAUCGGCGGCUCCAUCGCCCUGUCCUCCGAGAUCA GAAGCAGCUGAACGCCGAGAACAUCAACUUCGAGCGGAU CUCCGGCAACUCCCGGUACGACACCUCCUGAAGCUGGC CGAGAAGCUGGACCGGGAGAAGUCCAUCUCCAAGAUCGU GGUGGUGAACGGCGAGAAGGGCCUGGCCGACGCCGUGUC CGUGGGCGCCAUCGCCGCCAGGAGAACAUGCCCAUCAU CCUGUCCGACUCCGAGAACGGCACCGAGGUGGCCGACAA CUUCAUCGACUCCAAGGACAUCGCCAAGUCCUACGUGAU CGGCGGCACCUACUCCAUCUCCAACUCCGUGGAGCGGUC CCUGCCCCAGGCCACCCGGAUCGCCGGCUCCUCCCGGUCC GAGACCAACGCCAAGAUAUCGAGGAGUUCUACAAGGAC ACCGACAUCAAGAACAUCUACGUGACCAAGGACGGCACC AAGAACAAGAACGACCUGAUCGACUCCCUGGCCGUGGGC GUGCUGGCCGCCAAGAACUCCUCCCCCAUCGUGCUGGCC GGCAACAAGCUGGACACCAACCCAGAAGGACGUGCUGAAC	SEQ ID NO:80

	ACCAAGAUCAUCGACAAGGUGACCCAGAUCCGGCGGCCUG GGCAACGAGAACGUGGUGGAGGACAUCUGGACAUCACAG GAGGAGACCAAGUACACCGUGGAGACCAUCGACGAGCUG AACGCCGCCAUCAAGCGGGCCGACGCCAACGACAUCauc AAGUUCAAGCCCGAGAAGGAGAAGACCAUCAACAACUCC UUCUCCAUCGAGACCAAGAAGACCGUGACCAUCGAGCUG GACGGCCGGUACCGGCAGACCAUCACCCUGGACAUCCCC AACGGCAAGUUCAACAACUACGCCGAGAUCGAGGGCGGC GUGAAGCUGAAGAACAUCAAGAACGAGUCCUGGUGAAC AAGGGCUCCAUCACAGGACCGUGGACAUCUACGACGAGAAC GGCUGCAAGAUCGAGAACGAGUCCUCCGGCGAGAUCUGG UUCGUGACCAUCGUGGAGGAGGCCAACGACGUGUACAUC GUGAACUCCGGCGACAUCACCAAGAUCUCCAACAACUCC UCCUCCACCAUCAUCCGGAACUCCGGCAACAUCGACACC GUGACCGGCAAGAAGGAGCCCGCCAUCUCCGGCAACAAG CCCAAGGUGAACGACACCGAGAAGGAGACCAAGGCCGCC CGGGGCCUGAACCCCGGGUGGAGGCCUGCUCCGUGCCC AAGAAGGACUACGUGAUGAUCACCAUCCCCAACUCCCC AAGGACUCCCGGUACAAGAUCUACUACCGGGUGGUGUAC AACAAGCCCUACGCCAUGGACGUGGGCGACAAGAUAAC AUCGGCGAGUGGACCGUGGCCCCACCGACGAGGAGCCC UUCUGGAGAAGGCCAAGAACGGCUGCUACGUGGAGGCC GUGGAGGUGAACACCUCCACCAAGGAGGUGUCCCGGUGG GGCCGGACCCAGGCCACCGACGACGGCUUCUAAUAA<u>ACU</u> <u>AGUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGC</u> <u>CUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAUA</u> <u>CCAACUUAACAUACA AAAAUGUUGUCCCCCAAAUGUAG</u> <u>CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUA</u> <u>CAUUCUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
ZupT mRNA	AUGAACAACGUGAUCUUCGCCUUCUGAUCACCCUGCUG GCCGGCCUGUCCACCGGCAUCGGCUCCUGCAUCGCCUUC UUCGCCAAGAAGACCAACACCAAGUCCUGUCCAUCUCC CUGGGCUUCUCCGCCGGCGUGAUGAUCUACGUGUCCAUG AUCGAGAUUCUCCCCAAGGCCAGGACGCCUGACCAAG UCCAUGGGCGAGAAGCUGGGCUCCUGGUCCACCGUGAUC GCCUUCUUAUCGGCAUGGCCAUCAUCGCCGUGAUCGAC AAGCUGAUCCCCCAGGAGGAGAACCCCCACGAGAUCAAG AAGAUGGAGAACAUCGAGGAGAAGAACAUAAGAAGAA CAAGUCCUGUGCGGACCGGCAUCUUCACCGCCAUGGC CAUCGCCAUCCACAACUCCCCGAGGGCCUGGCCACCUU CAUCUCCGCCUGGACGACGUGACCAUCGCCAUCCCCAU CGCCAUCGCCAUCCACAACAUCCCCGAGGGCAU CUCCGUGUCCGUGCCCGUGUACUACGCCACCGGCGACAA GAAGAAGGCCUUCUACUACUCCUCCUGUCCGGCAUGUC CGAGCCCCUGGGCGCCAUCAUCGGCUACGUGCUGCUGCG GAACUCCUGAACGACAUCACCCUGGGCAUCGUGUUCGC CAUCGUGGGCGGAUCAUGGUGUUAUCUCCUGGACGA GCUGCUGCCUCCGCCGGGAGUACGGCGAGCACCAU GUCCAUCUACGGCCUGAUCGCCGGCAUGGGCGUGAUGGC CAUCUCCUGCUGCUGUUAAGUAAUA	SEQ ID NO: 81

ZupT mRNA	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGA</u> ACAACGUGAUCUUCGCCUUCUGAUCACCCUGCUGGCCG GCCUGUCCACCGGCAUCGGCUCCUGCAUCGCCUUCUUCG CCAAGAAGACCAACACCAAGUCCUGUCCAUCUCCUGG GCUUCUCCGCCGGCGUGAUGAUCUACGUGUCCAUGAUCG AGAUCUUCCTCAAGGCCAGGACGCCUGACCAAGUCCA UGGGCGAGAAGCUGGGCUCCUGGUCCACCGUGAUCGCCU UCUUCAUCGGCAUGGCCAUCAUCGCCGUGAUCGACAAGC UGAUCCCCCAGGAGGAGAACCCACGAGAUCAAGAAGA UGGAGAACAUCGAGGAGAAGAACAUCAAGAAGAACAAG UCCCUGCUGCGGACCGGCAUCUUCACCGCCAUGGCCAUC GCCAUCCACAACUUCCTCGAGGGCCUGGCCACCUUCAUC UCCGCCUGGACGACGUGACCAUCGCCAUCCCCAUCGCC AUCGCCAUCCCAUCCACAACAUCCCCGAGGGCAUCUCC GUGUCCGUGCCCGUGUACUACGCCACCGGCGACAAGAAG AAGGCCUUCUACUACUCCUCCUGUCCGGCAUGUCCGAG CCCCUGGGCGCCAUCAUCGGCUACGUGCUGCUGCGGAAC UUCCUGAACGACAUCACCCUGGGCAUCGUGUUCGCCAUC GUGGGCGGCAUCAUGGUGUUCUUCUCCUGGACGAGCUG CUGCCCUCCGCCCGGAGUACGGCGAGCACCAACUGUCC AUCUACGGCCUGAUCGCCGGCAUGGGCGUGAUGGCCAUC UCCCUGCUGCUGUUAAGUAAUAAACUAGUAGUGACUGA <u>CUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACC</u> <u>CGAAUGGAGUCUCUAAAGCUACAUAUACCAACUACACU</u> <u>UACAAAAGUUGUCCCCCAAAUGUAGCCAUCGUAUCU</u> <u>GCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAA</u>	SEQ ID NO: 82
Bile acid germinant receptor pseudoprotea se CspC mRNA	AUGGAGAAGUCCUACUGCAUCAUCUACCAGGGCGACAUC GAGUCCGCCUGCAGGAGAACGGCAUCAACCGCUACAUG GUGCUGAACUCCAGCUGGCCGUGAUCUACGUGCCCCUG GACUUCGACGAGACCAUCCUGAACAAUCAUCCAGGUG GCCUGGUGGGAGGAGUCCGAGCCCAUGUCCUCCUGAUC GAGAUACCAACAACGUGAACAAACGGCGAGACCAUACC ACCGCCGCCGAGACCGACUACAUCUACGAGAACCCCUAC AACGACAUCACCGGCCCGGCAUCCUGCUGGCCGUGAUC GACUCCGGCAUCGACUACCUGCACCCCGACUUCAUCAA GACGACGGCACCUCCAAGGUGCUGUACCUGUGGGACCAG GAGGCCAACCAACCCCCCCCCGAGGGCUUCAUCUUC GGCUCGAGUUCACCCGCUCGAGCUGAACAUCCGCAUC AACCGCAACGACGGCUCCUGUCCAGGACAACAUCGGC ACCGCACCCUGGUGUCCGGCAUCCUGGUGGGCAACGGC CGCAUCAACUCCAGUACCGCGGCAUACCACCGAGUCC GACCUGAUCGUGGUGAAGCUGAAGUCCUACACCGACACC UACUACGCCGGCCGCAUCAACUACUCCGUGUCCGACUUC CUGGCCGCCAUACCUACGUGACCAACAUCGCCCGCACC GAGAACAAGCCCCUGAUCAUCAACCUGACCAUCGGCGUG AAGUCCUCCGCCGUGGCCACCACCUCCAUCUGGACACC UUCAACAUCUGUCCUCCGCCGGCGUGGUGGUGUCC	SEQ ID NO: 83

	GGCGCCGGCAACCAGGGCAACACCGACAUCCACUACUCC GGCCGCUUCUCCUCCGUGGGCGAGGUGCAGGACGUGAUC AUCCAGGACGGCGACGACUACGCCCUGGACAUCACCCUG AACACCAACGGCCCCGACAAGGUGGGCGCCCAGAUCAUC UCCCCUCCGGCGAGGUGUCCACGACAUCCGCUACUCCC CCGACUUCUACAUCUACCGCGGCAAGUUCAACCUGGAGA ACACCACCUACGCCAUGCGCUUCAUCUACCCCUACAUC CCUCCGGCAAGGAGAACCUGGAGAUCCGCCUGCGCGACA UCAAGCCCGGCGUGUGGAUCCUGCGCCUGACCUCCGAGC UGAUCAUCAACGGCGAGUACGACAUCUACCUGCCCAACA AGAACCUGAUCGCCCCCGACACCCGCUUCCUGGACCCCG ACUCCGUGGGCCACCAUCACCAUGUACGCCGCCUCCGACG ACGUGAUCACCGUGGGCACCUUCAACAACAAGACCGACU CCAUGUGGAUCGGCUCCUCCAAGGGCCCCAUCCGCGGCC GCGGCAUCAAGCCCGACAUCGUGGCCUCCGGCGUGGACA UCAUCUCCACCUACAAGAACGGCACCUACAACACCGGCA CCGGCACCGGCGUGUCCUCCUCCAUCGUGACCGGCGUGC UGGCCCUGCUGAUGGAGUACCUGGAGAAGCAGGACAACG UGCCCCGCCUGUCCUGUACCCAGGUGCUGAAGACCU ACCUGAUCCUGGGCGCCACCAAGCUGGAGAUCAACCU ACCCCAACGUGUCCAGGGCUACGGCAUCCUGAACCUGA AGAACACCAUCCAGCAGAUCCGCCAACACCCUGUAAUAA	
Bile acid germinant receptor pseudoprotea se CspC mRNA	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAUUUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGG</u> AGAAGUCCUACUGCAUCAUCUACCAGGGCGACAUCGAGU CCGCCCUGCAGGAGAACGGCAUCAACCGCUACAUGGUGC UGAACUCCCAGCUGGGCCGUGAUCUACGUGCCCCUGGACU UCGACGAGACCAUCCUGAACAACAUAUCCAGGUGGCCU GGUGGGAGGAGUCCGAGCCCAUGUCCUCCCUGAUCGAGA UCACCAACAACGUGAACAACGGCGAGACCAUACCAACCG CCGCCGAGACCGACUACAUCUACGAGAACCCCUACAACG ACAUCACCGGCCCGCGGCAUCCUGCUGGCCGUGAUCGACU CCGGCAUCGACUACCUGCACCCCGACUUAUCAACGACG ACGGCACCUCCAAGGUGCUGUACCUGUGGGACCAAGG CCAACACCAACCCCCCCCCGAGGGCUUCAUCUUCGGCUC CGAGUUCACCCGCUCCGAGCUGAACAUCGCCAUCAACCG CAACGACGGCUCCUGUCCCAGGACAACAUCGGCACCCGG CACCCUGGUGUCCGGCAUCCUGGUGGGCAACGGCCGCAU CAACUCCCAGUACCGCGGCAUACCAACCGAGUCCGACCU GAUCGUGGUGAAGCUGAAGUCCUACACCGACACCUACUA CGCCGGCCGCAUCAACUACUCCGUGUCCGACUUCUCCGGC CGCAUACACCUACGUGACCAACAUCGCCCCGACCGAGAA CAAGCCCCUGAUAUCAACCUAGACCAUCGGCGUGAAGUC CUCCGCCGUGGGCCACCACCUCCAUCCUGGACACCUUCAAC AUCCUGUCCUCCGCCGGCGUGGUGGUGGUGUCCGGCGCC GGCAACCAGGGCAACACCGACAUCACUACUCCGGCCGC UUCUCCUCCGUGGGCGAGGUGCAGGACGUGAUAUCCAG GACGGCGACGACUACGCCUGGACAUCACCCUGAACACC AACGGCCCCGACAAGGUGGGCGCCAGAUCAUCUCCCCC UCCGGCGAGGUGUCCACGACAUCCGCUACUCCCCCGAC UUCUACAUCUACCGCGGCAAGUUCAACCUGGAGAACACC ACCUACGCCAUGCGCUUCAUCUACCCCUACAUCACCUCC GGCAAGGAGAACCUGGAGAUCCGCCUGCGCGACAUAAG	SEQ ID NO:84

	CCCGGCGUGUGGAUCCUGCGCCUGACCUCCGAGCUGAUC AUCAACGGCGGAGUACGACAUCUACCUGCCCAACAAGAAC CUGAUCGCCCCCGACACCCGCUUCCUGGACCCCGACUCCG UGGCCACCAUCACCAUGUACGCCGCCUCCGACGACGUGA UCACCGUGGGCACCUCUACAACAAGACCGACUCCAUGU GGAUCGGCUCCUCCAAGGGCCCCAUCCGCGGGCCGCGGCA UCAAGCCCCGACAUCGUGGCCUCCGGCGUGGACAUCUUC CCACCUACAAGAACGGCACCUCACAAACACCGGCACCGGCA CCGGCGUGUCCUCCUCCAUCGUGACCGGCGUGCUGGCCC UGCUGAUGGAGUACCUGGAGAAGCAGGACAACGUGCCCC GCCUGUCCCUGUUACCCAGGUGCUGAAGACCUACCUGA UCCUGGGCGCCACCAAGCUGGAGAUCUACACCUACCCCA ACGUGUCCCAAGGCUACGGCAUCCUGAACCUGAAGAACA CCAUCCAGCAGAUCCCAACACCCUGUAAUAAACUAGUA <u>GUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCA</u> <u>AGAACACCCGAAUGGAGUCUCUAAAGCUACAUAUACCAA</u> <u>CUUACACUUAACAAAUGUUGUCCCCCAAAUAGUAGCCAU</u> <u>UCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUU</u> <u>CUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
Modified Cspc mRNA	AUGGAGAAGUCCUACUGCAUCAUCUACCAGGGCGACAUC GAGUCCGCCCUGCAGGAGAACGGCAUCAACCGGUACAUG GUGCUGAACUCCAGCUGGCCGUGAUCUACGUGCCCCUG GACUUCGACGAGACCAUCCUGAACACAUCUACCAGGUG GCCUGGUGGGAGGAGUCCGAGCCCAUGUCCUCCUGAUC GAGAUACCAACAACGUGAACAAACGGCGAGACCAUACCC ACCGCCGCCGAGACCGACUACAUCUACGAGAACCCCUAC AACGACAUCACCGGCCGGGGCAUCCUGCUGGCCGUGAUC GACUCCGGCAUCGACUACCUGCACCCCGACUUCAUCAAC GACGACGGCACCUCUCCAAGGUGCUGUACCUGUGGGACCAG GAGGCCAACACCAACCCCCCCCCGAGGGCUUCAUCUUC GGCUCCGAGUUCACCCGGUCCGAGCUGAACAUCCGCAUC AACCGGAACGACGGCUCCUGUCCCAAGGACAACAUCGGC ACCGGCACCCUGGUGUCCGGCAUCCUGGUGGGCAACGGC CGGAUCAACUCCAGUACCGGGGCAUCCACCGAGUCC GACCUGAUCGUGGUAAGCUGAAGUCCUACACCGACACC UACUACGCCGGCCGGAUCGGCUACUCCGUGUCCGACUUC CUGGCCGCCAUACCUACGUGACCAACAUCGCCCCGACC GAGAACAAGCCCCUGAUAUCGGCCUGACCAUCGGCGUG AAGUCCUCCGCCGUGGCCACCACCUCCAUCUGGACACC UUCAACAUCUGUCCUCCGCCGGCGUGGUGGUGGUGUCC GGCGCCGGCAACCAGGGCAACACCGACAUCACUACUCC GGCCGGUUCUCCUCCUGGGCGAGGUGCAGGACGUGAUC AUCCAGGACGGCGACGACUACGCCUCCGACAUACCCUG AACACCAACGGCCCCGACAAGGUGGGCGCCAGAUCAUC UCCCCUCCGGCGAGGUGUCCACGACAUCCGGUACUCC CCCGACUUCUACAUCUACCGGGGCAAGUUAACCUAGGAG AACACCACCUACGCCAUGCGGUUCAUCUACCCUACAUC ACCUCCGGCAAGGAGAACCUGGAGAUCGGCGUGCGGGAC AUCAAGCCCCGGCGUGUGGAUCCUGCGGCUGACCUCCGAG CUGAUAUCAACGGCGAGUACGACAUCUACCUGCCCAAC AAGAACCUGAUCGCCCCCGACACCCGGUCCUGGACCCC GACUCCGUGGCCACCAUCACCAUGUACGCCGCCUCCGAC GACGUGAUCACCGUGGGCACCUUCAACAACAAGACCGAC	SEQ ID NO: 85

	UCCAUGUGGAUCGGCUCCUCCAAGGGCCCCAUCCGGGGC CGGGGCAUCAAGCCCGACAUCGUGGCCUCCGGCGUGGAC AUCAUCUCCACCUACAAGAACGGCACCUACAACACCGGC ACCGGCACCGGCGUGUCCUCCUCCAUCGUGACCGGCGUG CUGGGCCUGCUGAUGGAGUACCUGGAGAAGCAGGACAAC GUGCCCCGGCUGUCCUGUUCACCCAGGUGCUGAAGACC UACCUGAUCCUGGGCGCCACCAAGCUGGAGAUCUACACC UACCCCAACGUGUCCCAAGGGCUACGGCAUCCUGAACCUG AAGAACACCAUCCAGCAGAUCCGCAACACCCUGUAAUAA	
Modified Cspc mRNA	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAAUUGCAGCAAU</u> <u>UUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACC</u> AUGG AGAAGUCCUACUGCAUCAUCUACCAGGGCGACAUCGAGU CCGCCCUGCAGGAGAACGGCAUCAACCGGUACAUGGUGC UGAACUCCCAGCUGGGCCGUGAUCUACGUGCCCCUGGACU UCGACGAGACCAUCCUGAACAACAUAUCCAGGUGGCCU GGUGGGAGGAGUCCGAGCCCAUGUCCUCCUGAUCGAGA UCACCAACAACGUGAACAACGGCGAGACCAUACCACCG CCGCCGAGACCGACUACAUCUACGAGAACCCCUACAACG ACAUACCCGGCCGGGGCAUCCUGCUGGCCGUGAUCGACU CCGGCAUCGACUACCUGCACCCCGACUUAUCAACGACG ACGGCACCUCCAAGGUGCUGUACCUGUGGGACCAGGAGG CCAACACCAACCCCCCCCCCGAGGGGCUUCAUCUUCGGCUC CGAGUUCACCCGGUCCGAGCUGAACAUCGCCAUCAACCG GAACGACGGCUCCUGUCCCAGGACAACAUCGGCACCCGG CACCCUGGUGUCCGGCAUCCUGGUGGGCAACGGCCGGAU CAACUCCCAGUACCGGGGCAUACACCCGAGUCCGACCU GAUCGUGGUGAAGCUGAAGUCCUACACCGACACCUACUA CGCCGGCCGGAUCGGCUACUCCGUGUCCGACUUCUGGC CGCCAUCACCUACGUGACCAACAUCGCCCCGACCCGAGAA CAAGCCCCUGAUAUCGGCCUGACCAUCGGCGUGAAGUC CUCCGCCGUGGGCCACCACCUCCAUCCUGGACACCUUCAAC AUCCUGUCCUCCGCCGGCGUGGUGGUGGUGUCCGGCGCC GGCAACCAGGGCAACACCGACAUCCACUACUCCGGCCGG UUCUCCUCCGUGGGCGAGGUGCAGGACGUGAUCAUCCAG GACGGCGACGACUACGCCCUGGACAUCACCCUGAACACC AACGGCCCCGACAAGGUGGGCGCCCAGAUCAUCUCCCCC UCCGGCGAGGUGUCCACGACAUCCGGUACUCCCCCGAC UUCUACAUCUACCGGGGCAAGUUAACCUUGGAGAACACC ACCUACGCCAUGCGGUUCAUCUACCCCUACAUCACCUCC GGCAAGGAGAACCUGGAGAUCCGGCUGCGGGACAUAAG CCCGGCGUGUGGAUCCUGCGGCUGACCUCCGAGCUGAUC AUCAACGGCGAGUACGACAUCUACCUGCCCAACAAGAAC CUGAUCGCCCCCGACACCCGGUUCUUGGACCCCGACUCCG UGGCCACCAUACCAUGUACGCCGCCUCCGACGACGUGA UCACCGUGGGCACCUUCAACAACAAGACCGACUCCAUGU GGAUCGGCUCCUCCAAGGGCCCCAUCCGGGGCCGGGGCA UCAAGCCCCGACAUCGUGGCCUCCGGCGUGGACAUAUCU CCACCUACAAGAACGGCACCUACAACACCGGCACCCGCA CCGGCGUGUCCUCCUCCAUCGUGACCGGCGUGCUGGCC UGCUGAUGGAGUACCUGGAGAAGCAGGACAACGUGCCCC GGCUGUCCUGUUCACCCAGGUGCUGAAGACCUACCUGA UCCUGGGCGCCACCAAGCUGGAGAUCAACCUACCCCA ACUGUCCCAAGGGCUACGGCAUCCUGAACCUGAAGAACA	SEQ ID NO: 86

	<u>CCAUCCAGCAGAU</u> <u>CGCCAACACCCUGUAAUAAACUAGUA</u> <u>GUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCA</u> <u>AGAACACCCGAAUGGAGUCUCUAAGCUACAUAUACCAA</u> <u>CUUACACUUAACAAAUGUUGUCCCCCAAAUAGCCAU</u> <u>UCGUAUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUU</u> <u>CUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u>	
Modified CspC with IL-2 signal MRNA	AUGCGGAUGCAGCUGCUGCUGCUGAUCGCCCUGUCCC UGGCCCUGGUGACCAACUCCGAGAAGUCCUACUGCAUC AUCUACCAGGGCGACAUCGAGUCCGCCUGCAGGAGAAC GGCAUCAACCGGUACAUGGUGCUGAACUCCAGCUGGCC GUGAUCUACGUGCCCCUGGACUUCGACGAGACCAUCCUG AACAACAUAUCCAGGUGGCCUGGUGGGAGGAGUCCGAG CCCAUGUCCUCCUGAUCGAGAUCACCAACAACGUGAAC AACGGCGAGACCAUACACCGCCGCCGAGACCGACUAC AUCUACGAGAACCCCUACAACGACAUCACCGGCCGGGGC AUCCUGCUGGCCGUGAUCGACUCCGGCAUCGACUACCUG CACCCCGACUUCAUCAACGACGACGGCACCUCOAAGGUG CUGUACCUGUGGGACCAGGAGGCCAACACCAACCCCCC CCCGAGGGCUUCAUCUUCGGCUCCGAGUUCACCCGGUCC GAGCUGAACAUCCGAUCAACCGGAACGACGGCUCUCCUG UCCAGGACAACAUCGGCACCGGCACCCUGGUGUCCGGC AUCCUGGUGGGCAACGGCCGGAUCAACUCCAGUACCGG GGCAUACCAACCGAGUCCGACCUGAUCGUGGUGAAGCUG AAGUCCUACACCGACACCUACUACGCCGGCCGGAUCGGC UACUCCGUGUCCGACUUCUCCGGCCGCCAUCACCUACGUG ACCAACAUCGCCCCGACCGAGAACAAGCCCCUGAUAUC GGCCUGACCAUCGGCGUGAAGUCCUCCGCCUGGGCCACC ACCUCCAUCUCCUGGACACCUUCAACAUCUCCUGUCCUCCGCC GGCGUGGUGGUGGUGUCCGGCGCCGGCAACAGGGCAAC ACCGACAUCACUACUCCGGCCGGUUCUCCUCCGUGGGC GAGGUGCAGGACGUGAUAUCCAGGACGGCGACGACUAC GCCCUGGACAUCACCCUGAACACCAACGGCCCCGACAAG GUGGGCGCCCAGAUCAUCUCCCCUCCGGCGAGGUGUCC CACGACAUCCGGUACUCCCCCGACUUCUACAUCUACCGG GGCAAGUUAACCUUGGAGAACACCAACUACGCCAUCGCG UUCAUCUACCCUACAUAACCUCCGGCAAGGAGAACCUG GAGAUCCGGCUGCGGGACAUAAGCCCGGCGUGUGGAUC CUGCGGCUGACCUCCGAGCUGAUAUAACGGCGAGUAC GACAUCUACCUGCCCAACAAGAACCUGAUCGCCCCCGAC ACCCGGUUCUCCUGGACCCCGACUCCGUGGCCACCAUACC AUGUACGCCGCCUCCGACGACGUGAUAACCGUGGGCACC UUCAACAACAAGACCGACUCCAUGUGGAUCGGCUCCUCC AAGGGCCCCAUCCGGGGCCGGGGCAUAAGCCCGACAUC GUGGCCUCCGGCGUGGACAUAUCUCCACCUACAAGAAC GGCACCUAACAACCGGCACCGGCACCGGCGUGUCCUCC UCCAUCGUGACCGGCGUGCUGGCCUCCUGAUGGAGUAC CUGGAGAAGCAGGACAACGUGCCCCGGCUGUCCUGUUC ACCCAGGUGCUGAAGACCUACCUGAUCCUGGGCGCCACC AAGCUGGAGAUCAACCUACCCCAACGUGUCCAGGGC UACGGCAUCCUGAACCUGAAGAACACCAUCCAGCAGAU GCCAACACCCUGUAAUA	SEQ ID NO:87
Modified CspC with	<u>AGCAUAAAAGUCUCAACACAACAUAUACAAAACAAACGA</u> <u>AUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAU</u>	SEQ ID NO: 88

IL-2 signal MRNA	<u>UUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAA</u> <u>AAUUUUCACCAUUUACGAACGAUAGCGCUGCCACCAUGC</u> <u>GGAUGCAGCUGCUGCUGCUGAUCGCCUGUCCUGGC</u> <u>CCUGGUGACCAACUCCGAGAAGUCCUACUGCAUCAUCU</u> <u>ACCAGGGCGACAUCGAGUCCGCCUGCAGGAGAACGGCA</u> <u>UCAACCGGUACAUGGUGCUGAACUCCAGCUGGCCGUGA</u> <u>UCUACGUGCCCCUGGACUUCGACGAGACCAUCCUGAACA</u> <u>ACAUCAUCCAGGUGGCCUGGUGGGAGGAGUCCGAGCCCA</u> <u>UGUCCUCCCUGAUCGAGAUACCAACAACGUGAACAACG</u> <u>GCGAGACCAUACCAACCGCCGCCGAGACCGACUACAUCU</u> <u>ACGAGAACCCCUACAACGACAUCACCGGCCGGGGCAUCC</u> <u>UGCUGGCCGUGAUCGACUCCGGCAUCGACUACCGCACC</u> <u>CCGACUUCAUCAACGACGACGGCACCUCCAAGGUGCUGU</u> <u>ACCUUGUGGGACCAGGAGGCCAACACCAACCCCCCCCCG</u> <u>AGGGCUUCAUCUUCGGCUCCGAGUUCACCCGGUCCGAGC</u> <u>UGAACAUCCGCAUCAACCGGAACGACGGCUCCUGUCCC</u> <u>AGGACAACAUCGGCACCGGCACCCUGGUGUCCGGCAUCC</u> <u>UGGUGGGCAACGGCCGGAUCAACUCCAGUACCGGGGCA</u> <u>UCACCACCGAGUCCGACCUGAUCGUGGUGAAGCUGAAGU</u> <u>CCUACACCGACACCUACUACGCCGGCCGGAUCGGCUACU</u> <u>CCGUGUCCGACUUCUGGCCGCCAUCACCUACGUGACCA</u> <u>ACAUCGCCCGGACCGAGAACAAGCCCCUGAUCAUCCGCC</u> <u>UGACCAUCGGCGUGAAGUCCUCCGCCGUGGCCACCACCU</u> <u>CCAUCCUGGACACCUUCAACAUCUGUCCUCCGCCGGCG</u> <u>UGGUGGUGGUGUCCGGCGCCGGCAACCAGGGCAACACCG</u> <u>ACAUCCACUACUCCGGCCGGUUCUCCUCCUGGGCGAGG</u> <u>UGCAGGACGUGAUCAUCCAGGACGGCGACGACUACGCC</u> <u>UGGACAUCACCCUGAACACCAACGGCCCCGACAAGGUGG</u> <u>GCGCCCAGAUCAUCUCCCCUCCGGCGAGGUGUCCACG</u> <u>ACAUCCGGUACUCCCCGACUUCUACAUCUACCGGGGCA</u> <u>AGUUCAACCUGGAGAACACCACCUACGCCAUGCUGGUCA</u> <u>UCUACCCCUACAUCACCUCCGGCAAGGAGAACCUGGAGA</u> <u>UCCGGCUGCGGGACAUCAAGCCCGGCGUGUGGAUCCUGC</u> <u>GGCUGACCUCCGAGCUGAUCAUCAACGGCGAGUACGACA</u> <u>UCUACCUGCCCAACAAGAACCUGAUCGCCCCGACACCC</u> <u>GGUUCUGGACCCCGACUCCGUGGCCACCAUACCAUGU</u> <u>ACGCCGCCUCCGACGACGUGAUCACCGUGGGCACCUUCA</u> <u>ACAACAAGACCGACUCCAUGUGGAUCGGCUCCUCCAAGG</u> <u>GCCCCAUCCGGGGCCGGGCAUCAAGCCCGACAUCGUGG</u> <u>CCUCCGGCGUGGACAUCAUUCCACCUACAAGAACGGCA</u> <u>CCUACAACACCGGCACCGGCACCGGCGUGUCCUCCUCA</u> <u>UCGUGACCGGCGUGCUGGCCUUGCUGAUGGAGUACCUGG</u> <u>AGAAGCAGGACAACGUGCCCCGGCUGUCCUGUUCACCC</u> <u>AGGUGCUGAAGACCUACCUGAUCCUGGGCGCCACCAAGC</u> <u>UGGAGAUCAACACCUACCCCAACGUGUCCAGGGCUACG</u> <u>GCAUCCUGAACCUGAAGAACACCAUCCAGCAGAUCCCA</u> <u>ACACCCUGUAAUAAACUAGUAGUGACUGACUAGGAUCUG</u> <u>GUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGU</u> <u>CUCUAAGCUACAUAUACCAACUACACUACAAAAUGU</u> <u>UGUCCCCCAAAUGUAGCCAUUCGUUUCUCCUAAUA</u> <u>AAAAAGAAAGUUUCUACAUUCUAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</u> <u>AAAAAAAA</u>	
---------------------	--	--

In some embodiments, an RNA encoding a TcdA antigen comprises a nucleic acid sequence of SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof. In some
5 embodiments, an RNA encoding a TcdA antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, wherein one or more “t” nucleotides is
10 substituted with a pseudouridine or 1-methy-pseudouridine. In some embodiments, a nucleic acid molecule encoding a TcdA antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID
15 NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a nucleoside modified RNA corresponding to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44. In some embodiments, an RNA molecule encoding a TcdA antigen comprises at least a fragment of SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ
20 ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a nucleoside modified RNA corresponding to a fragment of SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44.

In some embodiments, an RNA encoding a TcdB antigen comprises a
30 nucleic acid sequence of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID

NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof. In some embodiments, an RNA encoding a TcdB antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof, wherein one or more “t” nucleotides is substituted with a pseudouridine or 1-methy-pseudouridine. In some embodiments, a nucleic acid molecule encoding a TcdB antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a nucleoside modified RNA corresponding to SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55. In some embodiments, an RNA molecule encoding a TcdB antigen comprises at least a fragment of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a nucleoside modified RNA corresponding to a fragment of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55.

In some embodiments, an RNA encoding a PPEP-1 antigen comprises a nucleic acid sequence of SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a fragment or variant thereof. In some embodiments, an RNA encoding a PPEP-1 antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ

ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a fragment or variant thereof, wherein one or more “t” nucleotides is substituted with a pseudouridine or 1-methy-pseudouridine. In some embodiments, a nucleic acid molecule encoding a PPEP-1 antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a nucleoside modified RNA corresponding to SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66. In some embodiments, an RNA molecule encoding a PPEP-1 antigen comprises at least a fragment of SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a nucleoside modified RNA corresponding to a fragment of SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66.

In some embodiments, an RNA encoding a CdeM antigen comprises a nucleic acid sequence of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74, or a fragment or variant thereof. In some embodiments, an RNA encoding a CdeM antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74, or a fragment or variant thereof, wherein one or more “t” nucleotides is substituted with a

pseudouridine or 1-methy-pseudouridine. In some embodiments, a nucleic acid molecule encoding a CdeM antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74, or a nucleoside modified RNA corresponding to SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74. In some embodiments, an RNA molecule encoding a CdeM antigen comprises at least a fragment of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74, or a nucleoside modified RNA corresponding to a fragment of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74.

In some embodiments, an RNA encoding a CWP84 antigen comprises a nucleic acid sequence of SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78, or a fragment or variant thereof. In some embodiments, an RNA encoding a CWP84 antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78, or a fragment or variant thereof, wherein one or more "t" nucleotides is substituted with a pseudouridine or 1-methy-pseudouridine. In some embodiments, a nucleic acid molecule encoding a CWP84 antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78, or a nucleoside modified RNA corresponding to SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78. In some embodiments, an RNA molecule encoding a CWP84 antigen comprises at least a fragment of SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%,

90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78, or a nucleoside modified RNA corresponding to a fragment of SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78.

5 In some embodiments, an RNA encoding a CWP66 antigen comprises a nucleic acid sequence of SEQ ID NO:79 or SEQ ID NO:80, or a fragment or variant thereof. In some embodiments, an RNA encoding a CWP66 antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ ID NO:79 or SEQ ID NO:80, or a fragment or variant thereof,
10 wherein one or more “t” nucleotides is substituted with a pseudouridine or 1-methyl-pseudouridine. In some embodiments, a nucleic acid molecule encoding a CWP66 antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:79 or SEQ ID NO:80, or a nucleoside modified RNA
15 corresponding to SEQ ID NO:79 or SEQ ID NO:80. In some embodiments, an RNA molecule encoding a CWP66 antigen comprises at least a fragment of SEQ ID NO:79 or SEQ ID NO:80 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:79 or SEQ ID NO:80, or a nucleoside modified RNA
20 corresponding to a fragment of SEQ ID NO:79 or SEQ ID NO:80.

 In some embodiments, an RNA encoding a ZupT antigen comprises a nucleic acid sequence of SEQ ID NO:81 or SEQ ID NO:82, or a fragment or variant thereof. In some embodiments, an RNA encoding a ZupT antigen comprises a nucleoside modified RNA molecule comprising a sequence corresponding to a nucleic acid sequence
25 of SEQ ID NO:81 or SEQ ID NO:82, or a fragment or variant thereof, wherein one or more “t” nucleotides is substituted with a pseudouridine or 1-methyl-pseudouridine. In some embodiments, a nucleic acid molecule encoding a ZupT antigen comprises a nucleic acid sequence having at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID
30 NO:81 or SEQ ID NO:82, or a nucleoside modified RNA corresponding SEQ ID NO:81 or SEQ ID NO:82. In some embodiments, an RNA molecule encoding a ZupT antigen

comprises at least a fragment of SEQ ID NO:81 or SEQ ID NO:82 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:81 or SEQ ID NO:82, or a nucleoside modified RNA corresponding to a fragment of SEQ ID NO:81 or

5 SEQ ID NO:82.

In some embodiments, an RNA encoding a CspC antigen comprises a nucleic acid sequence of SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88, or a fragment or variant thereof. In some embodiments, an RNA encoding a CspC antigen comprises a nucleoside modified RNA

10 molecule comprising a sequence corresponding to a nucleic acid sequence of SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88, or a fragment or variant thereof, wherein one or more “t” nucleotides is substituted with a pseudouridine or 1-methy-pseudouridine. In some embodiments, a nucleic acid molecule encoding a CspC antigen comprises a nucleic acid sequence having

15 at least 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88, or a nucleoside modified RNA corresponding to SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88. In some embodiments, an RNA molecule encoding a

20 CspC antigen comprises at least a fragment of SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88 comprising at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the full length of SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88, or a nucleoside

25 modified RNA corresponding to a fragment of SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88.

An RNA described herein can be made using any methodology in the art, including, but not limited to, in vitro transcription, chemical synthesis, or the like. In some embodiments, an RNA molecule is an RNA molecule made by in vitro transcription

30 in the presence of a mixture of nucleotides comprising pseudouridine triphosphate in place of uridine triphosphate. In some embodiments, a pseudouridine is m¹acp³Ψ (1-

methyl-3-(3-amino-3-carboxypropyl) pseudouridine). In some embodiments, a pseudouridine is $m^1\Psi$ (1-methylpseudouridine). In some embodiments, a pseudouridine is Ψm (2'-O-methylpseudouridine). In some embodiments, a pseudouridine is m^5D (5-methyldihydrouridine). In some embodiments, a pseudouridine is $m^3\Psi$ (3-methylpseudouridine). In some embodiments, a pseudouridine is a pseudouridine moiety that is not further modified.

An RNA sequence encoding at least one *C. difficile* antigen as described herein can alternatively comprise sequence variations with respect to an original nucleotide sequence, for example, substitutions, insertions and/or deletions of one or more nucleotides, with a condition that a resulting polynucleotide encodes a polypeptide according to the present disclosure. Therefore, the scope of the present disclosure includes nucleotide sequences that are substantially homologous or substantially identical to the nucleotide sequences recited herein and encode *C. difficile* antigens of the present disclosure.

A nucleotide sequence that is substantially homologous to an RNA molecule encoding an antigen can typically be isolated from a producer organism of the antigen based on information contained in a nucleotide sequence by means of introducing conservative or non-conservative substitutions, for example. In some embodiments, modifications include insertion of one or more nucleotides in a sequence, addition of one or more nucleotides in any end of a sequence, or deletion of one or more nucleotides in any end or inside a sequence. A degree of identity between two polyribonucleotides is determined using computer algorithms and methods that are widely known for the persons skilled in the art.

In some embodiments, an RNA molecule is operatively bound to a translational control element. In some embodiments, an RNA molecule further comprises an operatively bound regulatory sequence for expression of an RNA molecule, thus forming an expression cassette.

Vectors

A nucleic acid sequence(s) coding for *C. difficile* antigen(s) of the present disclosure can be obtained using recombinant methods known in the art, such as, for

example by screening libraries from cells expressing a gene, by deriving a gene from a vector known to include the same, or by isolating directly from cells and tissues containing the same, using standard techniques. Alternatively, a gene of interest can be produced synthetically.

5 In some embodiments, nucleic acid molecules can be cloned into a number of types of vectors. For example, a nucleic acid can be cloned into a vector including, but not limited to a plasmid, a phagemid, a phage derivative, an animal virus, a PCR-generated linear DNA sequence, and a cosmid. In some embodiments, a vector is an expression vector, replication vector, probe generation vector, sequencing vector and a
10 vector optimized for in vitro transcription.

 In some embodiments, nucleic acid molecules can be introduced into a host cell via methods known in the art including, colloidal dispersion systems, such as macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, carbohydrates, peptides,
15 cationic polymers, and liposomes. In some embodiments, a colloidal system for use as a delivery vehicle in vitro and in vivo is a liposome (e.g., an artificial membrane vesicle).

 In some embodiments, a non-viral delivery system is utilized. In some embodiments, a non-viral delivery system is a liposome. In some embodiments, a lipid formulation is used for introduction of nucleic acids (e.g., RNA) into a host cell (in vitro,
20 ex vivo or in vivo). In some embodiments, a nucleic acid (e.g., RNA) may be associated with a lipid. In some embodiments, a nucleic acid (e.g., RNA) associated with a lipid may be encapsulated in an aqueous interior of a liposome, interspersed within a lipid bilayer of a liposome, attached to a liposome via a linking molecule that is associated with both a liposome and an oligonucleotide, entrapped in a liposome, complexed with a liposome,
25 dispersed in a solution containing a lipid, mixed with a lipid, combined with a lipid, contained as a suspension in a lipid, contained or complexed with a micelle, or otherwise associated with a lipid. In some embodiments, lipid, lipid/RNA or lipid/expression vector associated compositions are not limited to any particular structure in solution. In some
30 embodiments, they may be present in a bilayer structure, as micelles, or with a “collapsed” structure, they may also simply be interspersed in a solution, possibly forming aggregates that are not uniform in size or shape. In some embodiments, a lipids

is a fatty substance which may be naturally occurring or a synthetic lipid. In some embodiments, lipids include fatty droplets that naturally occur in a cytoplasm as well as a class of compounds which contain long-chain aliphatic hydrocarbons and their derivatives, such as fatty acids, alcohols, amines, amino alcohols, and aldehydes.

5 In some embodiments, lipids suitable for use in accordance with the present disclosure can be obtained from commercial sources. For example, dimyristyl phosphatidylcholine (“DMPC”) can be obtained from Sigma, St. Louis, MO; dicetyl phosphate (“DCP”) can be obtained from K & K Laboratories (Plainview, NY); cholesterol (“Choi”) can be obtained from Calbiochem-Behring; dimyristyl
10 phosphatidylglycerol (“DMPG”) and other lipids may be obtained from Avanti Polar Lipids, Inc. (Birmingham, AL). In some embodiments, stock solutions of lipids in chloroform or chloroform/methanol can be stored at about -20°C. In some embodiments, chloroform is used as it is more readily evaporated than methanol.

 A variety of assays to confirm presence of an RNA sequence in a host cell
15 may be performed. In some embodiments, such assays include, for example, “molecular biological” assays well known to those of skill in the art, such as Northern blotting and RT-PCR; “biochemical” assays, such as detecting presence or absence of a particular peptide, e.g., by immunogenic means (ELISAs and Western blots) or by assays described herein to identify agents falling within the scope of the present disclosure.

20

In vitro transcribed RNA

 In some embodiments, a composition of the present disclosure comprises a combination of in vitro transcribed (IVT) RNA molecules encoding *C. difficile* antigens of the present disclosure. In some embodiments, an IVT RNA can be introduced to a cell
25 as a form of transient transfection. RNA is produced by in vitro transcription using a plasmid DNA template generated synthetically. DNA of interest from any source can be directly converted by PCR into a template for in vitro RNA synthesis using appropriate primers and RNA polymerase. A source of DNA can be, for example, genomic DNA, plasmid DNA, phage DNA, cDNA, synthetic DNA sequence or any other appropriate
30 source of DNA. In some embodiments, a desired template for in vitro transcription is a *C. difficile* antigen capable of inducing an adaptive immune response. In some

embodiments, a desired template for in vitro transcription is an adjuvant capable of enhancing an adaptive immune response.

In some embodiments, DNA to be used for PCR contains an open reading frame. DNA can be from a naturally occurring DNA sequence from the genome of an organism. In some embodiments, DNA is a full-length gene or a portion of a gene. A gene can include some or all of the 5' and/or 3' untranslated regions (UTRs). A gene can include exons and introns. In some embodiments, DNA to be used for PCR is a human gene. In another embodiment, DNA to be used for PCR is a human gene, including 5' and 3' UTRs. In another embodiment, DNA to be used for PCR is a gene from a pathogenic or commensal organism, including bacteria, viruses, parasites, and fungi. In another embodiment, DNA to be used for PCR is from a pathogenic or commensal organism, including bacteria, viruses, parasites, and fungi, including 5' and 3' UTRs. DNA can alternatively be an artificial DNA sequence that is not normally expressed in a naturally occurring organism. An exemplary artificial DNA sequence is one that contains portions of genes that are ligated together to form an open reading frame that encodes a fusion protein. Portions of DNA that are ligated together can be from a single organism or from more than one organism.

Genes that can be used as sources of DNA for PCR include genes that encode polypeptides that induce or enhance an adaptive immune response in an organism. In some instances, genes are useful for a short-term treatment. In some instances, genes have limited safety concerns regarding dosage of an expressed gene.

In some embodiments, a plasmid is used to generate a template for in vitro transcription of RNA, which is used for transfection.

Chemical structures with the ability to promote stability and/or translation efficiency may also be used. In some embodiments, RNA has 5' and 3' UTRs. In some embodiments, a 5' UTR is between zero and 3000 nucleotides in length. A length of 5' and 3' UTR sequences to be added to the coding region can be altered by different methods, including, but not limited to, designing primers for PCR that anneal to different regions of UTRs. Using this approach, one of ordinary skill in the art can modify 5' and 3' UTR lengths required to achieve optimal translation efficiency following transfection of a transcribed RNA.

5' and 3' UTRs can be the naturally occurring, endogenous 5' and 3' UTRs for the gene of interest. Alternatively, UTR sequences that are not endogenous to a gene of interest can be added by incorporating UTR sequences into the forward and reverse primers or by any other modifications of the template. The use of UTR sequences that are not endogenous to a gene of interest can be useful for modifying the stability and/or translation efficiency of an RNA. For example, AU-rich elements in 3' UTR sequences can decrease the stability of RNA. Therefore, 3' UTRs can be selected or designed to increase the stability of the transcribed RNA based on properties of UTRs that are well known in the art.

In some embodiments, a 5' UTR can contain a Kozak sequence of an endogenous gene. Alternatively, when a 5' UTR that is not endogenous to a gene of interest is being added by PCR as described above, a consensus Kozak sequence can be redesigned by adding a 5' UTR sequence. Kozak sequences can increase the efficiency of translation of some RNA transcripts but does not appear to be required for all RNAs to enable efficient translation. In some embodiments, a 5' UTR can be derived from an RNA virus whose RNA genome is stable in cells. In some embodiments, various nucleotide analogues can be used in a 3' or 5' UTR to impede exonuclease degradation of the RNA.

To enable synthesis of RNA from a DNA template, a promoter of transcription should be attached to the DNA template upstream of the sequence to be transcribed. When a sequence that functions as a promoter for an RNA polymerase is added to the 5' end of the forward primer, a RNA polymerase promoter becomes incorporated into the PCR product upstream of the open reading frame that is to be transcribed. In some embodiments, a promoter is a T7 RNA polymerase promoter, as described elsewhere herein. Other useful promoters include, but are not limited to, T3 and SP6 RNA polymerase promoters. Consensus nucleotide sequences for T7, T3 and SP6 promoters are known in the art.

In some embodiments, RNA has both a cap on the 5' end and a 3' poly(A) tail, which can determine ribosome binding, initiation of translation and stability of RNA in the cell. On a circular DNA template, for instance, plasmid DNA, RNA polymerase produces a long concatameric product, which is not suitable for expression in eukaryotic

cells. Transcription of plasmid DNA linearized at the end of a 3' UTR results in normal sized RNA, which is effective in eukaryotic transfection when it is polyadenylated after transcription.

On a linear DNA template, phage T7 RNA polymerase can extend the 3' end of the transcript beyond the last base of the template (Schenborn and Mierendorf, Nuc Acids Res., 13:6223-36 (1985); Nacheva and Berzal-Herranz, Eur. J. Biochem., 270:1485-65 (2003)).

One conventional method of integration of polyA/T stretches into a DNA template is molecular cloning. However, polyA/T sequence integrated into plasmid DNA can cause plasmid instability, which can be ameliorated through the use of recombination incompetent bacterial cells for plasmid propagation.

Poly(A) tails of RNAs can be further extended following in vitro transcription with the use of a poly(A) polymerase, such as E. coli polyA polymerase (E-PAP) or yeast polyA polymerase. In some embodiments, increasing the length of a poly(A) tail from 100 nucleotides to between 300 and 400 nucleotides results in about a two-fold increase in the translation efficiency of the RNA. Additionally, attachment of different chemical groups to the 3' end can increase RNA stability. Such attachment can contain modified/artificial nucleotides, aptamers and other compounds. For example, ATP analogs can be incorporated into the poly(A) tail using poly(A) polymerase. ATP analogs can further increase the stability of the RNA.

5' caps also provide stability to RNA molecules. In some embodiments, RNAs produced by the methods to include a 5' cap1 structure. Such cap1 structure can be generated using Vaccinia capping enzyme and 2'-O-methyltransferase enzymes (CellScript, Madison, WI). Alternatively, 5' cap is provided using techniques known in the art and described herein (Cougot, et al., Trends in Biochem. Sci., 29:436-444 (2001); Stepinski, et al., RNA, 7:1468-95 (2001); Elango, et al., Biochim. Biophys. Res. Commun., 330:958-966 (2005)).

RNA can be introduced into target cells using any of a number of different methods, for instance, commercially available methods which include, but are not limited to, electroporation (Amaza Nucleofector-II (Amaza Biosystems, Cologne, Germany)), (ECM 830 (BTX) (Harvard Instruments, Boston, Mass.) or the Gene Pulser II (BioRad,

Denver, Colo.), Multiporator (Eppendorf, Hamburg Germany), cationic liposome mediated transfection using lipofection, polymer encapsulation, peptide mediated transfection, or biolistic particle delivery systems such as “gene guns” (see, for example, Nishikawa, et al. Hum Gene Ther., 12(8):861-70 (2001)). In some embodiments RNA of the present disclosure is introduced to a cell with a method comprising the use of TransIT[®]-mRNA transfection Kit (Mirus, Madison WI), which, in some instances, provides high efficiency, low toxicity, transfection.

Nucleoside-modified RNA

In some embodiments, a composition of the present disclosure comprises a nucleoside-modified nucleic acid encoding a *C. difficile* antigen as described herein. In some embodiments, a composition of the present disclosure comprises a plurality of nucleoside-modified nucleic acid molecules encoding a plurality of *C. difficile* antigens as described herein.

For example, in some embodiments, a composition comprises a nucleoside-modified RNA. In some embodiments, a composition comprises a nucleoside-modified RNA. Nucleoside-modified RNA have particular advantages over non-modified RNA, including for example, increased stability, low or absent innate immunogenicity, and enhanced translation. Nucleoside-modified RNA useful in the present disclosure is further described in U.S. Patent Nos. 8,278,036, 8,691,966, and 8,835,108, each of which is incorporated by reference herein in its entirety.

In some embodiments, nucleoside-modified RNA does not activate any pathophysiologic pathways, translates very efficiently and almost immediately following delivery, and serve as templates for continuous protein production in vivo lasting for several days to weeks (Karikó et al., 2008, Mol Ther 16:1833-1840; Karikó et al., 2012, Mol Ther 20:948-953). The amount of RNA required to exert a physiological effect is small, making it applicable for human therapy. For example, as described herein, nucleoside-modified RNA encoding an antigen has demonstrated the ability to induce antigen-specific antibody production. For example, in some instances, antigen encoded by nucleoside-modified RNA induces greater production of antigen-specific antibody production as compared to antigen encoded by non-modified RNA.

In some instances, expressing a protein by delivering the encoding RNA has many benefits over methods that use protein, plasmid DNA or viral vectors. During RNA transfection, the coding sequence of the desired protein is the only substance delivered to cells, thus avoiding all the side effects associated with plasmid backbones, viral genes, and viral proteins. More importantly, unlike DNA- and viral-based vectors, the RNA does not carry the risk of being incorporated into the genome and protein production starts immediately after RNA delivery. For example, high levels of circulating proteins have been measured within 15 to 30 minutes of in vivo injection of the encoding RNA. In some embodiments, using RNA rather than the protein also has many advantages. Half-lives of proteins in the circulation or in tissues are often short, thus protein treatment would need frequent dosing, while RNA provides a template for continuous protein production for several days to weeks. Purification of proteins is problematic, and they can contain aggregates and other impurities that cause adverse effects (Kromminga and Schellekens, 2005, *Ann NY Acad Sci* 1050:257-265).

In some embodiments, the nucleoside-modified RNA comprises the naturally occurring modified-nucleoside pseudouridine. In some embodiments, inclusion of pseudouridine makes the RNA more stable, non-immunogenic, and highly translatable (Karikó et al., 2008, *Mol Ther* 16:1833-1840; Anderson et al., 2010, *Nucleic Acids Res* 38:5884-5892; Anderson et al., 2011, *Nucleic Acids Research* 39:9329-9338; Karikó et al., 2011, *Nucleic Acids Research* 39:e142; Karikó et al., 2012, *Mol Ther* 20:948-953; Karikó et al., 2005, *Immunity* 23:165-175).

It has been demonstrated that the presence of modified nucleosides, including pseudouridines in RNA suppress their innate immunogenicity (Karikó et al., 2005, *Immunity* 23:165-175). Further, protein-encoding, in vitro-transcribed RNA containing pseudouridine can be translated more efficiently than RNA containing no or other modified nucleosides (Karikó et al., 2008, *Mol Ther* 16:1833-1840). Subsequently, it is shown that the presence of pseudouridine improves the stability of RNA (Anderson et al., 2011, *Nucleic Acids Research* 39:9329-9338) and abates both activation of PKR and inhibition of translation (Anderson et al., 2010, *Nucleic Acids Res* 38:5884-5892).

Similar effects as described for pseudouridine have also been observed for RNA containing 1-methyl-pseudouridine.

In some embodiments, a nucleoside-modified nucleic acid molecule is a purified nucleoside-modified nucleic acid molecule. For example, in some embodiments, a composition is purified to remove double-stranded contaminants. In some instances, a preparative high-performance liquid chromatography (HPLC) purification procedure is used to obtain pseudouridine-containing RNA that has superior translational potential and no innate immunogenicity (Karikó et al., 2011, *Nucleic Acids Research* 39:e142). Administering HPLC-purified, pseudouridine-containing RNA coding for erythropoietin into mice and macaques resulted in a significant increase of serum EPO levels (Karikó et al., 2012, *Mol Ther* 20:948-953), thus confirming that pseudouridine-containing RNA is suitable for in vivo protein therapy. In some embodiments, a nucleoside-modified nucleic acid molecule is purified using non-HPLC methods. In some instances, a nucleoside-modified nucleic acid molecule is purified using chromatography methods, including but not limited to HPLC and fast protein liquid chromatography (FPLC). An exemplary FPLC-based purification procedure is described in Weissman et al., 2013, *Methods Mol Biol*, 969: 43-54. Exemplary purification procedures are also described in U.S. Patent Application Publication No. US2016/0032316, which is hereby incorporated by reference in its entirety.

The present disclosure encompasses RNA, oligoribonucleotide, and polyribonucleotide molecules comprising pseudouridine or a modified nucleoside. In some embodiments, a composition comprises an isolated nucleic acid encoding an antigen, wherein a nucleic acid comprises a pseudouridine or a modified nucleoside. In some embodiments, a composition comprises a vector, comprising an isolated nucleic acid encoding an antigen, adjuvant, or combination thereof, wherein the nucleic acid comprises a pseudouridine or a modified nucleoside.

In some embodiments, the nucleoside-modified RNA of the present disclosure is IVT RNA, as described elsewhere herein. For example, in some embodiments, a nucleoside-modified RNA is synthesized by T7 phage RNA polymerase. In some embodiments, a nucleoside-modified RNA is synthesized by SP6 phage RNA polymerase. In some embodiments, a nucleoside-modified RNA is synthesized by T3 phage RNA polymerase.

In some embodiments, a modified nucleoside is m¹acp³Ψ (1-methyl-3-(3-

amino-3-carboxypropyl) pseudouridine. In some embodiments, a modified nucleoside is $m^1\Psi$ (1-methylpseudouridine). In some embodiments, the modified nucleoside is Ψm (2'-O-methylpseudouridine). In some embodiments, a modified nucleoside is m^5D (5-methyldihydrouridine). In some embodiments, a modified nucleoside is $m^3\Psi$ (3-methylpseudouridine). In some embodiments, a modified nucleoside is a pseudouridine moiety that is not further modified. In some embodiments, a modified nucleoside is a monophosphate, diphosphate, or triphosphate of any of the above pseudouridines. In some embodiments, a modified nucleoside is any other pseudouridine-like nucleoside known in the art.

10 In some embodiments, a nucleoside that is modified in a nucleoside-modified RNA of the present disclosure is uridine (U). In some embodiments, a modified nucleoside is cytidine (C). In some embodiments, a modified nucleoside is adenosine (A). In some embodiments, a modified nucleoside is guanosine (G).

In some embodiments, a modified nucleoside of the present disclosure is m^5C (5-methylcytidine). In some embodiments, a modified nucleoside is m^5U (5-methyluridine). In some embodiments, a modified nucleoside is m^6A (N^6 -methyladenosine). In some embodiments, a modified nucleoside is s^2U (2-thiouridine). In some embodiments, a modified nucleoside is Ψ (pseudouridine). In some embodiments, a modified nucleoside is Um (2'-O-methyluridine).

20 In some embodiments, a modified nucleoside is m^1A (1-methyladenosine); m^2A (2-methyladenosine); Am (2'-O-methyladenosine); ms^2m^6A (2-methylthio- N^6 -methyladenosine); i^6A (N^6 -isopentenyladenosine); ms^2i^6A (2-methylthio- N^6 -isopentenyladenosine); io^6A (N^6 -(cis-hydroxyisopentenyl)adenosine); ms^2io^6A (2-methylthio- N^6 -(cis-hydroxyisopentenyl) adenosine); g^6A (N^6 -glycinylicarbamoyladenosine); t^6A (N^6 -threonylcarbamoyladenosine); ms^2t^6A (2-methylthio- N^6 -threonyl carbamoyladenosine); m^6t^6A (N^6 -methyl- N^6 -threonylcarbamoyladenosine); hn^6A (N^6 -hydroxynorvalylcarbamoyladenosine); ms^2hn^6A (2-methylthio- N^6 -hydroxynorvalyl carbamoyladenosine); $Ar(p)$ (2'-O-ribosyladenosine (phosphate)); I (inosine); m^1I (1-methylinosine); m^1Im (1,2'-O-dimethylinosine); m^3C (3-methylcytidine); Cm (2'-O-methylcytidine); s^2C (2-thiocytidine); ac^4C (N^4 -acetylcytidine); f^5C (5-formylcytidine); m^5Cm (5,2'-O-dimethylcytidine); ac^4Cm (N^4 -

25
30

acetyl-2'-O-methylcytidine); k²C (lysidine); m¹G (1-methylguanosine); m²G (N²-methylguanosine); m⁷G (7-methylguanosine); Gm (2'-O-methylguanosine); m²₂G (N²,N²-dimethylguanosine); m²Gm (N²,2'-O-dimethylguanosine); m²₂Gm (N²,N²,2'-O-trimethylguanosine); Gr(p) (2'-O-ribosylguanosine (phosphate)); yW (wybutosine);

5 ozyW (peroxywybutosine); OHyW (hydroxywybutosine); OHyW* (undermodified hydroxywybutosine); imG (wyosine); mimG (methylwyosine); Q (queuosine); oQ (epoxyqueuosine); galQ (galactosyl-queuosine); manQ (mannosyl-queuosine); preQ₀ (7-cyano-7-deazaguanosine); preQ₁ (7-aminomethyl-7-deazaguanosine); G⁺ (archaeosine); D (dihydrouridine); m⁵Um (5,2'-O-dimethyluridine); s⁴U (4-thiouridine); m⁵s²U (5-

10 methyl-2-thiouridine); s²Um (2-thio-2'-O-methyluridine); acp³U (3-(3-amino-3-carboxypropyl)uridine); ho⁵U (5-hydroxyuridine); mo⁵U (5-methoxyuridine); cmo⁵U (uridine 5-oxyacetic acid); mcmo⁵U (uridine 5-oxyacetic acid methyl ester); chm⁵U (5-(carboxyhydroxymethyl)uridine)); mchm⁵U (5-(carboxyhydroxymethyl)uridine methyl ester); mcm⁵U (5-methoxycarbonylmethyluridine); mcm⁵Um (5-

15 methoxycarbonylmethyl-2'-O-methyluridine); mcm⁵s²U (5-methoxycarbonylmethyl-2-thiouridine); nm⁵s²U (5-aminomethyl-2-thiouridine); mnm⁵U (5-methylaminomethyluridine); mnm⁵s²U (5-methylaminomethyl-2-thiouridine); mnm⁵se²U (5-methylaminomethyl-2-selenouridine); ncm⁵U (5-carbamoylmethyluridine); ncm⁵Um (5-carbamoylmethyl-2'-O-methyluridine); cmnm⁵U (5-

20 carboxymethylaminomethyluridine); cmnm⁵Um (5-carboxymethylaminomethyl-2'-O-methyluridine); cmnm⁵s²U (5-carboxymethylaminomethyl-2-thiouridine); m⁶₂A (N⁶,N⁶-dimethyladenosine); Im (2'-O-methylinosine); m⁴C (N⁴-methylcytidine); m⁴Cm (N⁴,2'-O-dimethylcytidine); hm⁵C (5-hydroxymethylcytidine); m³U (3-methyluridine); cm⁵U (5-carboxymethyluridine); m⁶Am (N⁶,2'-O-dimethyladenosine); m⁶₂Am (N⁶,N⁶,O-2'-

25 trimethyladenosine); m^{2,7}G (N²,7-dimethylguanosine); m^{2,2,7}G (N²,N²,7-trimethylguanosine); m³Um (3,2'-O-dimethyluridine); m⁵D (5-methyldihydrouridine); f⁵Cm (5-formyl-2'-O-methylcytidine); m¹Gm (1,2'-O-dimethylguanosine); m¹Am (1,2'-O-dimethyladenosine); τm⁵U (5-taurinomethyluridine); τm⁵s²U (5-taurinomethyl-2-thiouridine)); imG-14 (4-demethylwyosine); imG2 (isowyosine); or ac⁶A (N⁶-

30 acetyladenosine).

In some embodiments, a nucleoside-modified RNA of the present

disclosure comprises a combination of 2 or more of the above modifications. In some embodiments, a nucleoside-modified RNA comprises a combination of 3 or more of the above modifications. In some embodiments, a nucleoside-modified RNA comprises a combination of more than 3 of the above modifications.

5 In some embodiments, between 0.1% and 100% of residues in a nucleoside-modified RNA of the present disclosure are modified (e.g., either by the presence of pseudouridine, 1-methyl-pseudouridine, 5-methyl-uridine or another modified nucleoside base). In some embodiments, a fraction of modified residues is 0.1%. In some embodiments, a fraction of modified residues is 0.2%. In some
10 embodiments, a fraction is 0.3%. In some embodiments, a fraction is 0.4%. In some embodiments, a fraction is 0.5%. In some embodiments, a fraction is 0.6%. In some embodiments, a fraction is 0.7%. In some embodiments, a fraction is 0.8%. In some embodiments, a fraction is 0.9%. In some embodiments, a fraction is 1%. In some
15 embodiments, a fraction is 1.5%. In some embodiments, a fraction is 2%. In some embodiments, a fraction is 2.5%. In some embodiments, a fraction is 3%. In some embodiments, a fraction is 4%. In some embodiments, a fraction is 5%. In some
20 embodiments, a fraction is 6%. In some embodiments, a fraction is 7%. In some embodiments, a fraction is 8%. In some embodiments, a fraction is 9%. In some embodiments, a fraction is 10%. In some embodiments, a fraction is 12%. In some
25 embodiments, a fraction is 14%. In some embodiments, a fraction is 16%. In some embodiments, a fraction is 18%. In some embodiments, a fraction is 20%. In some embodiments, a fraction is 25%. In some embodiments, a fraction is 30%. In some
30 embodiments, a fraction is 35%. In some embodiments, a fraction is 40%. In some embodiments, a fraction is 45%. In some embodiments, a fraction is 50%. In some
embodiments, a fraction is 55%. In some embodiments, a fraction is 60%. In some
embodiments, a fraction is 65%. In some embodiments, a fraction is 70%. In some
embodiments, a fraction is 75%. In some embodiments, a fraction is 80%. In some
embodiments, a fraction is 85%. In some embodiments, a fraction is 90%. In some
embodiments, a fraction is 91%. In some embodiments, a fraction is 92%. In some
embodiments, a fraction is 93%. In some embodiments, a fraction is 94%. In some
embodiments, a fraction is 95%. In some embodiments, a fraction is 96%. In some

embodiments, a fraction is 97%. In some embodiments, a fraction is 98%. In some embodiments, a fraction is 99%. In some embodiments, a fraction is 100%.

In some embodiments, a fraction is less than 5%. In some embodiments, a fraction is less than 3%. In some embodiments, a fraction is less than 1%. In some
5 embodiments, a fraction is less than 2%. In some embodiments, a fraction is less than 4%. In some embodiments, a fraction is less than 6%. In some embodiments, a fraction is less than 8%. In some embodiments, a fraction is less than 10%. In some embodiments, a fraction is less than 12%. In some embodiments, a fraction is less than 15%. In some
10 embodiments, a fraction is less than 20%. In some embodiments, a fraction is less than 30%. In some embodiments, a fraction is less than 40%. In some embodiments, a fraction is less than 50%. In some embodiments, a fraction is less than 60%. In some
embodiments, a fraction is less than 70%.

In some embodiments, 0.1% of the residues of a given nucleoside (i.e., uridine, cytidine, guanosine, or adenosine) are modified. In some embodiments, a fraction
15 of modified residues is 0.2%. In some embodiments, a fraction is 0.3%. In some embodiments, a fraction is 0.4%. In some embodiments, a fraction is 0.5%. In some embodiments, a fraction is 0.6%. In some embodiments, a fraction is 0.7%. In some embodiments, a fraction is 0.8%. In some embodiments, a fraction is 0.9%. In some
embodiments, a fraction is 1%. In some embodiments, a fraction is 1.5%. In some
20 embodiments, a fraction is 2%. In some embodiments, a fraction is 2.5%. In some embodiments, a fraction is 3%. In some embodiments, a fraction is 4%. In some embodiments, a fraction is 5%. In some embodiments, a fraction is 6%. In some embodiments, a fraction is 7%. In some embodiments, a fraction is 8%. In some
embodiments, a fraction is 9%. In some embodiments, a fraction is 10%. In some
25 embodiments, a fraction is 12%. In some embodiments, a fraction is 14%. In some embodiments, a fraction is 16%. In some embodiments, a fraction is 18%. In some embodiments, a fraction is 20%. In some embodiments, a fraction is 25%. In some
embodiments, a fraction is 30%. In some embodiments, a fraction is 35%. In some
embodiments, a fraction is 40%. In some embodiments, a fraction is 45%. In some
30 embodiments, a fraction is 50%. In some embodiments, a fraction is 55%. In some
embodiments, a fraction is 60%. In some embodiments, a fraction is 65%. In some

embodiments, a fraction is 70%. In some embodiments, a fraction is 75%. In some
embodiments, a fraction is 80%. In some embodiments, a fraction is 85%. In some
embodiments, a fraction is 90%. In some embodiments, a fraction is 91%. In some
embodiments, a fraction is 92%. In some embodiments, a fraction is 93%. In some
5 embodiments, a fraction is 94%. In some embodiments, a fraction is 95%. In some
embodiments, a fraction is 96%. In some embodiments, a fraction is 97%. In some
embodiments, a fraction is 98%. In some embodiments, a fraction is 99%. In some
embodiments, a fraction is 100%. In some embodiments, a fraction of a given nucleotide
that is modified is less than 8%. In some embodiments, a fraction is less than 10%. In
10 some embodiments, a fraction is less than 5%. In some embodiments, a fraction is less
than 3%. In some embodiments, a fraction is less than 1%. In some embodiments, a
fraction is less than 2%. In some embodiments, a fraction is less than 4%. In some
embodiments, a fraction is less than 6%. In some embodiments, a fraction is less than
12%. In some embodiments, a fraction is less than 15%. In some embodiments, a fraction
15 is less than 20%. In some embodiments, a fraction is less than 30%. In some
embodiments, a fraction is less than 40%. In some embodiments, a fraction is less than
50%. In some embodiments, a fraction is less than 60%. In some embodiments, a fraction
is less than 70%.

In some embodiments, a composition comprises a purified preparation of
20 single-stranded nucleoside modified RNA. For example, in some embodiments, a
purified preparation of single-stranded nucleoside modified RNA is substantially free of
double stranded RNA (dsRNA). In some embodiments, a purified preparation is at least
90%, or at least 91%, or at least 92%, or at least 93 % or at least 94%, or at least 95%, or
at least 96%, or at least 97%, or at least 98%, or at least 99%, or at least 99.5%, or at least
25 99.9% single stranded nucleoside modified RNA, relative to all other nucleic acid
molecules (DNA, dsRNA, etc.).

In some embodiments, a nucleoside-modified RNA of the present
disclosure is translated in a cell more efficiently than an unmodified RNA molecule with
a same sequence. In some embodiments, a nucleoside-modified RNA exhibits enhanced
30 ability to be translated by a target cell. In some embodiments, translation is enhanced by
a factor of 2-fold relative to its unmodified counterpart. In some embodiments, translation

is enhanced by a 3-fold factor. In some embodiments, translation is enhanced by a 4-fold factor. In some embodiments, translation is enhanced by a 5-fold factor. In some embodiments, translation is enhanced by a 6-fold factor. In some embodiments, translation is enhanced by a 7-fold factor. In some embodiments, translation is enhanced by an 8-fold factor. In some embodiments, translation is enhanced by a 9-fold factor. In some embodiments, translation is enhanced by a 10-fold factor. In some embodiments, translation is enhanced by a 15-fold factor. In some embodiments, translation is enhanced by a 20-fold factor. In some embodiments, translation is enhanced by a 50-fold factor. In some embodiments, translation is enhanced by a 100-fold factor. In some embodiments, translation is enhanced by a 200-fold factor. In some embodiments, translation is enhanced by a 500-fold factor. In some embodiments, translation is enhanced by a 1000-fold factor. In some embodiments, translation is enhanced by a 2000-fold factor. In some embodiments, a factor is 10-1000-fold. In some embodiments, a factor is 10-100-fold. In some embodiments, a factor is 10-200-fold. In some embodiments, a factor is 10-300-fold. In some embodiments, a factor is 10-500-fold. In some embodiments, a factor is 20-1000-fold. In some embodiments, a factor is 30-1000-fold. In some embodiments, a factor is 50-1000-fold. In some embodiments, a factor is 100-1000-fold. In some embodiments, a factor is 200-1000-fold. In some embodiments, translation is enhanced by any other significant amount or range of amounts.

In some embodiments, a nucleoside-modified antigen-encoding RNA of the present disclosure induces a significantly more robust adaptive immune response as compared with an unmodified in vitro-synthesized RNA molecule of the same sequence. In some embodiments, a modified RNA molecule induces an adaptive immune response that is 2-fold greater than its unmodified counterpart. In some embodiments, an adaptive immune response is increased by a 3-fold factor. In some embodiments, an adaptive immune response is increased by a 4-fold factor. In some embodiments, an adaptive immune response is increased by a 5-fold factor. In some embodiments, an adaptive immune response is increased by a 6-fold factor. In some embodiments, an adaptive immune response is increased by a 7-fold factor. In some embodiments, an adaptive immune response is increased by an 8-fold factor. In some embodiments, an adaptive immune response is increased by a 9-fold factor. In some embodiments, an adaptive

immune response is increased by a 10-fold factor. In some embodiments, an adaptive immune response is increased by a 15-fold factor. In some embodiments, an adaptive immune response is increased by a 20-fold factor. In some embodiments, an adaptive immune response is increased by a 50-fold factor. In some embodiments, an adaptive immune response is increased by a 100-fold factor. In some embodiments, an adaptive immune response is increased by a 200-fold factor. In some embodiments, an adaptive immune response is increased by a 500-fold factor. In some embodiments, an adaptive immune response is increased by a 1000-fold factor. In some embodiments, an adaptive immune response is increased by a 2000-fold factor. In some embodiments, an adaptive immune response is increased by another fold difference.

In some embodiments, “induces significantly more robust adaptive immune response” refers to a detectable increase in an adaptive immune response. In some embodiments, the term refers to a fold increase in the adaptive immune response (e.g., 1 of the fold increases enumerated above). In some embodiments, the term refers to an increase such that the nucleoside-modified RNA can be administered at a lower dose or frequency than an unmodified RNA molecule while still inducing a similarly effective adaptive immune response. In some embodiments, the increase is such that the nucleoside-modified RNA can be administered using a single dose to induce an effective adaptive immune response.

In some embodiments, the nucleoside-modified RNA of the present disclosure exhibits significantly less innate immunogenicity than an unmodified in vitro-synthesized RNA molecule of the same sequence. In some embodiments, the modified RNA molecule exhibits an innate immune response that is 2-fold less than its unmodified counterpart. In some embodiments, innate immunogenicity is reduced by a 3-fold factor. In some embodiments, innate immunogenicity is reduced by a 4-fold factor. In some embodiments, innate immunogenicity is reduced by a 5-fold factor. In some embodiments, innate immunogenicity is reduced by a 6-fold factor. In some embodiments, innate immunogenicity is reduced by a 7-fold factor. In some embodiments, innate immunogenicity is reduced by an 8-fold factor. In some embodiments, innate immunogenicity is reduced by a 9-fold factor. In some embodiments, innate immunogenicity is reduced by a 10-fold factor. In some

embodiments, innate immunogenicity is reduced by a 15-fold factor. In some embodiments, innate immunogenicity is reduced by a 20-fold factor. In some embodiments, innate immunogenicity is reduced by a 50-fold factor. In some embodiments, innate immunogenicity is reduced by a 100-fold factor. In some
5 embodiments, innate immunogenicity is reduced by a 200-fold factor. In some embodiments, innate immunogenicity is reduced by a 500-fold factor. In some embodiments, innate immunogenicity is reduced by a 1000-fold factor. In some embodiments, innate immunogenicity is reduced by a 2000-fold factor. In some embodiments, innate immunogenicity is reduced by another fold difference.

10 In some embodiments, “exhibits significantly less innate immunogenicity” refers to a detectable decrease in innate immunogenicity. In some embodiments, the term refers to a fold decrease in innate immunogenicity (e.g., 1 of the fold decreases enumerated above). In some embodiments, the term refers to a decrease such that an effective amount of the nucleoside-modified RNA can be administered without triggering
15 a detectable innate immune response. In some embodiments, the term refers to a decrease such that the nucleoside-modified RNA can be repeatedly administered without eliciting an innate immune response sufficient to detectably reduce production of the protein encoded by the modified RNA. In some embodiments, the decrease is such that the nucleoside-modified RNA can be repeatedly administered without eliciting an innate
20 immune response sufficient to eliminate detectable production of the protein encoded by the modified RNA.

Lipid Nanoparticle

In some embodiments, delivery of nucleoside-modified RNA comprises
25 any suitable delivery method, including exemplary RNA transfection methods described elsewhere herein. In some embodiments, delivery of a nucleoside-modified RNA to a subject comprises mixing the nucleoside-modified RNA with a transfection reagent prior to the step of contacting. In some embodiments, a method of present disclosure further comprises administering nucleoside-modified RNA together with a transfection reagent.
30 In some embodiments, a transfection reagent is a cationic lipid reagent. In some embodiments, a transfection reagent is a cationic polymer reagent.

In some embodiments, a transfection reagent is a lipid-based transfection reagent. In some embodiments, a transfection reagent is a protein-based transfection reagent. In some embodiments, a transfection reagent is a carbohydrate-based transfection reagent. In some embodiments, a transfection reagent is a cationic lipid-based transfection reagent. In some embodiments, a transfection reagent is a cationic polymer-based transfection reagent. In some embodiments, a transfection reagent is a polyethyleneimine based transfection reagent. In some embodiments, a transfection reagent is calcium phosphate. In some embodiments, a transfection reagent is Lipofectin[®], Lipofectamine[®], or TransIT[®]. In some embodiments, a transfection reagent is any other transfection reagent known in the art.

In some embodiments, a transfection reagent forms a liposome. In some embodiments, liposomes increase intracellular stability, increase uptake efficiency and improve biological activity. In some embodiments, liposomes are hollow spherical vesicles composed of lipids arranged in a similar fashion as those lipids, which make up a cell membrane. In some embodiments, liposomes have an internal aqueous space for entrapping water-soluble compounds and range in size from 0.05 to several microns in diameter. In some embodiments, liposomes can deliver RNA to cells in a biologically active form.

In some embodiments, a composition comprises a LNP comprising one or more nucleic acid molecules described herein. In some embodiments, a composition comprises an LNP and one or more nucleoside-modified RNA molecules encoding one or more antigens, adjuvants, or a combination thereof. Exemplary adjuvants include, but are not limited to, alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), TNF α , TNF β , GM-CSF, epidermal growth factor (EGF), cutaneous T cell-attracting chemokine (CTACK), epithelial thymus-expressed chemokine (TECK), A proliferation-inducing ligand (APRIL), IL-2, mucosae-associated epithelial chemokine (MEC), IL-12, IL-15, IL-21, MHC, CD80, CD86. Other genes which may be useful adjuvants include those encoding: MCP-I, MIP-Ia, MIP-Ip, IL-8, RANTES, L-selectin, P-selectin, E-selectin, CD34, GlyCAM-1, MadCAM-1, LFA-I, VLA-I, Mac-1, pl50.95, PECAM, ICAM-I, ICAM-2, ICAM-3, CD2, LFA-3, M-CSF, G-CSF, IL-4, mutant forms of IL-18, CD40, CD40L, vascular growth factor, fibroblast growth factor, IL-7, nerve

growth factor, vascular endothelial growth factor, Fas, TNF receptor, Fit, Apo-1, p55, WSL-I, DR3, TRAMP, Apo-3, AIR, LARD, NGRF, DR4, DR5, KILLER, TRAIL-R2, TRICK2, DR6, Caspase ICE, Fos, c-jun, Sp-I, Ap-I, Ap-2, p38, p65Rel, MyD88, IRAK, TRAF6, Ikb, Inactive NIK, SAP K, SAP-I, JNK, interferon response genes, NFkB, Bax, TRAIL, TRAILrec, TRAILrecDRC5, TRAIL-R3, TRAIL-R4, RANK, RANK LIGAND, Ox40, Ox40 LIGAND, NKG2D, MICA, MICB, NKG2A, NKG2B, NKG2C, NKG2E, NKG2F, TAP 1, TAP2, anti-CTLA4-sc, anti-LAG3-Ig, anti-TIM3-Ig, and functional fragments thereof.

In some embodiments, the adjuvant comprises an RNA molecule (e.g., an mRNA) encoding alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), TNF α , TNF β , GM-CSF, epidermal growth factor (EGF), cutaneous T cell-attracting chemokine (CTACK), epithelial thymus-expressed chemokine (TECK), A proliferation-inducing ligand (APRIL), IL-2, mucosae-associated epithelial chemokine (MEC), IL-12, IL-15, IL-21, MHC, CD80, CD86. Other genes which may be useful adjuvants include those encoding: MCP-I, MIP-Ia, MIP-1p, IL-8, RANTES, L-selectin, P-selectin, E-selectin, CD34, GlyCAM-1, MadCAM-1, LFA-I, VLA-I, Mac-1, pl50.95, PECAM, ICAM-I, ICAM-2, ICAM-3, CD2, LFA-3, M-CSF, G-CSF, IL-4, mutant forms of IL-18, CD40, CD40L, vascular growth factor, fibroblast growth factor, IL-7, nerve growth factor, vascular endothelial growth factor, Fas, TNF receptor, Fit, Apo-1, p55, WSL-I, DR3, TRAMP, Apo-3, AIR, LARD, NGRF, DR4, DR5, KILLER, TRAIL-R2, TRICK2, DR6, Caspase ICE, Fos, c-jun, Sp-I, Ap-I, Ap-2, p38, p65Rel, MyD88, IRAK, TRAF6, Ikb, Inactive NIK, SAP K, SAP-I, JNK, interferon response genes, NFkB, Bax, TRAIL, TRAILrec, TRAILrecDRC5, TRAIL-R3, TRAIL-R4, RANK, RANK LIGAND, Ox40, Ox40 LIGAND, NKG2D, MICA, MICB, NKG2A, NKG2B, NKG2C, NKG2E, NKG2F, TAP 1, TAP2, anti-CTLA4-sc, anti-LAG3-Ig, or anti-TIM3-Ig, or functional fragments thereof. In some embodiments, the adjuvant is ASO3, Squalene, liposomes, chitosan, or Matrix M. In some embodiments, the adjuvant is alum.

In some embodiments, the LNP is a particle having at least one dimension on the order of nanometers (e.g., 1-1,000 nm). In some embodiments, a lipid nanoparticle comprises one or more lipids. In some embodiments, a lipid comprises a lipid of Formula (I), (II) or (III).

In some embodiments, lipid nanoparticles are included in a formulation comprising a nucleoside-modified RNA as described herein. In some embodiments, such lipid nanoparticles comprise a cationic lipid (e.g., a lipid of Formula (I), (II) or (III)) and one or more excipient selected from neutral lipids, charged lipids, steroids and polymer conjugated lipids (e.g., a pegylated lipid such as a pegylated lipid of structure (IV)). In some embodiments, the nucleoside-modified RNA is encapsulated in the lipid portion of a lipid nanoparticle or an aqueous space enveloped by some or all of the lipid portion of a lipid nanoparticle, thereby protecting it from enzymatic degradation or other undesirable effects induced by the mechanisms of the host organism or cells, e.g., an adverse immune response.

In some embodiments, lipid nanoparticles have a mean diameter of from about 30 nm to about 150 nm, from about 40 nm to about 150 nm, from about 50 nm to about 150 nm, from about 60 nm to about 130 nm, from about 70 nm to about 110 nm, from about 70 nm to about 100 nm, from about 80 nm to about 100 nm, from about 90 nm to about 100 nm, from about 70 to about 90 nm, from about 80 nm to about 90 nm, from about 70 nm to about 80 nm, or about 30 nm, 35 nm, 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm, 110 nm, 115 nm, 120 nm, 125 nm, 130 nm, 135 nm, 140 nm, 145 nm, or 150 nm, and are substantially non-toxic. In some embodiments, a nucleoside-modified RNA, when present in the lipid nanoparticles, is resistant in aqueous solution to degradation with a nuclease.

In some embodiments, an LNP may comprise any lipid capable of forming a particle to which the one or more RNA molecules are attached, or in which one or more RNA molecules are encapsulated.

In some embodiments, an LNP comprises one or more cationic lipids, and one or more stabilizing lipids. Stabilizing lipids include neutral lipids and pegylated lipids.

In some embodiments, an LNP comprises a cationic lipid. In some embodiments, a cationic lipid comprises any of a number of lipid species which carry a net positive charge at a selective pH, such as physiological pH. In some embodiments, such lipids include, but are not limited to, N,N-dioleoyl-N,N-dimethylammonium chloride

(DODAC); N-(2,3-dioleoyloxy)propyl)-N,N,N-trimethylammonium chloride (DOTMA); N,N-distearyl-N,N-dimethylammonium bromide (DDAB); N-(2,3-dioleoyloxy)propyl)-N,N,N-trimethylammonium chloride (DOTAP); 3-(N—(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), N-(1-(2,3-dioleoyloxy)propyl)-N-2-

5 (sperminecarboxamido)ethyl)-N,N-dimethylammonium trifluoroacetate (DOSPA), dioctadecylamidoglycyl carboxyspermine (DOGS), 1,2-dioleoyl-3-dimethylammonium propane (DODAP), N,N-dimethyl-2,3-dioleoyloxy)propylamine (DODMA), and N-(1,2-dimyrityloxyprop-3-yl)-N,N-dimethyl-N-hydroxyethyl ammonium bromide (DMRIE). Additionally, a number of commercial preparations of cationic lipids are available which

10 can be used in the present disclosure. In some embodiments, these include LIPOFECTIN[®] (commercially available cationic liposomes comprising DOTMA and 1,2-dioleoyl-sn-3-phosphoethanolamine (DOPE), from GIBCO/BRL, Grand Island, N.Y.); LIPOFECTAMINE[®] (commercially available cationic liposomes comprising N-(1-(2,3-dioleoyloxy)propyl)-N-(2-(sperminecarboxamido)ethyl)-N,N-dimethylammonium

15 trifluoroacetate (DOSPA) and (DOPE), from GIBCO/BRL); and TRANSFECTAM[®] (commercially available cationic lipids comprising dioctadecylamidoglycyl carboxyspermine (DOGS) in ethanol from Promega Corp., Madison, Wis.). In some embodiments, the following lipids are cationic and have a positive charge at below physiological pH: DODAP, DODMA, DMDMA, 1,2-dilinoleoyloxy-N,N-

20 dimethylaminopropane (DLinDMA), 1,2-dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA).

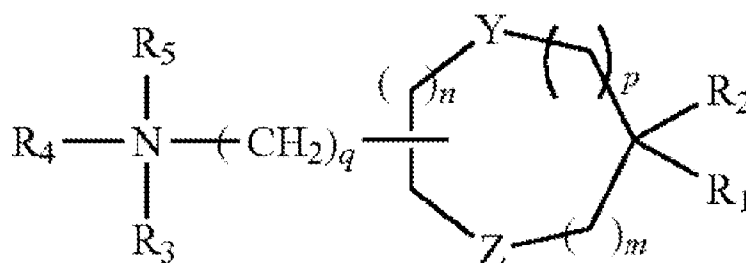
In some embodiments, a cationic lipid is an amino lipid. Suitable amino lipids useful in the present disclosure include those described in WO 2012/016184, incorporated herein by reference in its entirety. Representative amino lipids include, but

25 are not limited to, 1,2-dilinoleoyloxy-3-(dimethylamino)acetoxyp propane (DLin-DAC), 1,2-dilinoleoyloxy-3-morpholinopropane (DLin-MA), 1,2-dilinoleoyl-3-dimethylaminopropane (DLinDAP), 1,2-dilinoleylthio-3-dimethylaminopropane (DLin-S-DMA), 1-linoleoyl-2-linoleoyloxy-3-dimethylaminopropane (DLin-2-DMAP), 1,2-dilinoleoyloxy-3-trimethylaminopropane chloride salt (DLin-TMA.Cl), 1,2-dilinoleoyl-3-

30 trimethylaminopropane chloride salt (DLin-TAP.Cl), 1,2-dilinoleoyloxy-3-(N-methylpiperazino)propane (DLin-MPZ), 3-(N,N-dilinoleylamino)-1,2-propanediol

(DLinAP), 3-(N,N-dioleoylamino)-1,2-propanediol (DOAP), 1,2-dilinoleyloxo-3-(2-N,N-dimethylamino)ethoxypropane (DLin-EG-DMA), and 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA).

In some embodiments, suitable amino lipids include those having the
5 formula:



wherein R₁ and R₂ are either the same or different and independently optionally substituted C₁₀-C₂₄ alkyl, optionally substituted C₁₀-C₂₄ alkenyl, optionally substituted C₁₀-C₂₄ alkynyl, or optionally substituted C₁₀-C₂₄ acyl;

10 R₃ and R₄ are either the same or different and independently optionally substituted C₁-C₆ alkyl, optionally substituted C₂-C₆ alkenyl, or optionally substituted C₂-C₆ alkynyl or R₃ and R₄ may join to form an optionally substituted heterocyclic ring of 4 to 6 carbon atoms and 1 or 2 heteroatoms chosen from nitrogen and oxygen;

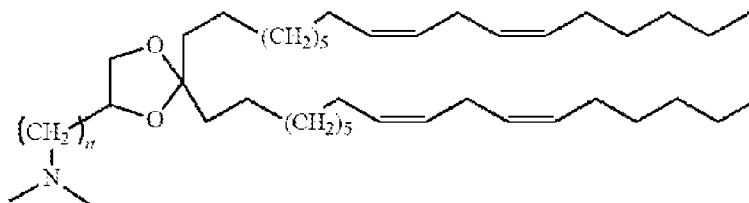
R₅ is either absent or present and when present is hydrogen or C₁-C₆ alkyl;
15 m, n, and p are either the same or different and independently either 0 or 1 with the proviso that m, n, and p are not simultaneously 0;

q is 0, 1, 2, 3, or 4; and

Y and Z are either the same or different and independently O, S, or NH.

In some embodiments, R₁ and R₂ are each linoleyl, and the amino lipid is a
20 dilinoleyl amino lipid. In some embodiments, the amino lipid is a dilinoleyl amino lipid.

In some embodiments, a representative useful dilinoleyl amino lipid has the formula:

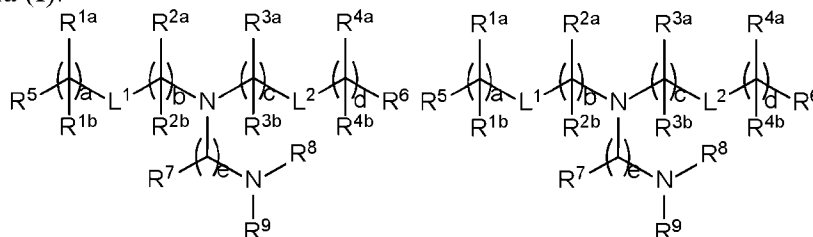


DLin-K-DMA

wherein n is 0, 1, 2, 3, or 4.

In some embodiments, a cationic lipid is a DLin-K-DMA. In some embodiments, a cationic lipid is DLin-KC2-DMA (DLin-K-DMA above, wherein n is 2).

5 In some embodiments, a cationic lipid component of a LNP has the structure of Formula (I):



(I)

or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

L^1 and L^2 are each independently $-O(C=O)-$, $-(C=O)O-$ or a carbon-carbon double bond;

R^{1a} and R^{1b} are, at each occurrence, independently either (a) H or C_1 - C_{12} alkyl, or (b) R^{1a} is H or C_1 - C_{12} alkyl, and R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{2a} and R^{2b} are, at each occurrence, independently either (a) H or C_1 - C_{12} alkyl, or (b) R^{2a} is H or C_1 - C_{12} alkyl, and R^{2b} together with the carbon atom to which it is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{3a} and R^{3b} are, at each occurrence, independently either (a) H or C_1 - C_{12} alkyl, or (b) R^{3a} is H or C_1 - C_{12} alkyl, and R^{3b} together with the carbon atom to which it is bound

bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{4a} and R^{4b} are, at each occurrence, independently either (a) H or C_1 - C_{12} alkyl, or (b) R^{4a} is H or C_1 - C_{12} alkyl, and R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^5 and R^6 are each independently methyl or cycloalkyl;

R^7 is, at each occurrence, independently H or C_1 - C_{12} alkyl;

R^8 and R^9 are each independently C_1 - C_{12} alkyl; or R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring comprising one nitrogen atom;

a and d are each independently an integer from 0 to 24;

b and c are each independently an integer from 1 to 24; and

e is 1 or 2.

In some embodiments of Formula (I), at least one of R^{1a} , R^{2a} , R^{3a} or R^{4a} is C_1 - C_{12} alkyl, or at least one of L^1 or L^2 is $-O(C=O)-$ or $-(C=O)O-$. In some embodiments, R^{1a} and R^{1b} are not isopropyl when a is 6 or n-butyl when a is 8.

In some embodiments of Formula (I), at least one of R^{1a} , R^{2a} , R^{3a} or R^{4a} is C_1 - C_{12} alkyl, or at least one of L^1 or L^2 is $-O(C=O)-$ or $-(C=O)O-$; and

R^{1a} and R^{1b} are not isopropyl when a is 6 or n-butyl when a is 8.

In some embodiments of Formula (I), R^8 and R^9 are each independently unsubstituted C_1 - C_{12} alkyl; or R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring comprising one nitrogen atom;

In some embodiments of Formula (I), any one of L^1 or L^2 may be $-O(C=O)-$ or a carbon-carbon double bond. L^1 and L^2 may each be $-O(C=O)-$ or may each be a carbon-carbon double bond.

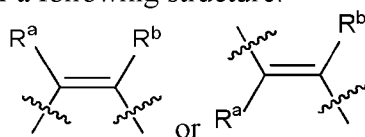
In some embodiments of Formula (I), one of L^1 or L^2 is $-O(C=O)-$. In some embodiments, both L^1 and L^2 are $-O(C=O)-$.

In some embodiments of Formula (I), one of L^1 or L^2 is $-(C=O)O-$. In some embodiments, both L^1 and L^2 are $-(C=O)O-$.

In some embodiments of Formula (I), one of L^1 or L^2 is a carbon-carbon double bond. In some embodiments, both L^1 and L^2 are a carbon-carbon double bond.

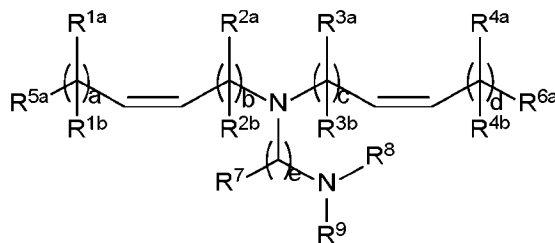
In some embodiments of Formula (I), one of L^1 or L^2 is $-O(C=O)-$ and the other of L^1 or L^2 is $-(C=O)O-$. In some embodiments, one of L^1 or L^2 is $-O(C=O)-$ and the other of L^1 or L^2 is a carbon-carbon double bond. In some embodiments, one of L^1 or L^2 is $-(C=O)O-$ and the other of L^1 or L^2 is a carbon-carbon double bond.

In some embodiments “carbon-carbon” double bond, as used throughout the present disclosure, refers to one of a following structure:



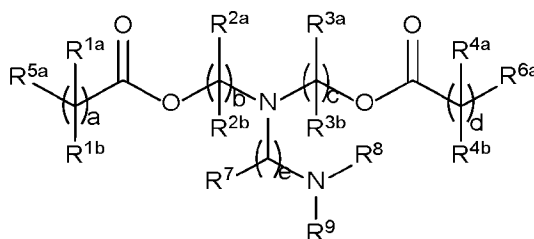
wherein R^a and R^b are, at each occurrence, independently H or a substituent. For example, in some embodiments R^a and R^b are, at each occurrence, independently H, C_1 - C_{12} alkyl or cycloalkyl, for example H or C_1 - C_{12} alkyl.

In some embodiments, a lipid compounds of Formula (I) have the following structure (Ia):



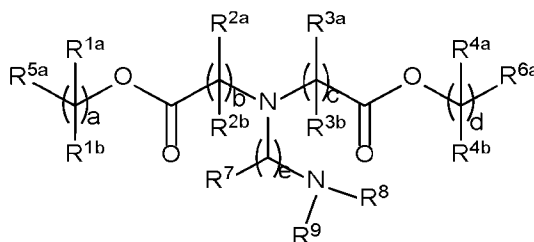
(Ia)

In some embodiments, a lipid compounds of Formula (I) have the following structure (Ib):



(Ib)

In some embodiments, a lipid compounds of Formula (I) have the following structure (Ic):



(Ic)

In some embodiments, a lipid compound of Formula (I), comprises a, b, c and d, wherein each are independently an integer from 2 to 12 or an integer from 4 to 12.

- 5 In some embodiments, a, b, c and d are each independently an integer from 8 to 12 or 5 to 9. In some embodiments, a is 0. In some embodiments, a is 1. In some embodiments, a is 2. In more embodiments, a is 3. In some embodiments, a is 4. In some embodiments, a is 5. In some embodiments, a is 6. In some embodiments, a is 7. In some embodiments, a is 8. In some embodiments, a is 9. In some embodiments, a is 10. In some embodiments, a is 11. In some embodiments, a is 12. In some embodiments, a is 13. In some
10 embodiments, a is 14. In some embodiments, a is 15. In some embodiments, a is 16.

- In some embodiments of Formula (I), b is 1. In some embodiments, b is 2. In some embodiments, b is 3. In some embodiments, b is 4. In some embodiments, b is 5. In some embodiments, b is 6. In some embodiments, b is 7. In some embodiments, b is 8.
15 In some embodiments, b is 9. In some embodiments, b is 10. In some embodiments, b is 11. In some embodiments, b is 12. In some embodiments, b is 13. In some embodiments, b is 14. In some embodiments, b is 15. In some embodiments, b is 16.

- In some embodiments of Formula (I), c is 1. In some embodiments, c is 2. In some embodiments, c is 3. In some embodiments, c is 4. In some embodiments, c is 5.
20 In some embodiments, c is 6. In some embodiments, c is 7. In some embodiments, c is 8. In some embodiments, c is 9. In some embodiments, c is 10. In some embodiments, c is 11. In some embodiments, c is 12. In some embodiments, c is 13. In some embodiments, c is 14. In some embodiments, c is 15. In some embodiments, c is 16.

- In some embodiments of Formula (I), d is 0. In some embodiments, d is 1.
25 In some embodiments, d is 2. In some embodiments, d is 3. In some embodiments, d is 4. In some embodiments, d is 5. In some embodiments, d is 6. In some embodiments, d is 7.

In some embodiments, d is 8. In some embodiments, d is 9. In some embodiments, d is 10. In some embodiments, d is 11. In some embodiments, d is 12. In some embodiments, d is 13. In some embodiments, d is 14. In some embodiments, d is 15. In some embodiments, d is 16.

5 In some embodiments, a and d of Formula (I) are the same. In some embodiments, b and c are the same. In some embodiments, a and d are the same and b and c are the same.

 In some embodiments, the sum of a and b and sum of c and d in Formula (I) are factors which may be varied to obtain a lipid of Formula (I) having desired
10 properties. In some embodiments, a and b are chosen such that their sum is an integer ranging from 14 to 24. In some embodiments, c and d are chosen such that their sum is an integer ranging from 14 to 24. In some embodiments, the sum of a and b and the sum of c and d are the same. In some embodiments, the sum of a and b and the sum of c and d are both the same integer which may range from 14 to 24. In some embodiments, a, b, c and
15 d are selected such the sum of a and b and the sum of c and d is 12 or greater.

 In some embodiments, e is 1. In some embodiments, e is 2.

 In some embodiments, I substituents at R^{1a} , R^{2a} , R^{3a} and R^{4a} of Formula (I) are not particularly limited. In some embodiments, R^{1a} , R^{2a} , R^{3a} and R^{4a} are H at each occurrence. In some embodiments, at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C₁-C₁₂ alkyl. In
20 some embodiments, at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C₁-C₈ alkyl. In some embodiments, at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C₁-C₆ alkyl. In some embodiments, the C₁-C₈ alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

 In some embodiments, of Formula (I), R^{1a} , R^{1b} , R^{4a} and R^{4b} are C₁-C₁₂
25 alkyl at each occurrence.

 In some embodiments, of Formula (I), at least one of R^{1b} , R^{2b} , R^{3b} and R^{4b} is H or R^{1b} , R^{2b} , R^{3b} and R^{4b} are H at each occurrence.

 In some embodiments, of Formula (I), R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which
30 it is bound to form a carbon-carbon double bond. In some embodiments of the foregoing

R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

The substituents at R^5 and R^6 of Formula (I) are not particularly limited in the foregoing embodiments. In some embodiments, one or both of R^5 or R^6 is methyl. In some embodiments one or both of R^5 or R^6 is cycloalkyl for example cyclohexyl. In some embodiments, a cycloalkyl may be substituted or not substituted. In some embodiments a cycloalkyl is substituted with C_1 - C_{12} alkyl, for example tert-butyl.

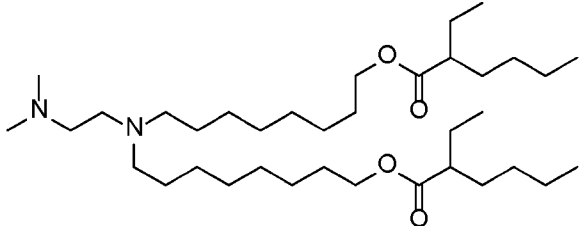
The substituents at R^7 are not particularly limited in the foregoing embodiments of Formula (I). In some embodiments, at least one R^7 is H. In some embodiments, R^7 is H at each occurrence. In some embodiments R^7 is C_1 - C_{12} alkyl.

In some embodiments of Formula (I), one of R^8 or R^9 is methyl. In some embodiments, both R^8 and R^9 are methyl.

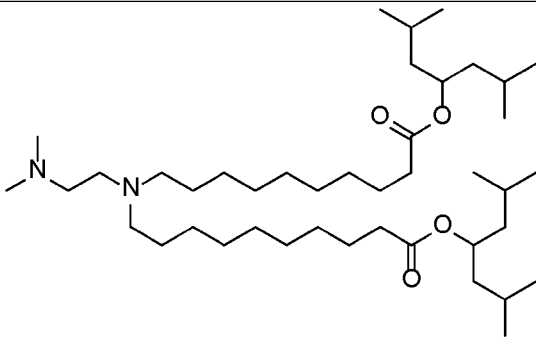
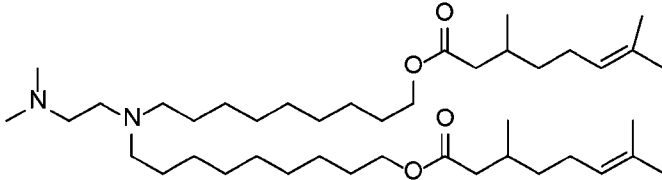
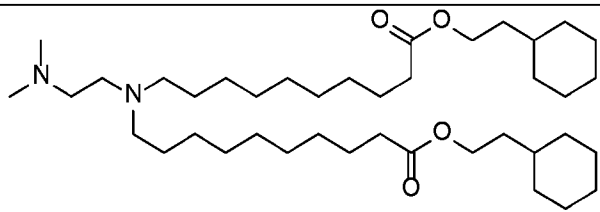
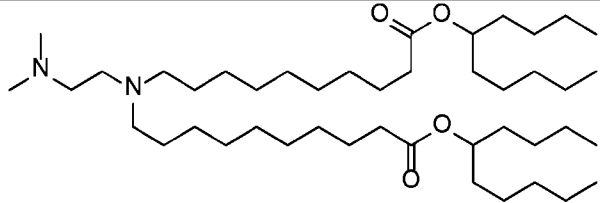
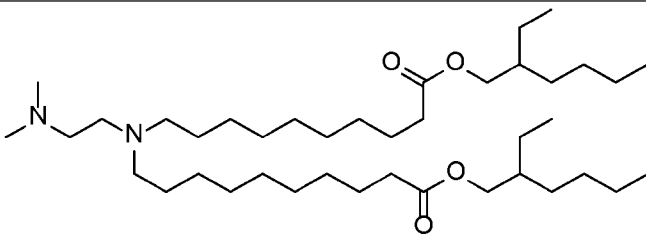
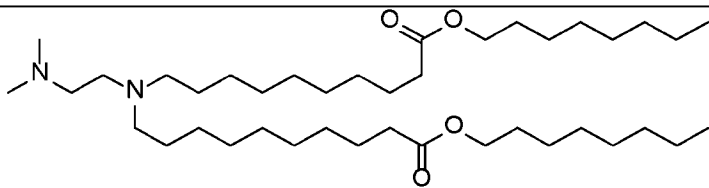
In some embodiments of Formula (I), R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring. In some embodiments, R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5-membered heterocyclic ring, for example a pyrrolidinyl ring.

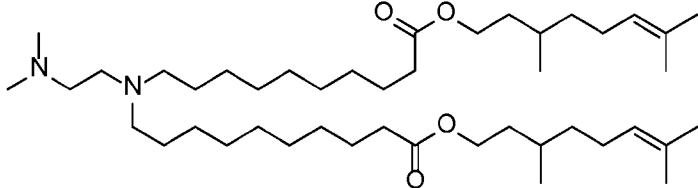
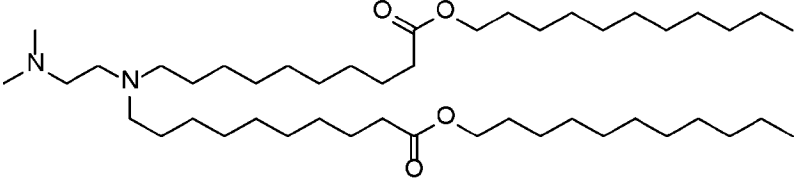
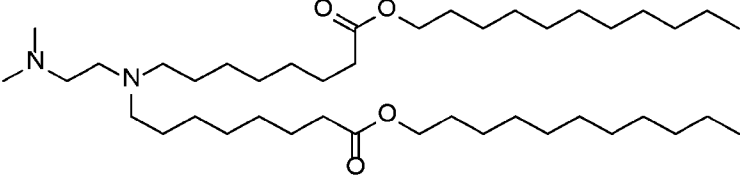
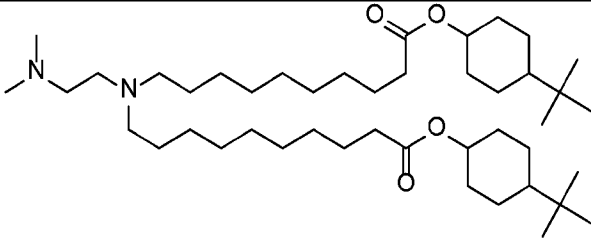
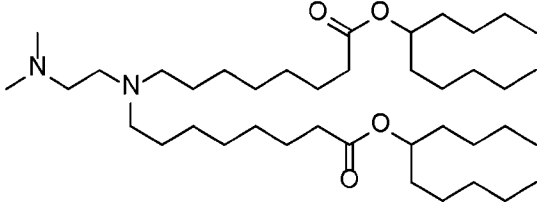
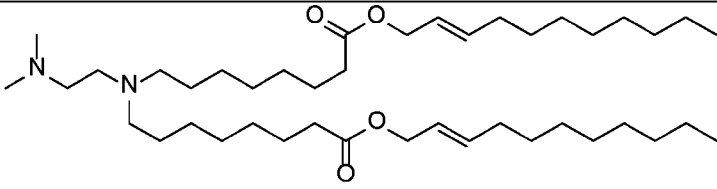
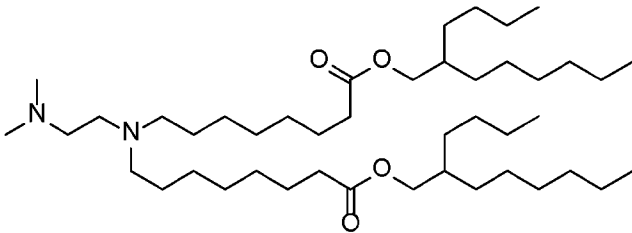
In some embodiments, the lipid of Formula (I) has one of the structures set forth in Table 3 below.

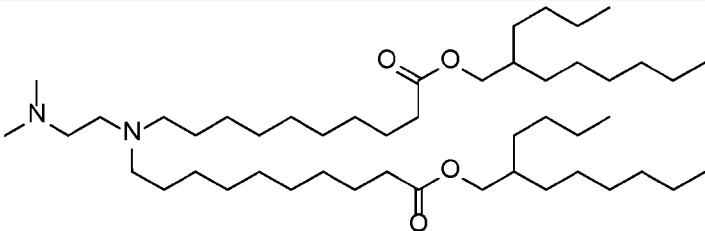
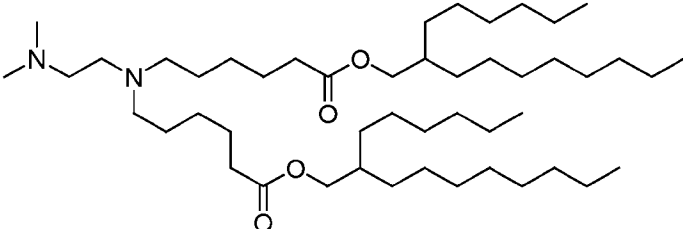
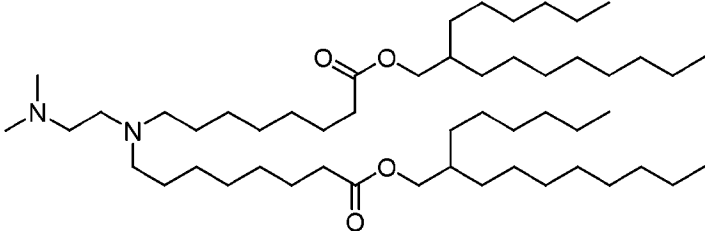
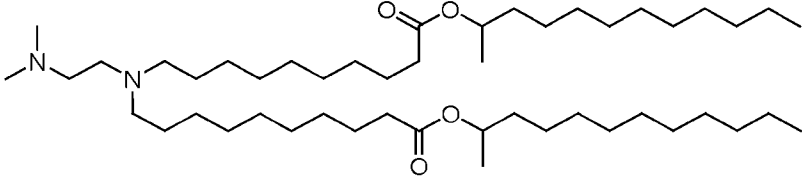
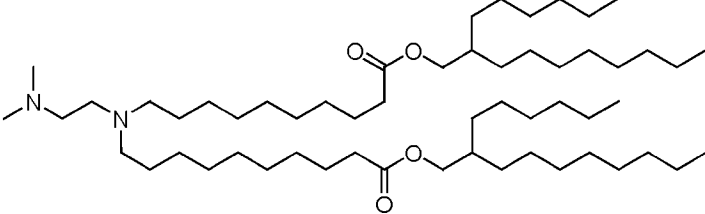
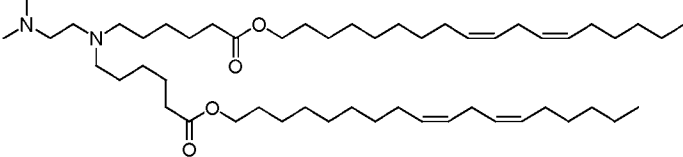
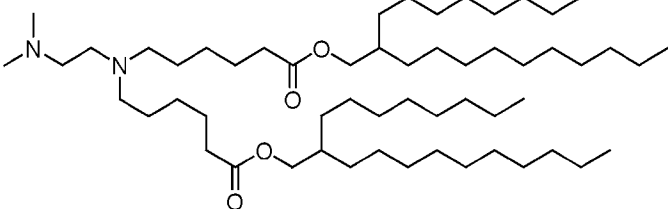
Table 3: Representative Lipids of Formula (I).

No.	Structure	Prep. Method
I-1		B

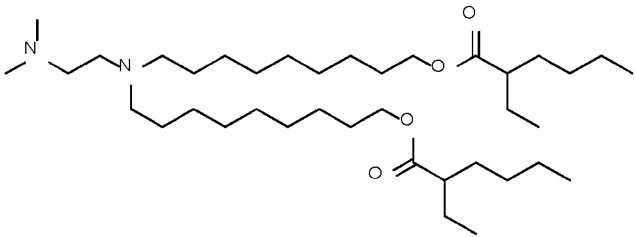
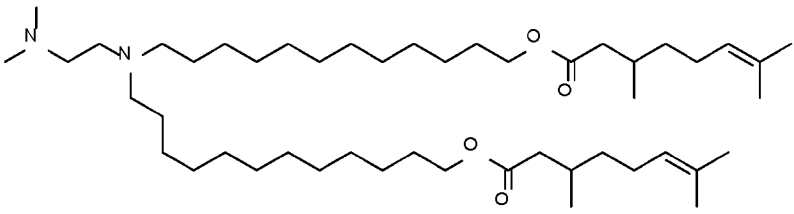
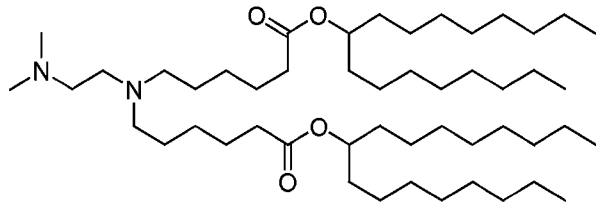
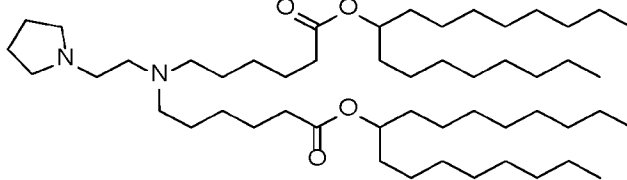
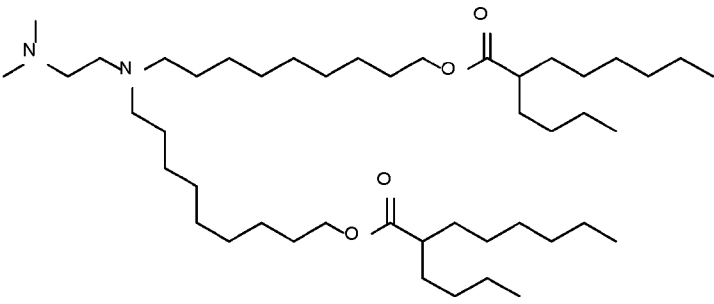
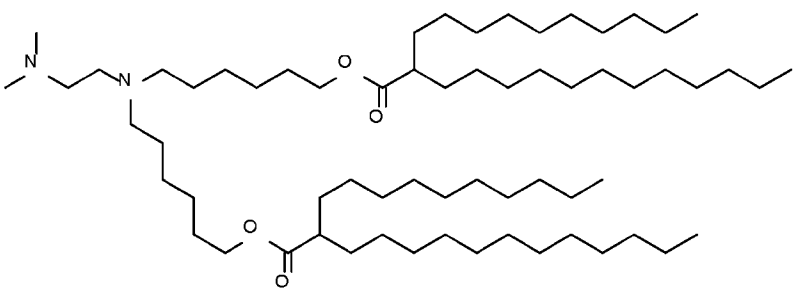
No.	Structure	Prep. Method
I-2		A
I-3		A
I-4		B
I-5		B
I-6		B
I-7		A

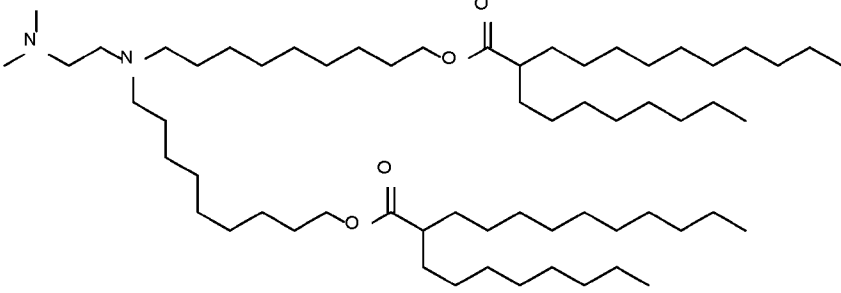
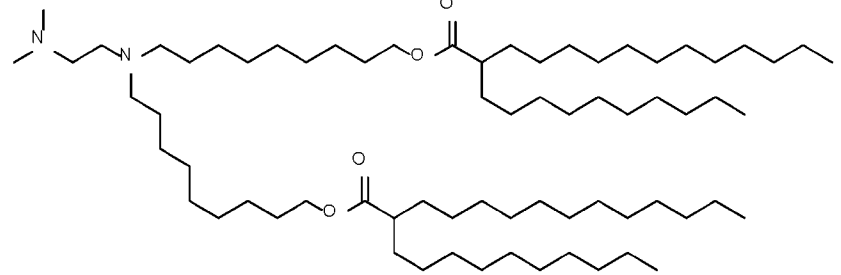
No.	Structure	Prep. Method
I-8		A
I-9		B
I-10		A
I-11		A
I-12		A
I-13		A

No.	Structure	Prep. Method
I-14		A
I-15		A
I-16		A
I-17		A
I-18		A
I-19		A
I-20		A

No.	Structure	Prep. Method
I-21		A
I-22		A
I-23		A
I-24		A
I-25		A
I-26		A
I-27		A

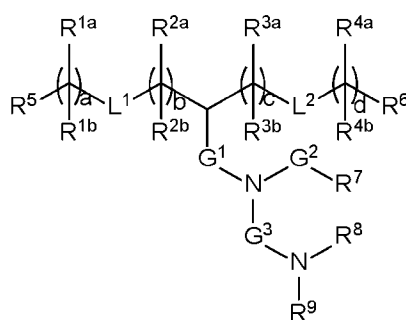
No.	Structure	Prep. Method
I-28		A
I-29		A
I-30		A
I-31		C
I-32		C
I-33		C

No.	Structure	Prep. Method
I-34		B
I-35		B
I-36		C
I-37		C
I-38		B
I-39		B

No.	Structure	Prep. Method
I-40		B
I-41		B

In some embodiments, LNPs comprise a lipid of Formula (I), a nucleoside-modified RNA and one or more excipients selected from neutral lipids, steroids and pegylated lipids. In some embodiments, a lipid of Formula (I) is compound I-5. In some embodiments, a lipid of Formula (I) is compound I-6.

In some embodiments, a cationic lipid component of an LNP has the structure of Formula (II):



(II)

or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

L^1 and L^2 are each independently $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$,

-S(O)_x-, -S-S-, -C(=O)S-, -SC(=O)-, -NR^aC(=O)-, -C(=O)NR^a-, -NR^aC(=O)NR^a,

-OC(=O)NR^a-, -NR^aC(=O)O-, or a direct bond;

G¹ is C₁-C₂ alkylene, -(C=O)-, -O(C=O)-, -SC(=O)-, -NR^aC(=O)- or a
5 direct bond;

G² is -C(=O)-, -(C=O)O-, -C(=O)S-, -C(=O)NR^a or a direct bond;

G³ is C₁-C₆ alkylene;

R^a is H or C₁-C₁₂ alkyl;

R^{1a} and R^{1b} are, at each occurrence, independently either: (a) H or C₁-C₁₂
10 alkyl; or (b) R^{1a} is H or C₁-C₁₂ alkyl, and R^{1b} together with the carbon atom to which it is
bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to
form a carbon-carbon double bond;

R^{2a} and R^{2b} are, at each occurrence, independently either: (a) H or C₁-C₁₂
alkyl; or (b) R^{2a} is H or C₁-C₁₂ alkyl, and R^{2b} together with the carbon atom to which it is
15 bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to
form a carbon-carbon double bond;

R^{3a} and R^{3b} are, at each occurrence, independently either: (a) H or C₁-C₁₂
alkyl; or (b) R^{3a} is H or C₁-C₁₂ alkyl, and R^{3b} together with the carbon atom to which it is
bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to
20 form a carbon-carbon double bond;

R^{4a} and R^{4b} are, at each occurrence, independently either: (a) H or C₁-C₁₂
alkyl; or (b) R^{4a} is H or C₁-C₁₂ alkyl, and R^{4b} together with the carbon atom to which it is
bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to
form a carbon-carbon double bond;

25 R⁵ and R⁶ are each independently H or methyl;

R⁷ is C₄-C₂₀ alkyl;

R⁸ and R⁹ are each independently C₁-C₁₂ alkyl; or R⁸ and R⁹, together with
the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic
ring;

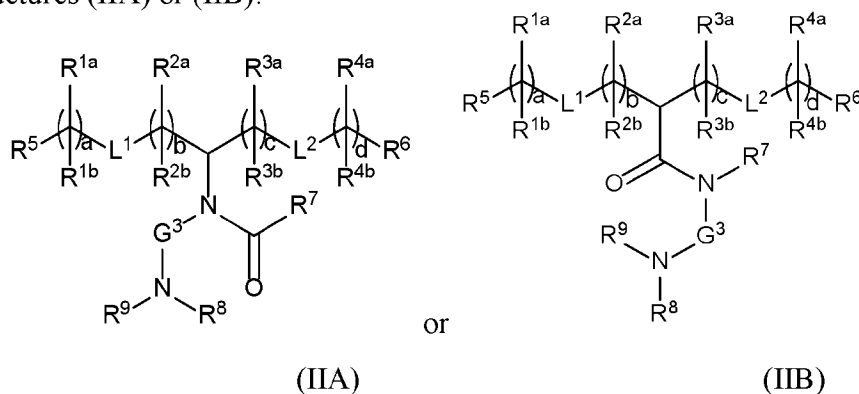
30 a, b, c and d are each independently an integer from 1 to 24; and

x is 0, 1 or 2.

In some embodiments of Formula (II), L^1 and L^2 are each independently $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$ or a direct bond. In some embodiments, G^1 and G^2 are each independently $-(\text{C}=\text{O})-$ or a direct bond. In some embodiments, L^1 and L^2 are
 5 each independently $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$ or a direct bond; and G^1 and G^2 are each independently $-(\text{C}=\text{O})-$ or a direct bond.

In some embodiments of Formula (II), L^1 and L^2 are each independently $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $-\text{SC}(\text{=O})-$, $-\text{NR}^a-$, $-\text{NR}^a\text{C}(\text{=O})-$,
 10 $-\text{C}(\text{=O})\text{NR}^a-$, $-\text{NR}^a\text{C}(\text{=O})\text{NR}^a-$, $-\text{OC}(\text{=O})\text{NR}^a-$, $-\text{NR}^a\text{C}(\text{=O})\text{O}-$, $-\text{NR}^a\text{S}(\text{O})_x\text{N}$
 R^a- ,
 $-\text{NR}^a\text{S}(\text{O})_x-$ or $-\text{S}(\text{O})_x\text{NR}^a-$.

In some embodiments of Formula (II), a lipid compound has one of the following structures (IIA) or (IIB):



In some embodiments of Formula (II), a lipid compound has structure (IIA). In some embodiments, a lipid compound has structure (IIB).

In some embodiments of Formula (II), one of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$. For example, in some embodiments, each of L^1 and L^2 are $-\text{O}(\text{C}=\text{O})-$.

20 In some embodiments of Formula (II), one of L^1 or L^2 is $-(\text{C}=\text{O})\text{O}-$. For example, in some embodiments, each of L^1 and L^2 is $-(\text{C}=\text{O})\text{O}-$.

In some embodiments of Formula (II), one of L^1 or L^2 is a direct bond. As used herein, a “direct bond” means the group (e.g., L^1 or L^2) is absent. For example, in some embodiments each of L^1 and L^2 is a direct bond.

25 In some embodiments of Formula (II), for at least one occurrence of R^{1a}

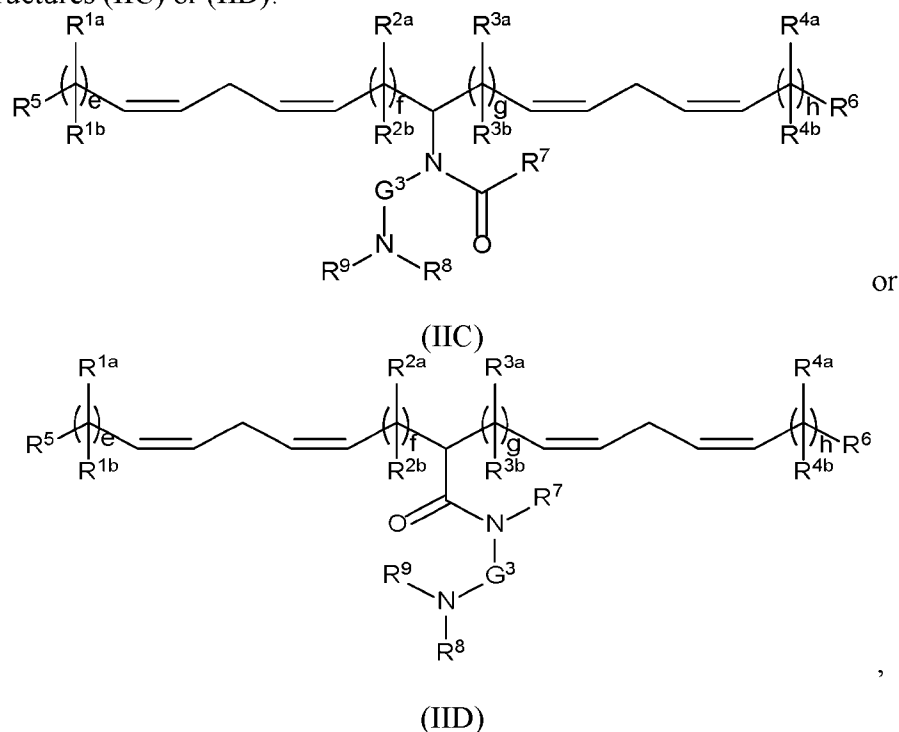
and R^{1b} , R^{1a} is H or C_1 - C_{12} alkyl, and R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In some embodiments of Formula (II), for at least one occurrence of R^{4a} and R^{4b} , R^{4a} is H or C_1 - C_{12} alkyl, and R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In some embodiments of Formula (II), for at least one occurrence of R^{2a} and R^{2b} , R^{2a} is H or C_1 - C_{12} alkyl, and R^{2b} together with the carbon atom to which it is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In some embodiments of Formula (II), for at least one occurrence of R^{3a} and R^{3b} , R^{3a} is H or C_1 - C_{12} alkyl, and R^{3b} together with the carbon atom to which it is bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In some embodiments of Formula (II), a lipid compound has one of the following structures (IIC) or (IID):



wherein e, f, g and h are each independently an integer from 1 to 12.

In some embodiments of Formula (II), the lipid compound has structure (IIC). In some embodiments, the lipid compound has structure (IID).

In some embodiments of structures (IIC) or (IID), e, f, g and h are each
5 independently an integer from 4 to 10.

In some embodiments of Formula (II), a, b, c and d are each independently an integer from 2 to 12 or an integer from 4 to 12. In some embodiments, a, b, c and d are each independently an integer from 8 to 12 or 5 to 9. In some embodiments, a is 0. In some embodiments, a is 1. In some embodiments, a is 2. In more embodiments, a is 3. In
10 some embodiments, a is 4. In some embodiments, a is 5. In some embodiments, a is 6. In more embodiments, a is 7. In some embodiments, a is 8. In some embodiments, a is 9. In some embodiments, a is 10. In some embodiments, a is 11. In some embodiments, a is 12. In some embodiments, a is 13. In some embodiments, a is 14. In some embodiments, a is 15. In some embodiments, a is 16.

In some embodiments of Formula (II), b is 1. In some embodiments, b is
15 2. In some embodiments, b is 3. In yet some embodiments, b is 4. In some embodiments, b is 5. In some embodiments, b is 6. In some embodiments, b is 7. In some embodiments, b is 8. In some embodiments, b is 9. In some embodiments, b is 10. In some embodiments, b is 11. In some embodiments, b is 12. In some embodiments, b is 13. In
20 some embodiments, b is 14. In some embodiments, b is 15. In some embodiments, b is 16.

In some embodiments of Formula (II), c is 1. In some embodiments, c is 2. In some embodiments, c is 3. In some embodiments, c is 4. In some embodiments, c is 5. In some embodiments, c is 6. In some embodiments, c is 7. In some embodiments, c is 8.
25 In some embodiments, c is 9. In some embodiments, c is 10. In some embodiments, c is 11. In some embodiments, c is 12. In some embodiments, c is 13. In some embodiments, c is 14. In some embodiments, c is 15. In some embodiments, c is 16.

In some embodiments of Formula (II), d is 0. In some embodiments, d is
30 1. In some embodiments, d is 2. In some embodiments, d is 3. In some embodiments, d is 4. In some embodiments, d is 5. In some embodiments, d is 6. In some embodiments, d is

7. In some embodiments, d is 8. In some embodiments, d is 9. In some embodiments, d is 10. In some embodiments, d is 11. In some embodiments, d is 12. In some embodiments, d is 13. In some embodiments, d is 14. In some embodiments, d is 15. In some embodiments, d is 16.

5 In some embodiments of Formula (II), e is 1. In some embodiments, e is 2. In some embodiments, e is 3. In some embodiments, e is 4. In some embodiments, e is 5. In some embodiments, e is 6. In some embodiments, e is 7. In some embodiments, e is 8. In some embodiments, e is 9. In some embodiments, e is 10. In some embodiments, e is 11. In some embodiments, e is 12.

10 In some embodiments of Formula (II), f is 1. In some embodiments, f is 2. In some embodiments, f is 3. In some embodiments, f is 4. In some embodiments, f is 5. In some embodiments, f is 6. In some embodiments, f is 7. In some embodiments, f is 8. In some embodiments, f is 9. In some embodiments, f is 10. In some embodiments, f is 11. In some embodiments, f is 12.

15 In some embodiments of Formula (II), g is 1. In some embodiments, g is 2. In some embodiments, g is 3. In some embodiments, g is 4. In some embodiments, g is 5. In some embodiments, g is 6. In some embodiments, g is 7. In some embodiments, g is 8. In some embodiments, g is 9. In some embodiments, g is 10. In some embodiments, g is 11. In some embodiments, g is 12.

20 In some embodiments of Formula (II), h is 1. In some embodiments, e is 2. In some embodiments, h is 3. In some embodiments, h is 4. In some embodiments, e is 5. In some embodiments, h is 6. In some embodiments, h is 7. In some embodiments, h is 8. In some embodiments, h is 9. In some embodiments, h is 10. In some embodiments, h is 11. In some embodiments, h is 12.

25 In some embodiments of Formula (II), a and d are the same. In some embodiments, b and c are the same. In some embodiments and a and d are the same and b and c are the same.

 In some embodiments, the sum of a and b and the sum of c and d of Formula (II) are factors which may be varied to obtain a lipid having the desired
30 properties. In some embodiments, a and b are chosen such that their sum is an integer

ranging from 14 to 24. In some embodiments, c and d are chosen such that their sum is an integer ranging from 14 to 24. In some embodiments, the sum of a and b and the sum of c and d are the same. In some embodiments, the sum of a and b and the sum of c and d are both the same integer which may range from 14 to 24. In some embodiments, a, b, c and d are selected such that the sum of a and b and the sum of c and d is 12 or greater.

In some embodiments, the substituents at R^{1a} , R^{2a} , R^{3a} and R^{4a} of Formula (II) are not particularly limited. In some embodiments, at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is H. In some embodiments R^{1a} , R^{2a} , R^{3a} and R^{4a} are H at each occurrence. In some embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C₁-C₁₂ alkyl. In some embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C₁-C₈ alkyl. In some embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C₁-C₆ alkyl. In some of the foregoing embodiments, the C₁-C₈ alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In some embodiments of Formula (II), R^{1a} , R^{1b} , R^{4a} and R^{4b} are C₁-C₁₂ alkyl at each occurrence.

In some embodiments of Formula (II), at least one of R^{1b} , R^{2b} , R^{3b} and R^{4b} is H or R^{1b} , R^{2b} , R^{3b} and R^{4b} are H at each occurrence.

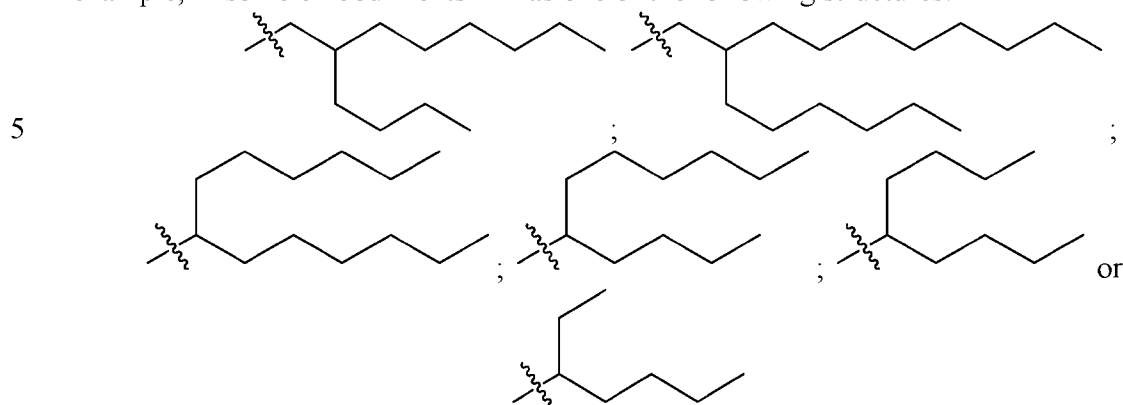
In some embodiments of Formula (II), R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond. In some embodiments of the foregoing R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In some embodiments, the substituents at R^5 and R^6 of Formula (II) are not particularly limited in the foregoing embodiments. In some embodiments, one of R^5 or R^6 is methyl. In some embodiments each of R^5 or R^6 is methyl.

In some embodiments, the substituents at R^7 of Formula (II) are not particularly limited in the foregoing embodiments. In some embodiments R^7 is C₆-C₁₆ alkyl. In some embodiments, R^7 is C₆-C₉ alkyl. In some embodiments, R^7 is substituted with $-(C=O)OR^b$, $-O(C=O)R^b$, $-C(=O)R^b$, $-OR^b$, $-S(O)_xR^b$, $-S-SR^b$, $-C(=O)SR^b$, $-SC(=O)R^b$, $-NR^aR^b$, $-NR^aC(=O)R^b$, $-C(=O)NR^aR^b$, $-NR^aC(=O)NR^aR^b$, $-OC(=O)NR^aR^b$, $-NR^aC(=O)OR^b$, $-NR^aS(O)_xNR^aR^b$, $-NR^aS(O)_xR^b$

or $-S(O)_xNR^aR^b$, wherein: R^a is H or C_1 - C_{12} alkyl; R^b is C_1 - C_{15} alkyl; and x is 0, 1 or 2. For example, in some embodiments, R^7 is substituted with $-(C=O)OR^b$ or $-O(C=O)R^b$.

In some embodiments of Formula (II), R^b is branched C_1 - C_{15} alkyl. For example, in some embodiments R^b has one of the following structures:



In some embodiments of Formula (II), one of R^8 or R^9 is methyl. In some embodiments, both R^8 and R^9 are methyl.

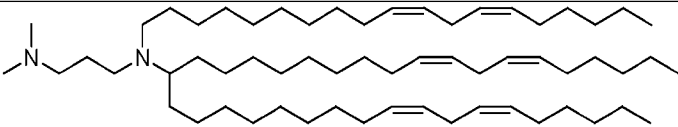
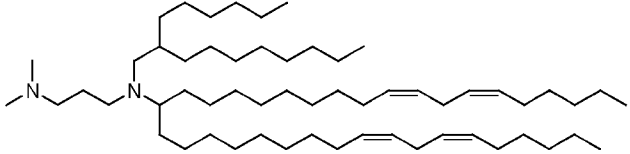
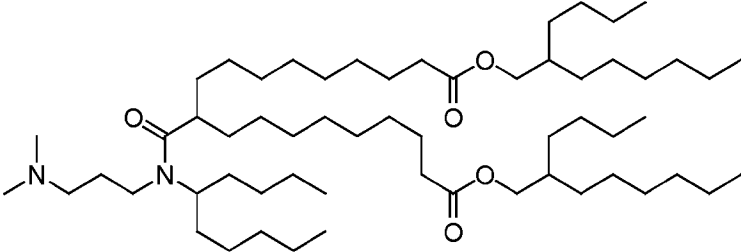
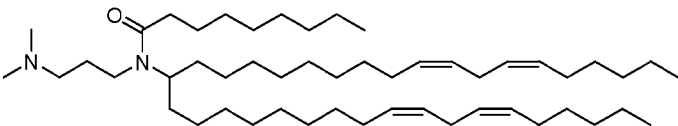
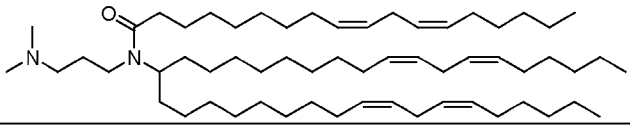
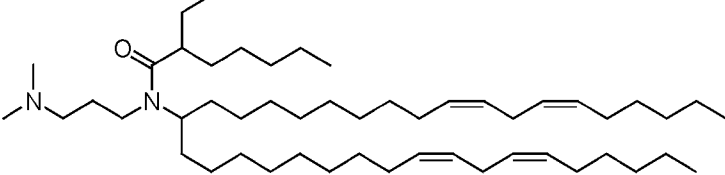
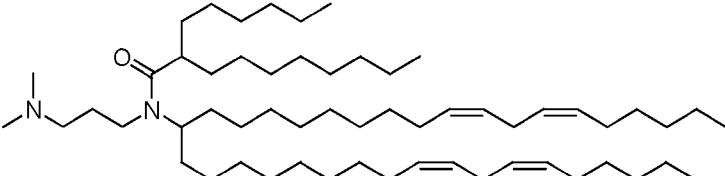
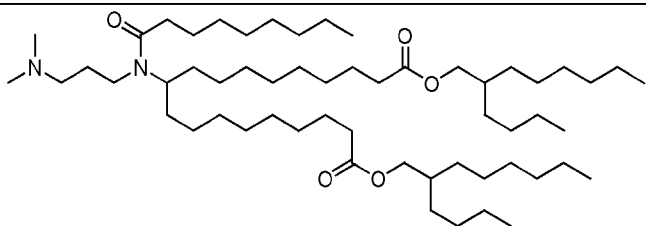
10 In some embodiments of Formula (II), R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring. In some embodiments, R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5-membered heterocyclic ring, for example a pyrrolidinyl ring. In some
15 form a 6-membered heterocyclic ring, for example a piperazinyl ring.

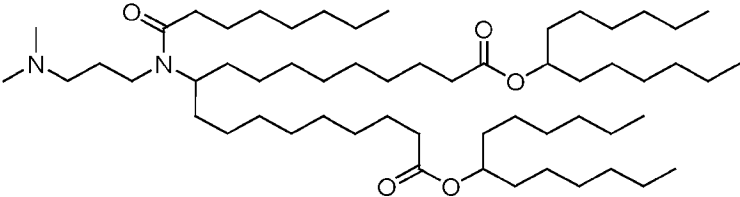
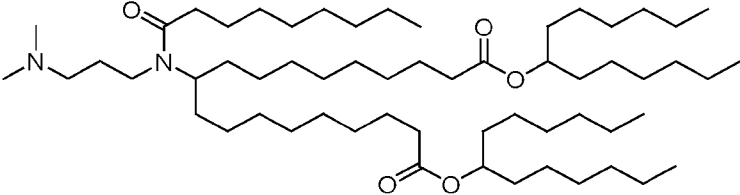
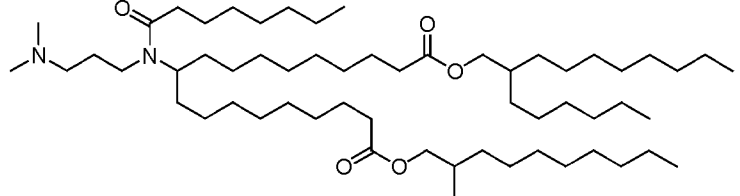
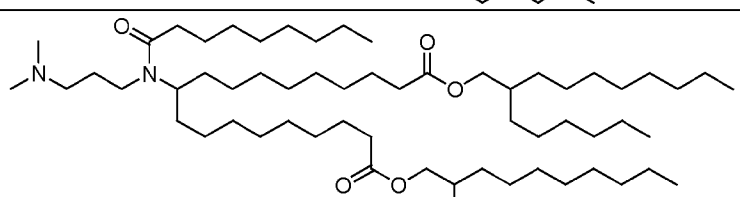
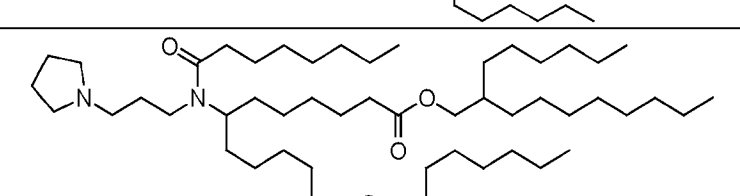
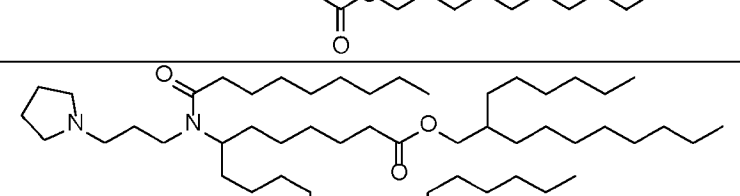
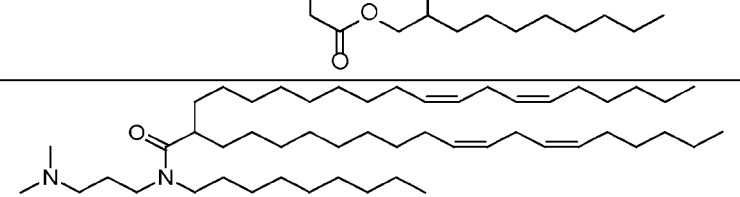
In some embodiments, G^3 is C_2 - C_4 alkylene, for example C_3 alkylene.

In some embodiments, a lipid compound has one of the structures set forth in Table 4 below.

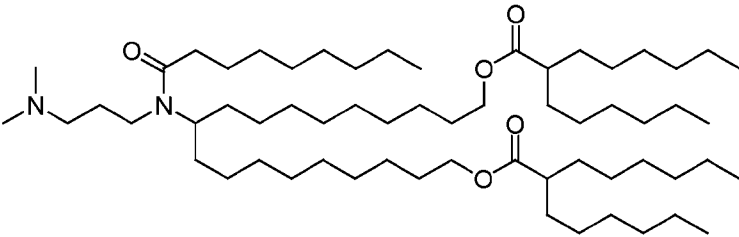
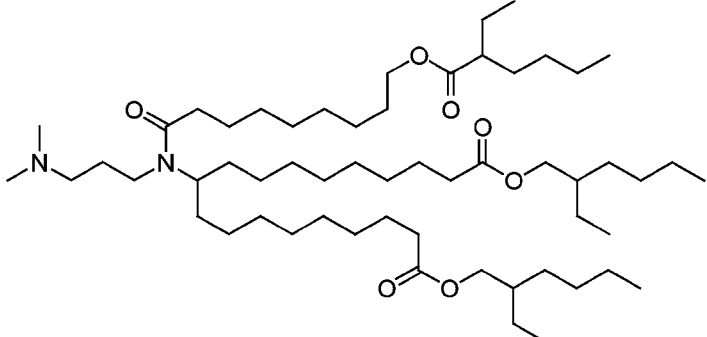
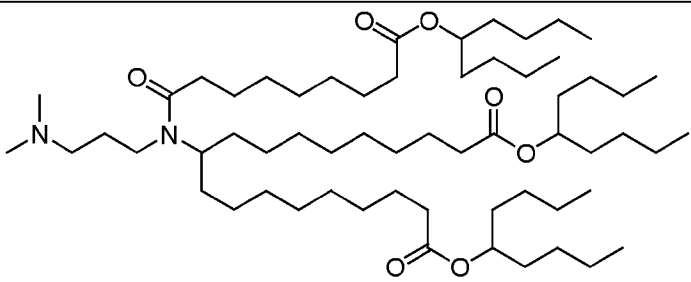
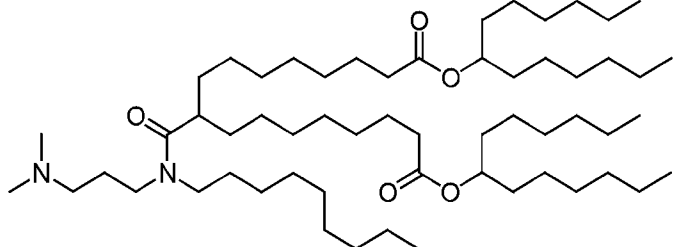
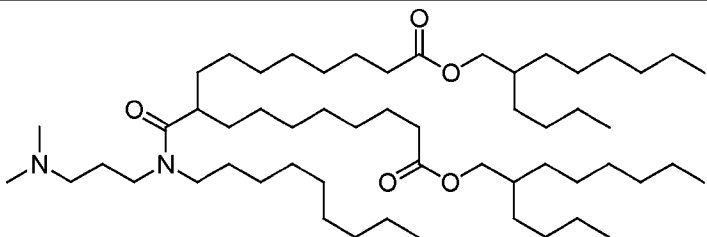
20 Table 4: Representative Lipids of Formula (II).

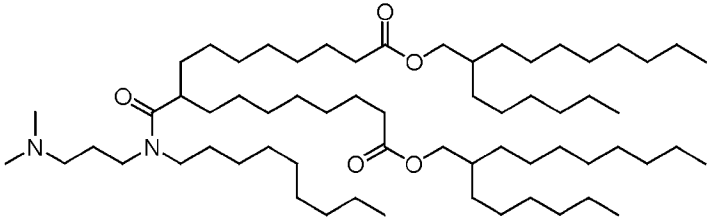
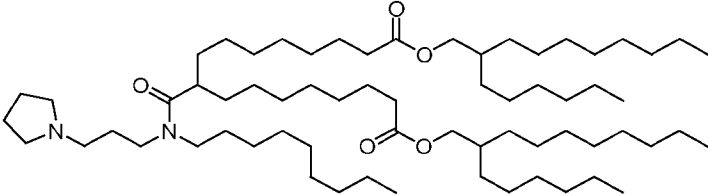
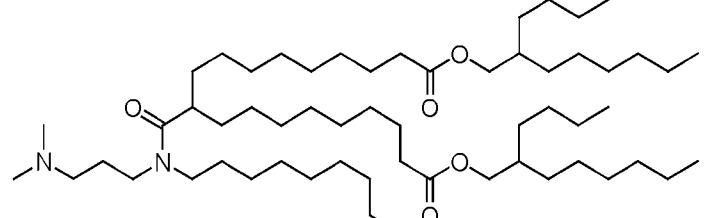
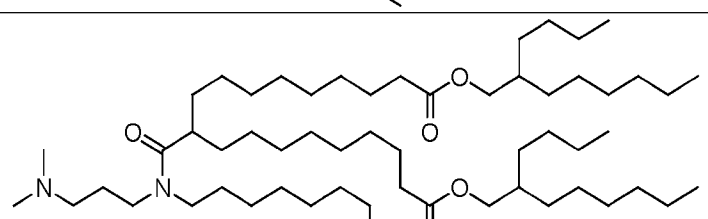
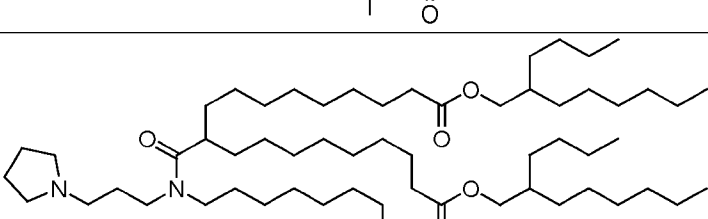
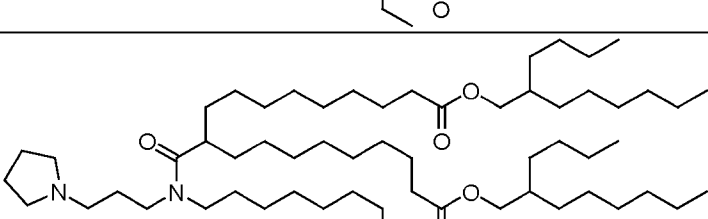
No.	Structure	Prep. Method
II-1		D

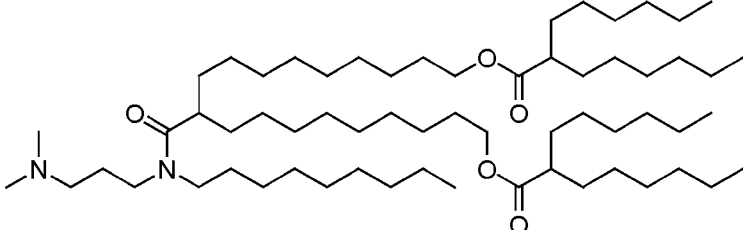
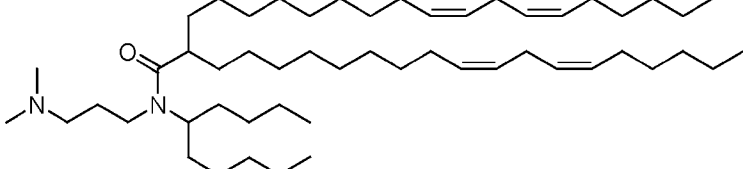
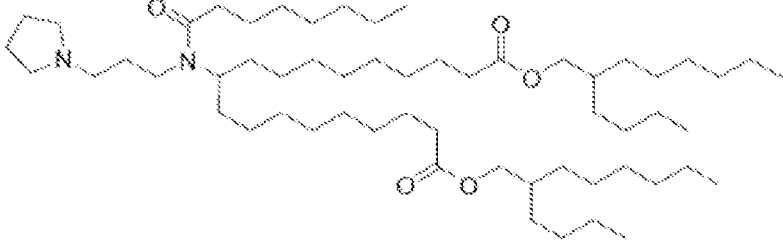
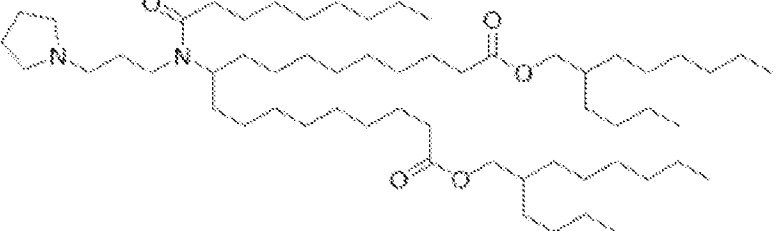
No.	Structure	Prep. Method
II-2		D
II-3		D
II-4		E
II-5		D
II-6		D
II-7		D
II-8		D
II-9		D

No.	Structure	Prep. Method
II-10		D
II-11		D
II-12		D
II-13		D
II-14		D
II-15		D
II-16		E

No.	Structure	Prep. Method
II-17		D
II-18		D
II-19		D
II-20		D
II-21		D

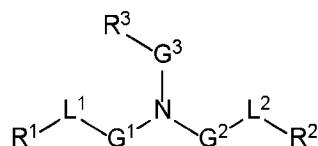
No.	Structure	Prep. Method
II-22		D
II-23		D
II-24		D
II-25		E
II-26		E

No.	Structure	Prep. Method
II-27		E
II-28		E
II-29		E
II-30		E
II-31		E
II-32		E

No.	Structure	Prep. Method
II-33		E
II-34		E
II-35		D
II-36		D

In some embodiments, LNPs comprise a lipid of Formula (II), a nucleoside-modified RNA and one or more excipient selected from neutral lipids, steroids and pegylated lipids. In some embodiments, a lipid of Formula (II) is compound II-9. In some
5 embodiments, a lipid of Formula (II) is compound II-10. In some embodiments, a lipid of Formula (II) is compound II-11. In some embodiments, a lipid of Formula (II) is compound II-12. In some embodiments, a lipid of Formula (II) is compound II-32.

In some embodiments, a cationic lipid component of the LNPs has the structure of Formula (III):



(III)

or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

- 5 one of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $\text{SC}(\text{=O})-$, $-\text{NR}^a\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{NR}^a-$, $\text{NR}^a\text{C}(\text{=O})\text{NR}^a-$, $-\text{OC}(\text{=O})\text{NR}^a-$

or

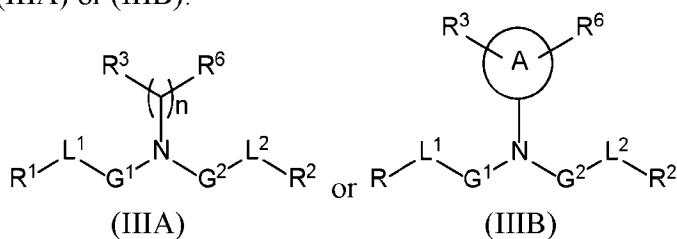
- $-\text{NR}^a\text{C}(\text{=O})\text{O}-$, and the other of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $\text{SC}(\text{=O})-$, $-\text{NR}^a\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{NR}^a-$, $\text{NR}^a\text{C}(\text{=O})\text{NR}^a-$, $-\text{OC}(\text{=O})\text{NR}^a-$ or $-\text{NR}^a\text{C}(\text{=O})\text{O}-$ or a direct bond;

G^1 and G^2 are each independently unsubstituted C_1 - C_{12} alkylene or C_1 - C_{12} alkenylene;

- G^3 is C_1 - C_{24} alkylene, C_1 - C_{24} alkenylene, C_3 - C_8 cycloalkylene, C_3 - C_8 cycloalkenylene;

- R^a is H or C_1 - C_{12} alkyl;
 R^1 and R^2 are each independently C_6 - C_{24} alkyl or C_6 - C_{24} alkenyl;
 R^3 is H, OR^5 , CN, $-\text{C}(\text{=O})\text{OR}^4$, $-\text{OC}(\text{=O})\text{R}^4$ or $-\text{NR}^5\text{C}(\text{=O})\text{R}^4$;
 R^4 is C_1 - C_{12} alkyl;
 R^5 is H or C_1 - C_6 alkyl; and
 x is 0, 1 or 2.

In some of the foregoing embodiments of Formula (III), the lipid has one of the following structures (IIIA) or (IIIB):



wherein:

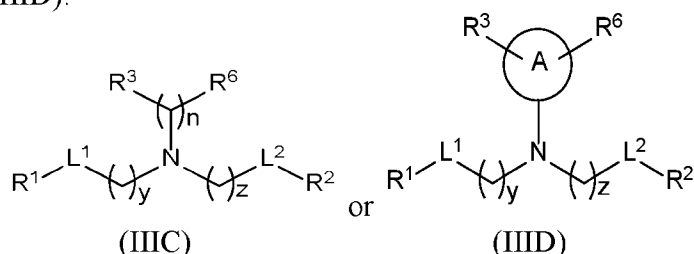
A is a 3 to 8-membered cycloalkyl or cycloalkylene ring;

R⁶ is, at each occurrence, independently H, OH or C₁-C₂₄ alkyl;

n is an integer ranging from 1 to 15.

In some embodiments of Formula (III), the lipid has structure (IIIA), and in some
5 embodiments, the lipid has structure (IIIB).

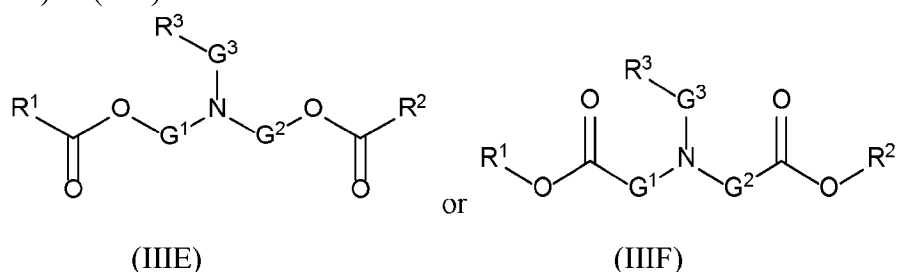
In some embodiments of Formula (III), the lipid has one of the following
structures (IIIC) or (IIID):



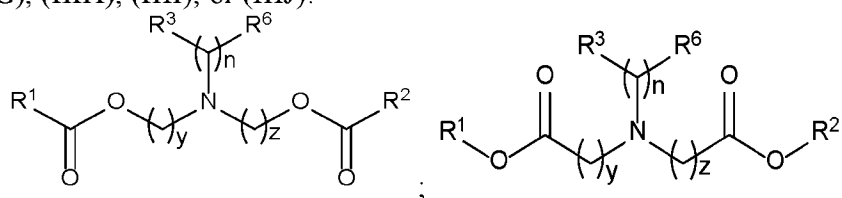
10 wherein y and z are each independently integers ranging from 1 to 12.

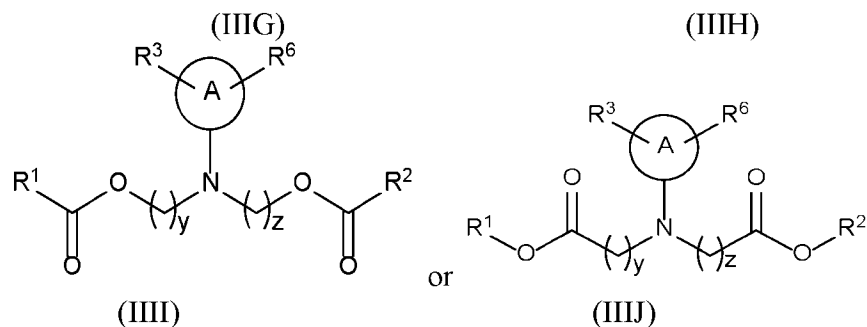
In any of the foregoing embodiments of Formula (III), one of L¹ or L²
is -O(C=O)-. For example, in some embodiments each of L¹ and L² are -O(C=O)-. In
some embodiments of any of the foregoing, L¹ and L² are each
independently -(C=O)O- or -O(C=O)-. For example, in some embodiments each of L¹
15 and L² is -(C=O)O-.

In some embodiments of Formula (III), the lipid has one of the following
structures (IIIE) or (IIIF):



20 In some embodiments of Formula (III), the lipid has one of the following
structures (IIIG), (IIIH), (IIII), or (IIIJ):





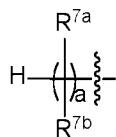
In some of Formula (III), n is an integer ranging from 2 to 12, for example from 2 to 8 or from 2 to 4. For example, in some embodiments, n is 3, 4, 5 or 6. In some embodiments, n is 3. In some embodiments, n is 4. In some embodiments, n is 5. In some embodiments, n is 6.

In some embodiments of Formula (III), y and z are each independently an integer ranging from 2 to 10. For example, in some embodiments, y and z are each independently an integer ranging from 4 to 9 or from 4 to 6.

In some embodiments of Formula (III), R^6 is H. In other of the foregoing embodiments, R^6 is C_1 - C_{24} alkyl. In some embodiments, R^6 is OH.

In some embodiments of Formula (III), G^3 is unsubstituted. In some embodiments, G^3 is substituted. In some embodiments, G^3 is linear C_1 - C_{24} alkylene or linear C_1 - C_{24} alkenylene.

In some embodiments of Formula (III), R^1 or R^2 , or both, is C_6 - C_{24} alkenyl. For example, in some embodiments, R^1 and R^2 each, independently have the following structure:



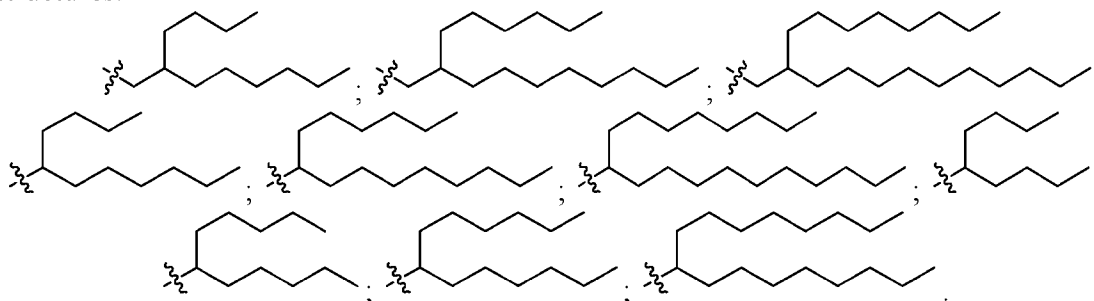
wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C_1 - C_{12} alkyl; and a is an integer from 2 to 12,

wherein R^{7a} , R^{7b} and a are each selected such that R^1 and R^2 each independently comprise from 6 to 20 carbon atoms. For example, in some embodiments a is an integer ranging from 5 to 9 or from 8 to 12.

In some embodiments of Formula (III), at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence. In other some embodiments of the foregoing, at least one occurrence of R^{7b} is C_1 - C_8 alkyl. For example, in some embodiments, C_1 - C_8 alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In some embodiments of Formula (III), R^1 or R^2 , or both, has one of the following structures:

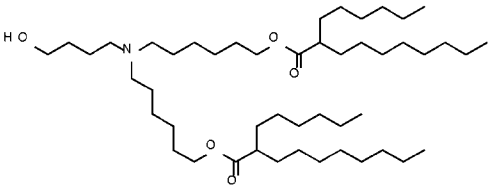
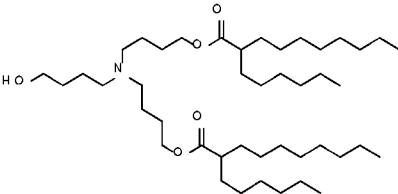
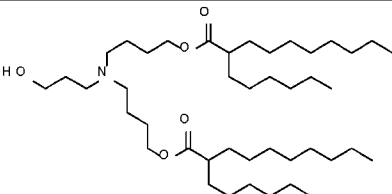
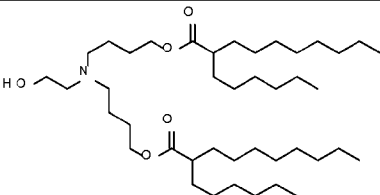
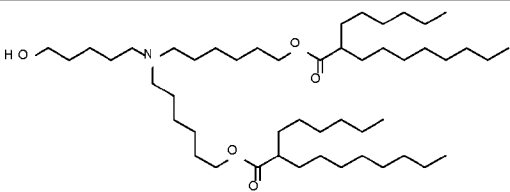
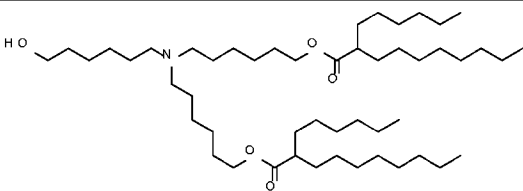
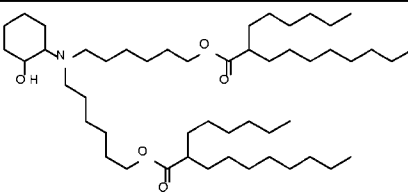


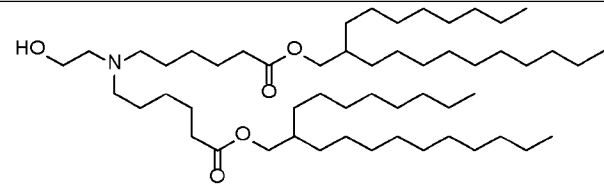
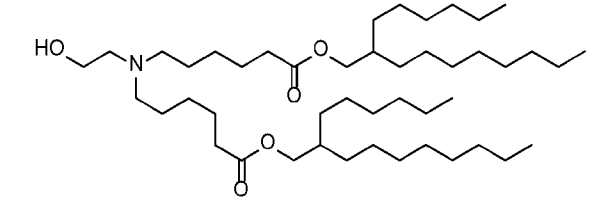
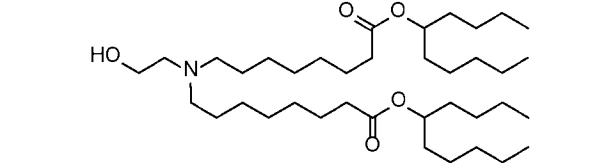
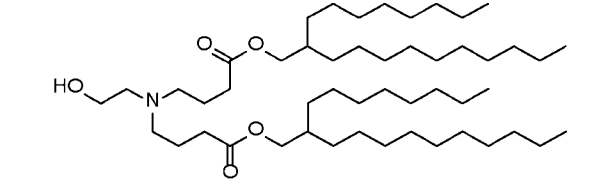
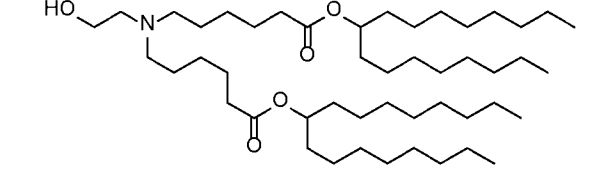
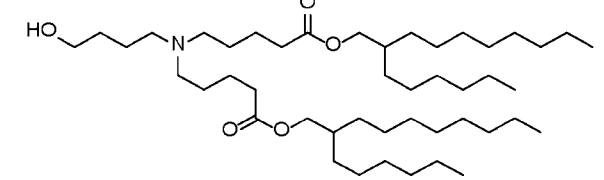
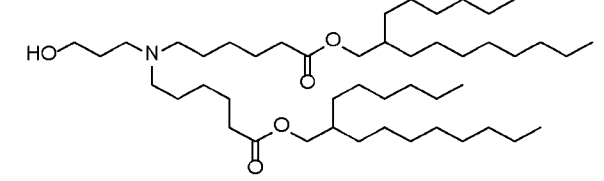
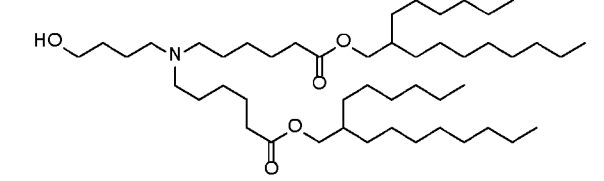
In some of the foregoing embodiments of Formula (III), R^3 is OH, CN, $-C(=O)OR^4$, $-OC(=O)R^4$ or $-NHC(=O)R^4$. In some embodiments, R^4 is methyl or ethyl.

In various some embodiments, the cationic lipid of Formula (III) has one of the structures set forth in Table 5 below.

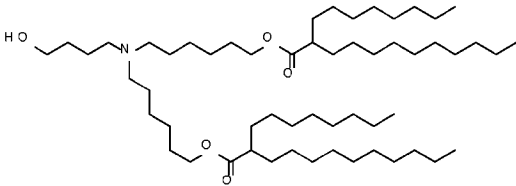
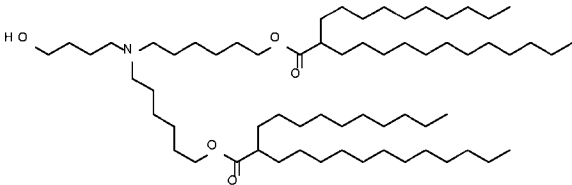
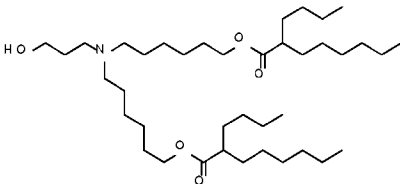
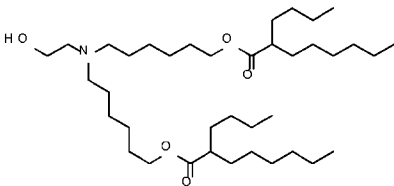
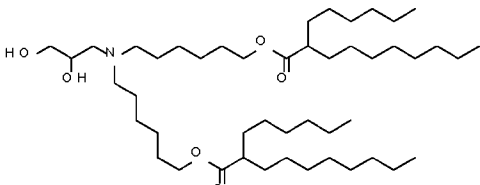
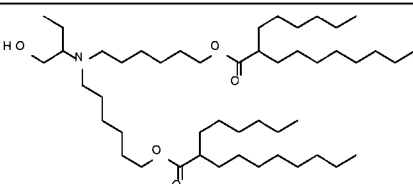
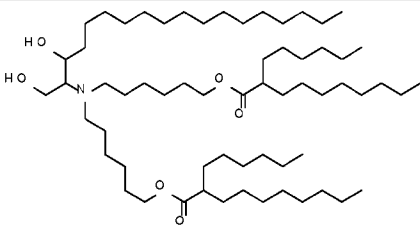
Table 5: Representative Compounds of Formula (III).

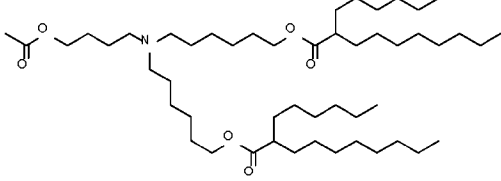
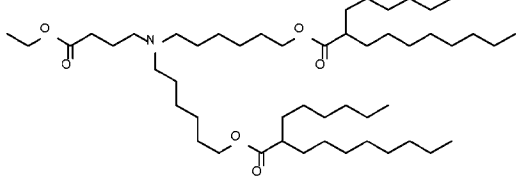
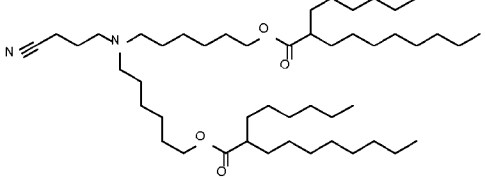
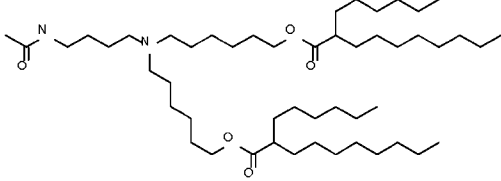
No.	Structure	Prep. Method
III-1		F
III-2		F

No.	Structure	Prep. Method
III-3		F
III-4		F
III-5		F
III-6		F
III-7		F
III-8		F
III-9		F

No.	Structure	Prep. Method
III-10		F
III-11		F
III-12		F
III-13		F
III-14		F
III-15		F
III-16		F
III-17		F

No.	Structure	Prep. Method
III-18		F
III-19		F
III-20		F
III-21		F
III-22		F
III-23		F
III-24		F
III-25		F

No.	Structure	Prep. Method
III-26		F
III-27		F
III-28		F
III-29		F
III-30		F
III-31		F
III-32		F

No.	Structure	Prep. Method
III-33		F
III-34		F
III-35		F
III-36		F

In some embodiments, LNPs comprise a lipid of Formula (III), a nucleoside-modified RNA and one or more excipient selected from neutral lipids, steroids and pegylated lipids. In some embodiments, a lipid of Formula (III) is compound III-3. In some embodiments, a lipid of Formula (III) is compound III-7.

In some embodiments, a cationic lipid is present in a LNP in an amount from about 30 to about 95 mole percent. In some embodiments, a cationic lipid is present in a LNP in an amount from about 30 to about 70 mole percent. In some embodiments, a cationic lipid is present in a LNP in an amount from about 40 to about 60 mole percent. In some embodiments, a cationic lipid is present in a LNP in an amount of about 50 mole percent. In some embodiments, a LNP comprises only cationic lipids.

In some embodiments, a LNP comprises one or more additional lipids which

stabilize the formation of particles during their formation.

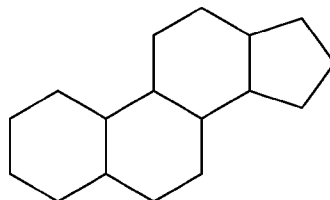
In some embodiments, suitable stabilizing lipids include neutral lipids and anionic lipids.

In some embodiments, exemplary anionic lipids include, but are not limited to, phosphatidylglycerol, cardiolipin, diacylphosphatidylserine, diacylphosphatidic acid, N-dodecanoylphosphatidylethanolamines, N-succinylphosphatidylethanolamines, N-glutarylphosphatidylethanolamines, lysylphosphatidylglycerols, palmitoyloleyolphosphatidylglycerol (POPG), and other anionic modifying groups joined to neutral lipids.

In some embodiments, exemplary neutral lipids include, for example, distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dioleoylphosphatidylglycerol (DOPG), dipalmitoylphosphatidylglycerol (DPPG), dioleoyl-phosphatidylethanolamine (DOPE), palmitoyloleyolphosphatidylcholine (POPC), palmitoyloleyolphosphatidylethanolamine (POPE) and dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (DOPE-mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidylethanolamine (DSPE), 16-O-monomethyl PE, 16-O-dimethyl PE, 18-1-trans PE, 1-stearioyl-2-oleoyl-phosphatidylethanol amine (SOPE), and 1,2-dielaoidyl-sn-glycero-3-phosphoethanolamine (transDOPE). In some embodiments, the neutral lipid is 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC).

In some embodiments, LNPs comprise a neutral lipid selected from DSPC, DPPC, DMPC, DOPC, POPC, DOPE and SM. In some embodiments, the molar ratio of the cationic lipid (e.g., lipid of Formula (I)) to the neutral lipid ranges from about 2:1 to about 8:1.

In some embodiments, LNPs further comprise a steroid or steroid analogue. In some embodiments, a “steroid” is a compound comprising the following carbon skeleton:



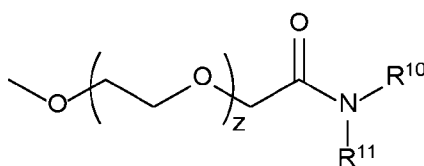
In some embodiments, a steroid or steroid analogue is cholesterol. In embodiments, the molar ratio of the cationic lipid (e.g., lipid of Formula (I)) to cholesterol ranges from about 2:1 to 1:1.

In some embodiments, a LNP comprises glycolipids (e.g., monosialoganglioside GM₁). In some embodiments, a LNP comprises a sterol, such as cholesterol.

In some embodiments, LNPs comprise a polymer conjugated lipid.

In some embodiments, a LNP comprises an additional, stabilizing -lipid which is a polyethylene glycol-lipid (pegylated lipid). Suitable polyethylene glycol-lipids include PEG-modified phosphatidylethanolamine, PEG-modified phosphatidic acid, PEG-modified ceramides (e.g., PEG-CerC14 or PEG-CerC20), PEG-modified dialkylamines, PEG-modified diacylglycerols, PEG-modified dialkylglycerols. Representative polyethylene glycol-lipids include PEG-c-DOMG, PEG-c-DMA, and PEG-s-DMG. In some embodiments, a polyethylene glycol-lipid is N-[(methoxy poly(ethylene glycol)₂₀₀₀)carbonyl]-1,2-dimyristyloxypropyl-3-amine (PEG-c-DMA). In some embodiments, a polyethylene glycol-lipid is PEG-c-DOMG). In some embodiments, LNPs comprise a pegylated diacylglycerol (PEG-DAG) such as 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-DMG), a pegylated phosphatidylethanolamine (PEG-PE), a PEG succinate diacylglycerol (PEG-S-DAG) such as 4-O-(2',3'-di(tetradecanoyloxy)propyl)-1-O-(ω-methoxy(polyethoxy)ethyl)butanedioate (PEG-S-DMG), a pegylated ceramide (PEG-cer), or a PEG dialkoxypentylcarbamate such as ω-methoxy(polyethoxy)ethyl-N-(2,3-di(tetradecanoyloxy)propyl)carbamate or 2,3-di(tetradecanoyloxy)propyl-N-(ω-methoxy(polyethoxy)ethyl)carbamate. In some embodiments, the molar ratio of the cationic lipid to the pegylated lipid ranges from about 100:1 to about 25:1.

In some embodiments, LNPs comprise a pegylated lipid having the following structure (IV):



(IV)

or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, wherein:

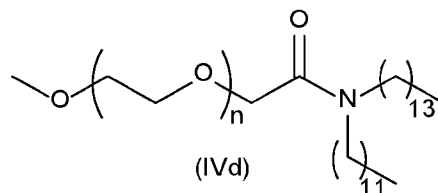
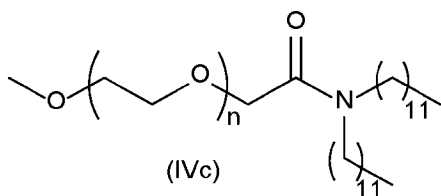
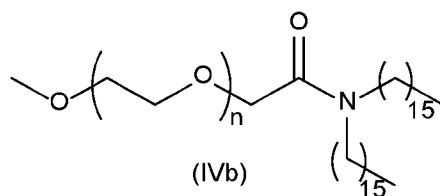
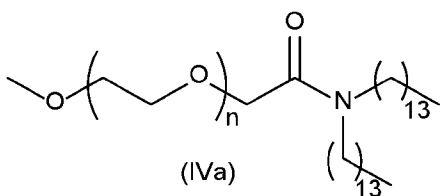
R^{10} and R^{11} are each independently a straight or branched, saturated or unsaturated alkyl chain containing from 10 to 30 carbon atoms, wherein the alkyl chain is optionally interrupted by one or more ester bonds; and

5 z has mean value ranging from 30 to 60.

In some embodiments of the pegylated lipid (IV), R^{10} and R^{11} are not both n-octadecyl when z is 42. In some some embodiments, R^{10} and R^{11} are each independently a straight or branched, saturated or unsaturated alkyl chain containing from 10 to 18 carbon atoms. In some embodiments, R^{10} and R^{11} are each independently a straight or
10 branched, saturated or unsaturated alkyl chain containing from 12 to 16 carbon atoms. In some embodiments, R^{10} and R^{11} are each independently a straight or branched, saturated or unsaturated alkyl chain containing 12 carbon atoms. In some embodiments, R^{10} and R^{11} are each independently a straight or branched, saturated or unsaturated alkyl chain containing 14 carbon atoms. In some embodiments, R^{10} and R^{11} are each independently a
15 straight or branched, saturated or unsaturated alkyl chain containing 16 carbon atoms. In some embodiments, R^{10} and R^{11} are each independently a straight or branched, saturated or unsaturated alkyl chain containing 18 carbon atoms. In still some embodiments, R^{10} is a straight or branched, saturated or unsaturated alkyl chain containing 12 carbon atoms and R^{11} is a straight or branched, saturated or unsaturated alkyl chain containing 14
20 carbon atoms.

In some embodiments, z spans a range that is selected such that the PEG portion of (II) has an average molecular weight of about 400 to about 6000 g/mol. In some embodiments, the average z is about 45.

In some embodiments, a pegylated lipid has one of the following structures:



wherein n is an integer selected such that the average molecular weight of the pegylated lipid is about 2500 g/mol.

In some embodiments, an additional lipid is present in a LNP in an amount from about 1 to about 10 mole percent. In some embodiments, the additional lipid is present in a LNP in an amount from about 1 to about 5 mole percent. In some embodiments, an additional lipid is present in a LNP in about 1 mole percent or about 1.5 mole percent.

In some embodiments, LNPs comprise a lipid of Formula (I), a nucleoside-modified RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, a lipid of Formula (I) is compound I-6. In some embodiments, the neutral lipid is DSPC. In some embodiments, a steroid is cholesterol. In still some embodiments, a pegylated lipid is compound IVa.

In some embodiments, a LNP comprises one or more targeting moieties, which are capable of targeting a LNP to a cell or cell population. For example, in some embodiments, a targeting moiety is a ligand, which directs a LNP to a receptor found on a cell surface.

In some embodiments, a LNP comprises one or more internalization domains. For example, in some embodiments, a LNP comprises one or more domains, which bind to a cell to induce internalization of the LNP. For example, in some embodiments, one or more internalization domains bind to a receptor found on a cell surface to induce receptor-mediated uptake of a LNP. In some embodiments, a LNP is capable of binding a biomolecule in vivo, where a LNP-bound biomolecule can then be recognized by a cell-

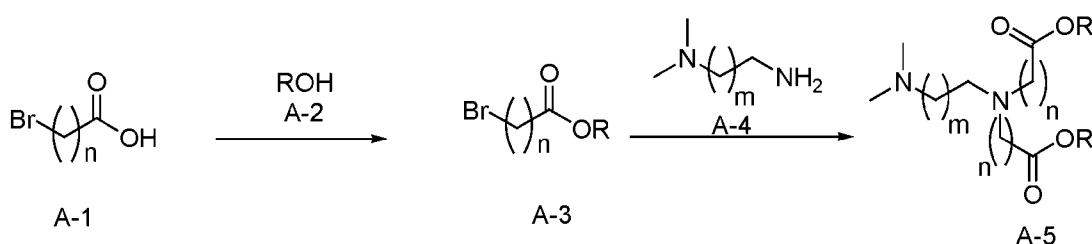
surface receptor to induce internalization. For example, in some embodiments, a LNP binds systemic ApoE, which leads to the uptake of the LNP and associated cargo.

Other exemplary LNPs and their manufacture are described in the art, for example in U.S. Patent Application Publication No. US20120276209, Semple et al., 2010, Nat Biotechnol., 28(2):172-176; Akinc et al., 2010, Mol Ther., 18(7): 1357-1364; Basha et al., 2011, Mol Ther, 19(12): 2186-2200; Leung et al., 2012, J Phys Chem C Nanomater Interfaces, 116(34): 18440-18450; Lee et al., 2012, Int J Cancer., 131(5): E781-90; Belliveau et al., 2012, Mol Ther nucleic Acids, 1: e37; Jayaraman et al., 2012, Angew Chem Int Ed Engl., 51(34): 8529-8533; Mui et al., 2013, Mol Ther Nucleic Acids. 2, e139; Maier et al., 2013, Mol Ther., 21(8): 1570-1578; and Tam et al., 2013, Nanomedicine, 9(5): 665-74, each of which are incorporated by reference in their entirety.

The following Reaction Schemes illustrate methods to make lipids of Formula (I), (II) or (III).

15

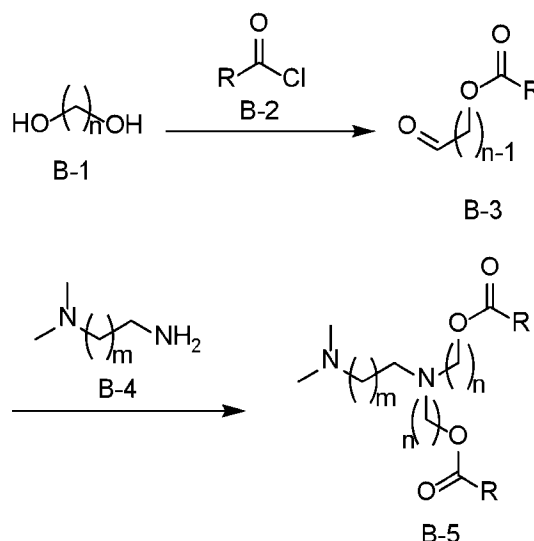
GENERAL REACTION SCHEME 1



Embodiments of the lipid of Formula (I) (e.g., compound A-5) can be prepared according to General Reaction Scheme 1 ("Method A"), wherein R is a saturated or unsaturated C₁-C₂₄ alkyl or saturated or unsaturated cycloalkyl, m is 0 or 1 and n is an integer from 1 to 24. Referring to General Reaction Scheme 1, compounds of structure A-1 can be purchased from commercial sources or prepared according to methods familiar to one of ordinary skill in the art. A mixture of A-1, A-2 and DMAP is treated with DCC to give the bromide A-3. A mixture of the bromide A-3, a base (e.g., N,N-diisopropylethylamine) and the N,N-dimethyldiamine A-4 is heated at a temperature and time sufficient to produce A-5 after any necessarily workup and or purification step.

25

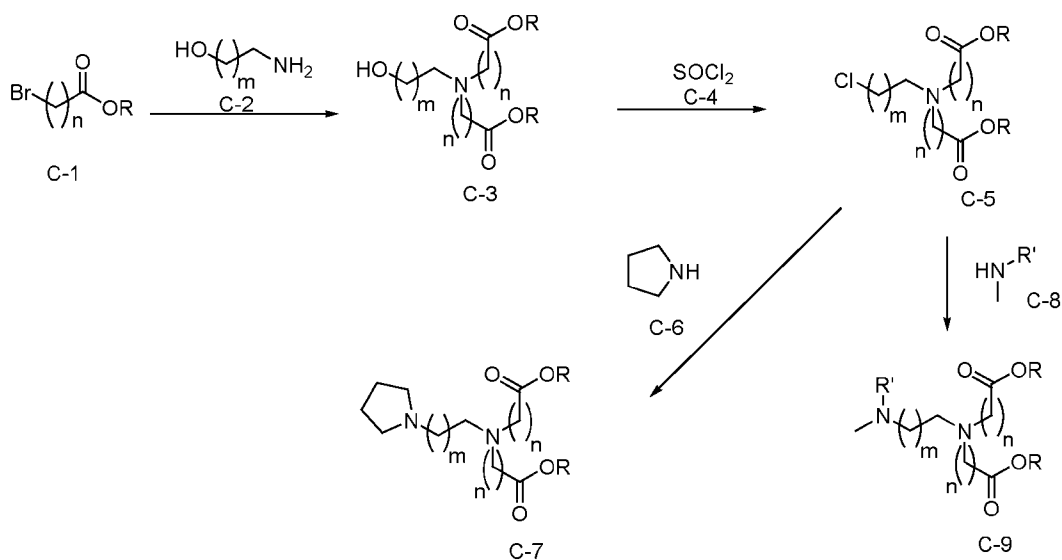
GENERAL REACTION SCHEME 2



Some embodiments of the compound of Formula (I) (e.g., compound B-5) can be prepared according to General Reaction Scheme 2 (“Method B”), wherein R is a saturated or unsaturated C₁-C₂₄ alkyl or saturated or unsaturated cycloalkyl, m is 0 or 1 and n is an integer from 1 to 24. As shown in General Reaction Scheme 2, compounds of structure B-1 can be purchased from commercial sources or prepared according to methods familiar to one of ordinary skill in the art. A solution of B-1 (1 equivalent) is treated with acid chloride B-2 (1 equivalent) and a base (e.g., triethylamine). A crude product is treated with an oxidizing agent (e.g., pyridinium chlorochromate) and intermediate product B-3 is recovered. A solution of crude B-3, an acid (e.g., acetic acid), and N,N-dimethylaminoamine B-4 is then treated with a reducing agent (e.g., sodium triacetoxyborohydride) to obtain B-5 after any necessary work up and/or purification.

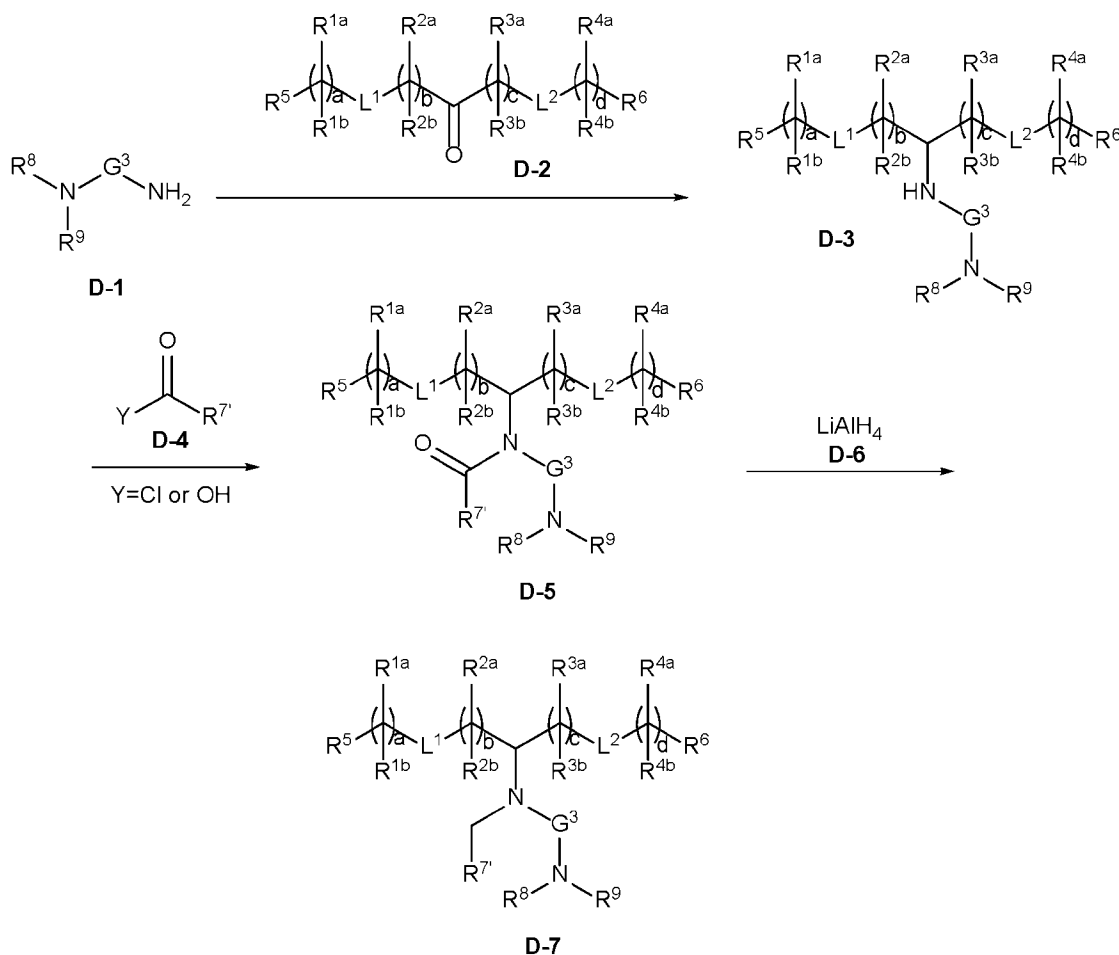
It should be noted that although starting materials A-1 and B-1 are depicted above as including only saturated methylene carbons, starting materials which include carbon-carbon double bonds may also be employed for preparation of compounds which include carbon-carbon double bonds.

GENERAL REACTION SCHEME 3



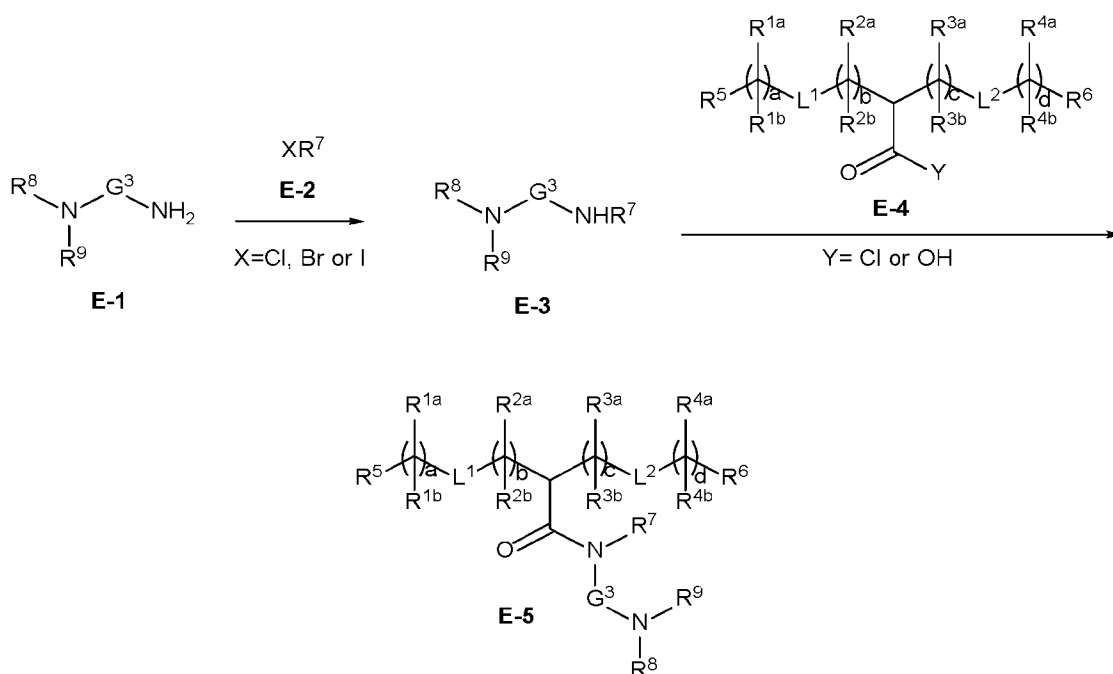
Some embodiments of the lipid of Formula (I) (e.g., compound C-7 or C-9) can be prepared according to General Reaction Scheme 3 ("Method C"), wherein R is a saturated or unsaturated C₁-C₂₄ alkyl or saturated or unsaturated cycloalkyl, m is 0 or 1 and n is an integer from 1 to 24. Referring to General Reaction Scheme 3, compounds of structure C-1 can be purchased from commercial sources or prepared according to methods familiar to one of ordinary skill in the art.

GENERAL REACTION SCHEME 4



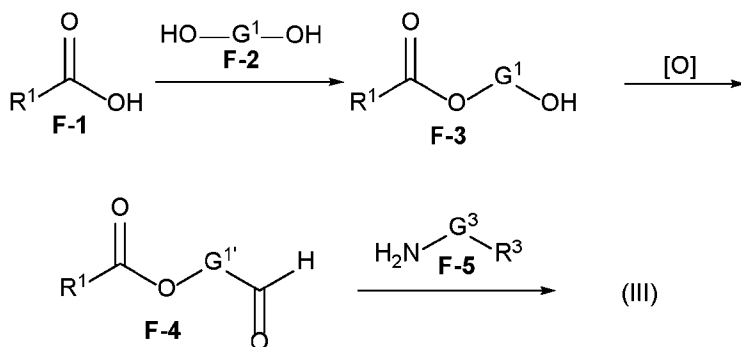
Embodiments of the compound of Formula (II) (e.g., compounds D-5 and D-7) can be prepared according to General Reaction Scheme 4 ("Method D"), wherein R^{1a}, R^{1b}, R^{2a}, R^{2b}, R^{3a}, R^{3b}, R^{4a}, R^{4b}, R⁵, R⁶, R⁸, R⁹, L¹, L², G¹, G², G³, a, b, c and d are as defined herein, and R^{7'} represents R⁷ or a C₃-C₁₉ alkyl. Referring to General Reaction Scheme 1, compounds of structure D-1 and D-2 can be purchased from commercial sources or prepared according to methods familiar to one of ordinary skill in the art. A solution of D-1 and D-2 is treated with a reducing agent (e.g., sodium triacetoxyborohydride) to obtain D-3 after any necessary work up. A solution of D-3 and a base (e.g. trimethylamine, DMAP) is treated with acyl chloride D-4 (or carboxylic acid and DCC) to obtain D-5 after any necessary work up and/or purification. D-5 can be reduced with LiAlH₄ D-6 to give D-7 after any necessary work up and/or purification.

GENERAL REACTION SCHEME 5



Embodiments of the lipid of Formula (II) (e.g., compound E-5) can be prepared according to General Reaction Scheme 5 ("Method E"), wherein R^{1a}, R^{1b}, R^{2a}, R^{2b}, R^{3a}, R^{3b}, R^{4a}, R^{4b}, R⁵, R⁶, R⁷, R⁸, R⁹, L¹, L², G³, a, b, c and d are as defined herein. Referring to General Reaction Scheme 2, compounds of structure E-1 and E-2 can be purchased from commercial sources or prepared according to methods familiar to one of ordinary skill in the art. A mixture of E-1 (in excess), E-2 and a base (e.g., potassium carbonate) is heated to obtain E-3 after any necessary work up. A solution of E-3 and a base (e.g. trimethylamine, DMAP) is treated with acyl chloride E-4 (or carboxylic acid and DCC) to obtain E-5 after any necessary work up and/or purification.

GENERAL REACTION SCHEME 6



General Reaction Scheme 6 provides an exemplary method (Method F) for

preparation of Lipids of Formula (III). G^1 , G^3 , R^1 and R^3 in General Reaction Scheme 6 are as defined herein for Formula (III), and $G1'$ refers to a one-carbon shorter homologue of $G1$. Compounds of structure F-1 are purchased or prepared according to methods known in the art. Reaction of F-1 with diol F-2 under appropriate condensation conditions (e.g., DCC) yields ester/alcohol F-3, which can then be oxidized (e.g., PCC) to aldehyde F-4. Reaction of F-4 with amine F-5 under reductive amination conditions yields a lipid of Formula (III).

It should be noted that various alternative strategies for preparation of lipids of Formula (III) are available to those of ordinary skill in the art. For example, other lipids of Formula (III) wherein L^1 and L^2 are other than ester can be prepared according to analogous methods using the appropriate starting material. Further, General Reaction Scheme 6 depicts preparation of a lipids of Formula (III), wherein G^1 and G^2 are the same; however, this is not a required aspect of the present disclosure and modifications to the above reaction scheme are possible to yield compounds wherein G^1 and G^2 are different.

It will be appreciated by those skilled in the art that in the process described herein the functional groups of intermediate compounds may need to be protected by suitable protecting groups. In some embodiments, such functional groups include hydroxy, amino, mercapto and carboxylic acid. In some embodiments, suitable protecting groups for hydroxy include trialkylsilyl or diarylalkylsilyl (for example, *t*-butyldimethylsilyl, *t*-butyldiphenylsilyl or trimethylsilyl), tetrahydropyranyl, benzyl, and the like. In some embodiments, suitable protecting groups for amino, amidino and guanidino include *t*-butoxycarbonyl, benzyloxycarbonyl, and the like. In some embodiments, suitable protecting groups for mercapto include $-C(O)-R''$ (where R'' is alkyl, aryl or arylalkyl), *p*-methoxybenzyl, trityl and the like. In some embodiments, suitable protecting groups for carboxylic acid include alkyl, aryl or arylalkyl esters. In some embodiments, protecting groups may be added or removed in accordance with standard techniques, which are known to one skilled in the art and as described herein. The use of protecting groups is described in detail in Green, T.W. and P.G.M. Wutz, *Protective Groups in Organic Synthesis* (1999), 3rd Ed., Wiley. As one of skill in the art would appreciate, the protecting group may also be a polymer resin such as a Wang resin,

Rink resin or a 2-chlorotriptyl-chloride resin.

In some embodiments, an LNP comprises a cationic lipid. In some embodiments, a cationic lipid may comprise from about 10 mol % to about 100 mol %, about 20 mol % to about 100 mol %, about 30 mol % to about 100 mol %, about 40 mol
5 % to about 100 mol %, or about 50 mol % to about 100 mol % of total lipid present in a lipid composition utilized in accordance with the present disclosure.

In some embodiments, an LNP comprises a PEG lipid. In some embodiments, a PEG lipid may comprise from about 10 mol % to about 100 mol %, about 20 mol % to about 100 mol %, about 30 mol % to about 100 mol %, about 40 mol
10 % to about 100 mol %, or about 50 mol % to about 100 mol % of total lipid present in a lipid composition utilized in accordance with the present disclosure.

In some embodiments, an LNP comprises a phospholipid. In some embodiments, a phospholipid may comprise from about 10 mol % to about 100 mol %, about 20 mol % to about 100 mol %, about 30 mol % to about 100 mol %, about 40 mol
15 % to about 100 mol %, or about 50 mol % to about 100 mol % of total lipid present in a lipid composition utilized in accordance with the present disclosure.

In some embodiments, an LNP comprises cholesterol. In some embodiments, cholesterol may comprise from about 10 mol % to about 100 mol %, about 20 mol % to about 100 mol %, about 30 mol % to about 100 mol %, about 40 mol % to
20 about 100 mol %, or about 50 mol % to about 100 mol % of total lipid present in a lipid composition utilized in accordance with the present disclosure.

Self-Assembling Nanoparticles

In one embodiment, one or more *C. difficile* antigen is incorporated into a
25 self-assembling peptide nanoparticle (SAPN). Self-assembling protein nanoparticles (SAPN) may be formed by the assembly of one or more polypeptide chains comprising at least one antigen and at least one protein oligomerization domain. Without limitation, the SAPN of the invention may self-assemble into a tetrahedron, a cube, an octahedron, a dodecahedron, or an icosahedron. The SAPN of the invention may be used as an efficient
30 means for presenting one or more *C. difficile* antigen. In one embodiment, the invention exploits ferritin, a ubiquitous iron storage protein, that self-assembles into spherical

nanoparticles and serves as a scaffold to express a *C. difficile* antigen. Therefore, in one embodiment, the invention includes compositions comprising a fusion molecule comprising a *C. difficile* antigen described herein linked to an oligomerization domain. In some embodiments, the fusion molecule comprises a polypeptide fusion molecule comprising a fusion of a *C. difficile* antigen described herein linked to an oligomerization domain. In some embodiments, the fusion molecule comprises a polynucleotide fusion molecule comprising a fusion of a nucleic acid molecule encoding a *C. difficile* antigen described herein (e.g., an mRNA) linked to a nucleic acid molecule encoding an oligomerization domain. In some embodiments, the oligomerization domain comprises ferritin, or a fragment or variant thereof.

Pharmaceutical Compositions

The present disclosure provides compositions, e.g., pharmaceutical compositions (e.g., immunogenic compositions, e.g., vaccines) comprising one or more RNA molecules described herein. In some embodiments, pharmaceutical compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In some embodiments, such preparatory methods include a step of bringing an RNA molecule into association with a carrier or one or more other accessory ingredients, and then, if necessary or desirable, shaping or packaging a composition into a desired single- or multi-dose unit.

Although description of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions which are suitable for ethical administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to subjects of all sorts.

Modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various subjects is well understood, and the ordinarily skilled veterinary pharmacologist can design and perform such modification with merely ordinary, if any, experimentation. In some embodiments, a pharmaceutical composition is administered to a human or other primate, or another mammal including commercially relevant mammals such as non-human primates, cattle, pigs, horses, sheep, cats, and dogs.

In some embodiments, pharmaceutical compositions described herein may be prepared, packaged, or sold in formulations suitable for ophthalmic, oral, rectal, vaginal, parenteral, topical, pulmonary, intranasal, buccal, intravenous, intracerebroventricular, intradermal, intramuscular, or another route of administration. In some embodiments, a pharmaceutical composition is formulated to include projected nanoparticles, liposomal preparations, resealed erythrocytes containing an active ingredient, and immunogenic-based formulations.

In some embodiments, a pharmaceutical composition may be prepared, packaged, or sold in bulk, as a single unit dose, or as a plurality of single unit doses. As used herein, a “unit dose” is a discrete amount of a pharmaceutical composition comprising a predetermined amount of an active ingredient (e.g., an RNA molecule). In some embodiments, an amount of an active ingredient is generally equal to a dosage of an active ingredient, which would be administered to a subject, or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

In some embodiments, relative amounts of an active ingredient, the pharmaceutically acceptable carrier, and any additional ingredients in a pharmaceutical composition of the present disclosure will vary, depending upon identity, size, and condition of a subject treated and further depending upon a route by which a composition is to be administered. In some embodiments, a composition may comprise between 0.1% and 100% (w/w) active ingredient.

In some embodiments, in addition to an active ingredient, a pharmaceutical composition of the present disclosure may further comprise one or more additional pharmaceutically active agents.

In addition to the active ingredient, a pharmaceutical composition of the invention may further comprise one or more additional adjuvants. Exemplary adjuvants include, but are not limited to, aluminum-based adjuvant and monophosphoryl lipid A.

In still further embodiments, a pharmaceutical composition of the invention may further comprise a biological response modifier, a chemokine, a cytokine, a γ -chain receptor cytokine, A proliferation-inducing ligand (APRIL), IL-2, IL-7, IL-12, IL-15, and IL-21, or an immune checkpoint agonist or antagonist, or a nucleic acid molecule encoding the same.

In some embodiments, a pharmaceutical composition is or comprises a controlled- or sustained -release formulation. Controlled- or sustained-release formulations of a pharmaceutical composition of the present disclosure may be made using conventional technology.

5 As described herein, “parenteral administration” of a pharmaceutical composition includes any route of administration characterized by physical breaching of a tissue of a subject and administration of a pharmaceutical composition through a breach in a tissue. In some embodiments, parenteral administration thus includes, but is not limited to, administration of a pharmaceutical composition by injection of a composition,
10 by application of a composition through a surgical incision, by application of a composition through a tissue-penetrating non-surgical wound, and the like. In some embodiments,, parenteral administration comprises intraocular, intravitreal, subcutaneous, intraperitoneal, intramuscular, intradermal, intrasternal injection, intratumoral, intravenous, intracerebroventricular and kidney dialytic infusion techniques.

15 In some embodiments, formulations of a pharmaceutical composition suitable for parenteral administration comprise an active ingredient combined with a pharmaceutically acceptable carrier, such as sterile water or sterile isotonic saline. In some embodiments, a formulation may be prepared, packaged, or sold in a form suitable for bolus administration or for continuous administration. Injectable formulations may be
20 prepared, packaged, or sold in unit dosage form, such as in ampules or in multi-dose containers containing a preservative. In some embodiments, a pharmaceutical composition comprises suspensions, solutions, emulsions in oily or aqueous vehicles, pastes, and implantable sustained-release or biodegradable formulations. In some embodiments, a pharmaceutical composition further comprises one or more additional
25 ingredients including, but not limited to, suspending, stabilizing, or dispersing agents. In some embodiments, a pharmaceutical composition formulated for parenteral administration, comprises an active ingredient provided in dry (i.e. powder or granular) form for reconstitution with a suitable vehicle (e.g. sterile pyrogen-free water) prior to parenteral administration of a reconstituted composition.

30 In some embodiments, a pharmaceutical composition may be prepared, packaged, or sold in a form of a sterile injectable aqueous or oily suspension or solution.

In some embodiments, a suspension or solution may be formulated according to the known art, and may comprise, in addition to an active ingredient, additional ingredients such as the dispersing agents, wetting agents, or suspending agents described herein. In some embodiments, a sterile injectable formulation may be prepared using a non-toxic
5 parenterally-acceptable diluent or solvent, such as water or 1,3-butane diol, for example. In some embodiments, acceptable diluents and solvents comprise Ringer's solution, isotonic sodium chloride solution, and fixed oils such as synthetic mono- or di-glycerides. In some embodiments, a parentally administrable formulation of a pharmaceutical composition comprises an active ingredient in microcrystalline form, in a liposomal
10 preparation, or as a component of a biodegradable polymer systems. In some embodiments, compositions for sustained release or implantation may comprise pharmaceutically acceptable polymeric or hydrophobic materials such as an emulsion, an ion exchange resin, a sparingly soluble polymer, or a sparingly soluble salt.

In some embodiments, a pharmaceutical composition may be prepared,
15 packaged, or sold in a formulation suitable for pulmonary administration via a buccal cavity. In some embodiments, a pharmaceutical composition suitable for pulmonary administration comprises dry particles which comprise the active ingredient and which have a diameter in the range from about 0.5 to about 7 nanometers. In some
20 embodiments, a formulation may comprise dry particles which comprise an active ingredient and which have a diameter in the range from about 1 to about 6 nanometers. In some embodiments, compositions are in a form of dry powders for administration using a device comprising a dry powder reservoir to which a stream of propellant may be directed to disperse a powder or using a self-propelling solvent/powder-dispensing
25 container such as a device comprising an active ingredient dissolved or suspended in a low-boiling propellant in a sealed container. In some embodiments, such powders comprise particles wherein at least 98% of the particles by weight have a diameter greater than 0.5 nanometers and at least 95% of the particles by number have a diameter less than 7 nanometers. In some embodiments, at least 95% of particles by weight have a diameter greater than 1 nanometer and at least 90% of particles by number have a diameter less
30 than 6 nanometers. In some embodiments, dry powder compositions include a solid fine powder diluent such as sugar and are conveniently provided in a unit dose form.

In some embodiments, low boiling propellants generally include liquid propellants having a boiling point of below 65°F at atmospheric pressure. In some embodiments, a propellant may constitute 50 to 99.9% (w/w) of the composition, and an active ingredient may constitute 0.1 to 20% (w/w) of the composition. In some
5 embodiments, a propellant may further comprise additional ingredients such as a liquid non-ionic or solid anionic surfactant or a solid diluent (in some instances having a particle size of the same order as particles comprising the active ingredient).

In some embodiments, formulations of a pharmaceutical composition suitable for parenteral administration comprise an active ingredient combined with a
10 pharmaceutically acceptable carrier, such as sterile water or sterile isotonic saline. In some embodiments, such formulations may be prepared, packaged, or sold in a form suitable for bolus administration or for continuous administration. In some embodiments, injectable formulations may be prepared, packaged, or sold in unit dosage form, such as in ampules or in multi-dose containers containing a preservative. In some embodiments,
15 formulations for parenteral administration include, but are not limited to, suspensions, solutions, emulsions in oily or aqueous vehicles, pastes, and implantable sustained-release or biodegradable formulations. In some embodiments, such formulations may further comprise one or more additional ingredients including, but not limited to, suspending, stabilizing, or dispersing agents. In some embodiments, a formulation for
20 parenteral administration, comprise an active ingredient in dry (i.e., powder or granular) form for reconstitution with a suitable vehicle (e.g., sterile pyrogen-free water) prior to parenteral administration of the reconstituted composition.

In some embodiments, a pharmaceutical compositions may be prepared, packaged, or sold in a form of a sterile injectable aqueous or oily suspension or solution.
25 In some embodiments, a suspension or solution may be formulated according to the known art, and may comprise, in addition to an active ingredient, additional ingredients such as the dispersing agents, wetting agents, or suspending agents described herein. In some embodiments, a sterile injectable formulation may be prepared using a non-toxic parenterally-acceptable diluent or solvent, such as water or 1,3-butane diol, for example.
30 In some embodiments, other acceptable diluents and solvents include, but are not limited to, Ringer's solution, isotonic sodium chloride solution, and fixed oils such as synthetic

mono- or di-glycerides. In some embodiments, other parentally administrable formulations that are useful include those that comprise an active ingredient in microcrystalline form, in a liposomal preparation, or as a component of a biodegradable polymer system. In some embodiments, compositions for sustained release or
5 implantation may comprise pharmaceutically acceptable polymeric or hydrophobic materials such as an emulsion, an ion exchange resin, a sparingly soluble polymer, or a sparingly soluble salt.

Methods of Treatment or Prevention

10 Among other things, in some embodiments, technologies of the present disclosure are used to treat a disease or disorder. In some embodiments, the present disclosure provides methods of treating *C. difficile* infection or treating or preventing a disease or disorder associated with *C. difficile* infection using compositions of the present disclosure. In some embodiments, the present disclosure provides methods of treating *C.*
15 *difficile* infection. In some embodiments, the present disclosure provides methods for treating or preventing a disease or disorder associated with *C. difficile* infection. In some embodiment, compositions described herein can induce an immune response in a subject.

 In some embodiments, provided methods include administering a composition comprising at least one lipid nanoparticle (LNP) comprising at least one
20 nucleoside-modified RNA molecule encoding at least one antigen, wherein the at least one antigen comprises toxin A (TcdA) or a fragment thereof comprising at least a receptor binding domain (RBD), toxin B (TcdB) or a fragment thereof comprising at least a RBD, Pro-Pro endopeptidase 1 (PPEP-1/Zmp1), exosporium morphogenic protein CdeM, cell surface protein cwp84 (CWP84), CWP66, ZupT, Bile acid germinant receptor
25 pseudoprotease CspC (CspC), or any combination thereof to a subject in need thereof to treat a disease or disorder or to prevent or reduce at least one symptom of a disease associated with *C. difficile* infection.

 Exemplary diseases that can be treated or prevented according to methods of the present disclosure include, but are not limited to, colitis and diarrhea.

30 In some embodiments, the present disclosure provides methods of inducing an adaptive immune response against *C. difficile* in a subject comprising

administering at least one RNA molecule (e.g., a nucleoside modified RNA molecule) encoding a *C. difficile* antigen or a composition (e.g., an LNP) comprising at least one RNA molecule (e.g., a nucleoside modified RNA molecule) encoding a *C. difficile* antigen.

5 In some embodiments, provided methods provide immunity in a subject to multiple strains of *C. difficile*, *C. difficile* infection, or to a disease or disorder associated with *C. difficile* including, but not limited to, colitis and diarrhea. The present disclosure thus provides a method of treating or preventing the infection, disease, or disorder associated with *C. difficile*.

10 In some embodiments, a composition is administered to a human subject having an infection, disease, or disorder associated with *C. difficile*. In some embodiments, a composition is administered to a subject at risk for developing an infection, disease, or disorder associated with *C. difficile*. In some embodiments, a composition may be administered to a subject who is at risk for being in contact with *C.*
15 *difficile*. In some embodiments, a composition is administered to a subject who is on antibiotics that increase risk of contracting *C. difficile*. Populations of interest for administration of a composition described herein (e.g., an immunogenic composition, e.g., a vaccine) include, but are not limited to, older adults (e.g., over 65 years of age), those with a recent stay at a hospital or nursing home, those with a weakened immune
20 system, such as people with HIV/AIDS, cancer, organ transplant patients taking immunosuppressive drugs, individuals taking high-risk antibiotics, hospitalized individuals, individuals with gastrointestinal or autoimmune disorders that heighten risk of *C. difficile* infection (e.g. inflammatory bowel disease), those with repeated episodes of recurrent and/or refractory *C. difficile* infection, children, or gravid subjects, and other
25 populations at heightened risk for developing *C. difficile* infection.

 In some embodiments, a composition is administered to a non-human animal subject having an infection, disease, or disorder associated with *C. difficile*. In some embodiments, a composition is administered to a non-human animal subject at risk for developing an infection, disease, or disorder associated with *C. difficile*. In some
30 embodiments, a composition may be administered to a subject who is at risk for being in contact with *C. difficile*. In some embodiments, a composition is administered to a non-

human animal subject who is on antibiotics that increase the risk of contracting *C. difficile*. Non-human animal subjects of interest for administration of an exemplary RNA-LNP a composition described herein (e.g., an immunogenic composition, e.g., a vaccine) include, but are not limited to, dogs, cats, horses, pigs, calves, hamsters, guinea pigs, rats, and rabbits.

5 In some embodiments, provided methods comprise administering a composition comprising one or more nucleoside-modified nucleic acid molecules encoding one or more *C. difficile* antigens selected from TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, and any combination thereof. In some embodiments, provided methods comprise administering a composition comprising nucleoside-modified nucleic acid molecules encoding TcdA, TcdB, PPEP-1, and CdeM, or fragments or variants thereof. In some embodiments, provided methods comprise administering a composition comprising nucleoside-modified nucleic acid molecules encoding TcdA and TcdB or fragments or variants thereof. In some
10
15
embodiments, provided methods comprise administering a composition comprising nucleoside-modified nucleic acid molecules encoding TcdA, TcdB, and PPEP-1 or fragments or variants thereof.

In some embodiments, provided methods comprise administering one or more compositions, each composition comprising one or more nucleoside-modified nucleic acid molecules encoding one or more *C. difficile* antigens comprising TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, or fragment thereof, or any combination thereof. In some embodiments, provided methods comprise administering a first composition comprising one or more nucleoside-modified nucleic acid molecules encoding one or more *C. difficile* antigens that comprise TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, or fragment thereof, or any combination thereof, and administering a second composition comprising one or more nucleoside-modified nucleic acid molecules encoding one or more *C. difficile* antigens that comprise TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, or fragment thereof, or any combination thereof. In some embodiments, provided methods comprise administering a single
20
25
30
composition. In some embodiments, provided methods comprise administering a plurality of compositions, each composition comprising one or more nucleoside-modified nucleic

acid molecules encoding one or more *C. difficile* antigens described herein. In some embodiments, provided methods comprise a staggered administration of a plurality of compositions.

In some embodiments, provided methods comprise administering a first LNP comprising a nucleoside-modified RNA molecule encoding TcdA, or a fragment thereof, and a second LNP comprising a nucleoside-modified RNA molecule encoding TcdB, or a fragment thereof. In some embodiments, provided methods comprise administering a first LNP comprising a nucleoside-modified RNA molecule encoding TcdA, or a fragment thereof, a second LNP comprising a nucleoside-modified RNA molecule encoding TcdB, or a fragment thereof, and a third LNP comprising a nucleoside-modified RNA molecule encoding PPEP-1, or a fragment thereof. In some embodiments, provided methods comprise administering a single composition comprising a plurality of LNPs. In some embodiments, provides methods comprise administering a plurality of compositions comprising a plurality of LNPs. In some embodiments, the method comprises a staggered administration of a plurality of compositions comprising a plurality of LNPs.

In some embodiments, provided methods comprise administering a first LNP comprising a nucleoside-modified RNA comprising a nucleotide sequence corresponding to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, and a second LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof. In some embodiments, provided methods comprise administering a first LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, a second LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:45, SEQ ID NO:46, SEQ

ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof, and a third LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a fragment or variant thereof.

In some embodiments, provided methods comprise administering a first LNP comprising a nucleoside-modified RNA comprising a nucleotide sequence corresponding to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, and a second LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof. In some embodiments, provided methods comprise administering a first LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, a second LNP comprising a nucleoside-modified RNA comprising a nucleotide sequence corresponding to SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof, and a third LNP comprising a nucleoside-modified RNA molecule comprising a nucleotide sequence corresponding to SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a fragment or variant thereof.

In some embodiments, provided methods comprise administering a single composition comprising a plurality of LNPs. In some embodiments, provided methods comprise administering a plurality of compositions. In some embodiments, provided methods comprises a staggered administration of a plurality of compositions.

In some embodiments, provided methods comprise sustained expression of a *C. difficile* antigen described herein for at least several days following administration. In some embodiments, provided methods comprise sustained expression of a *C. difficile* antigen described herein for at least 2 weeks following administration. In some embodiments, provided methods comprise sustained expression of a *C. difficile* described herein for at least 1 month following administration. In some embodiments, provided methods comprise transient expression of a *C. difficile* antigen described herein.

In some embodiments, provided methods comprises administering nucleoside-modified RNA, which provides stable expression of a *C. difficile* antigen described herein in a subject. In some embodiments, administration of nucleoside-modified RNA results in little to no innate immune response, while inducing an effective adaptive immune response in a subject.

In some embodiments, a method described herein provides sustained protection against *C. difficile* in a subject. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for more than 2 weeks. In some embodiments, the method provides sustained protection against *C. difficile* in a subject for 1 month or more. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for 2 months or more. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for 3 months or more. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for 4 months or more. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for 5 months or more. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for 6 months or more. In some embodiments, a method provides sustained protection against *C. difficile* in a subject for 1 year or more.

In some embodiments, a single immunization of a composition induces a sustained protection against *C. difficile* in a subject for 1 month or more, 2 months or more, 3 months or more, 4 months or more, 5 months or more, 6 months or more, or 1 year or more.

Administration of a composition described herein can be achieved in a number of different ways, using methods known in the art. In some embodiments, a

method comprises systemic administration of a subject, including for example enteral or parenteral administration. In some embodiments, a method comprises intradermal delivery of a composition. In some embodiments, a method comprises intravenous delivery of a composition. In some embodiments, a method comprises intramuscular delivery of a composition. In some embodiments, a method comprises subcutaneous delivery of a composition. In some embodiment, a method comprises inhalation of a composition. In some embodiments, a method comprises intranasal delivery of a composition.

10 In some embodiments, a composition is administered to a subject in conjunction with another agent.

Among other things, in some embodiments, technologies of the present disclosure may be administered to a subject according to a particular dosing regimen.

In some embodiments, provided compositions (e.g., pharmaceutical compositions) are administered to deliver a dose of from 1 ng/kg/day to 100 mg/kg/day.
15 In some embodiments, a composition is administered at a dose, which results in a concentration of from 10 nM to 10 μ M in a subject.

In some embodiments, provided compositions are administered in RNA doses from 0.01 μ g to about 50 mg per kilogram of body weight of a subject. In some embodiments, provided compositions are administered in RNA doses from about 0.1 μ g to about 10 mg per kilogram of body weight of a subject. In some embodiments, the dosage will vary from about 1 μ g to about 1 mg per kilogram of body weight of the mammal. Among other things, a dosage administered may depend upon any number of factors, including but not limited to, a type of subject (e.g., mammal) and type of disease state being treated, age of a subject and route of administration.

25 In some embodiments, a composition may be administered as frequently as several times daily, or it may be administered less frequently, such as once a day, once a week, once every two weeks, once a month, or even less frequently, such as once every several months, several years, or even less frequently, such as every 10-20 years, 15-30 years, or even less frequently, such as every 50-100 years. A suitable frequency of a dose
30 will be readily apparent to a skilled artisan and will depend upon any number of factors, such as, but not limited to, type and severity of a disease being treated, type and age of a

mammal, etc.

In some embodiments, administration of a composition described herein (e.g., an immunogenic composition, pharmaceutical composition, or vaccine) comprises a single administration. In some embodiments, administration of a composition is
5 boosted by multiple administrations.

In some embodiments, provided compositions may be administered prophylactically (i.e., to prevent a disease or disorder) or therapeutically (i.e., to treat a disease or disorder) to subjects suffering from, or at risk of (or susceptible to) developing a disease or disorder. In some embodiments, such subjects may be identified using
10 standard clinical methods. In some embodiments, prophylactic administration occurs prior to manifestation of overt clinical symptoms of disease, such that a disease or disorder is prevented or alternatively delayed in its progression. In some embodiments, the term “prevent” encompasses any activity, which reduces burden of mortality or morbidity from disease. In some embodiments, prevention can occur at primary,
15 secondary and tertiary prevention levels. In some embodiments, while primary prevention avoids the development of a disease, secondary and tertiary levels of prevention encompass activities aimed at preventing progression of a disease and the emergence of symptoms as well as reducing negative impact of an already established disease by restoring function and reducing disease-related complications.

20

EXPERIMENTAL EXAMPLES

The present disclosure is further described in detail by reference to the following experimental examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified. Thus, the
25 present disclosure should in no way be construed as being limited to the following examples, but rather should be construed to encompass any and all variations which become evident as a result of the teaching provided herein.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and
30 utilize the present disclosure and practice the claimed methods. The following working examples therefore are not to be construed as limiting in any way the remainder of the

disclosure.

Example 1: Production of an exemplary multi-valent modified RNA-LNP vaccine to protect against *Clostridioides difficile* infection.

5 The present Example describes identification, selection and/or characterization of compositions and related methods useful for inducing an immune response against *C. difficile* infection (CDI) in a subject as described herein. Moreover, the present Example describes development of an RNA-LNP vaccine useful for inducing an immune response against *C. difficile* infection (CDI) in a subject.

10 Presented herein is development of a nucleoside-modified RNA-LNP vaccine against an enteric pathogen using a multivalent strategy targeting secreted virulence factors. This work is also the first examination of an anti-*C. difficile* RNA-LNP vaccine on the composition of the intestinal microbiota. This approach was designed to prevent *C. difficile*-associated disease by targeting TcdA and TcdB, as well as the
15 metalloprotease and virulence factor, Pro-Pro endopeptidase 1 (PPEP-1/Zmp1). A multivalent strategy against *C. difficile* toxins and PPEP-1 induces a robust cellular and humoral immune response that effectively protects mice from lethal toxin and bacterial challenges and hinders *C. difficile* colonization and virulence in a mouse model. In addition, despite vaccination eliciting a humoral response and transcytosis of antibodies
20 into the gut lumen, this is not be associated with changes in the microbiota.

Study design

 LNP-encapsulated RNA molecules encoding TcdA, TcdB, and PPEP-1 were generated. Design of RNA molecules used in the present Example comprised
25 optimizing nucleotide sequences for efficient translation and stability.

 A series of murine studies were conducted to assess efficacy of an exemplary RNA-LNP vaccination strategy in inducing a protective immune response against *C. difficile* infection. First, cellular and humoral immune responses at systemic and mucosal sites following vaccination using flow cytometry and ELISA were
30 evaluated. Two well-established *in vivo* murine models of toxin challenge and *C. difficile* infection were employed. Mice were randomly assigned to different experimental groups,

with a minimum of five and a maximum of ten mice per group. This random assignment ensured an unbiased distribution of animals across groups. All experiments were repeated independently two to three times to ensure reproducibility. Investigators were not blinded to treatment groups.

5 To assess safety of exemplary RNA-LNP vaccines on composition of intestinal microbiota, 16S rRNA gene sequencing was performed on fecal samples from each experimental group. Microbiota samples were collected from two independent experiments and a minimum of two and maximum of 5 cages per group to minimize cage and litter effects. This analysis allowed for detection of any alterations in gut microbiota
10 associated with vaccination strategies used in the present Example.

Ethical considerations were strictly followed throughout the study. The animal experiments were approved by the Animal Care and Use Committees of the Children's Hospital of Philadelphia (IAC 18-001316) and the University of Pennsylvania (IACUC-803941).

15

RNA design and production

Amino acid sequences of a TcdA and TcdB receptor binding domain (RBD) containing a combined repetitive oligopeptides (CROP) motifs were obtained from GenBank accession numbers P16154.2 and P18177.3 respectively. Putative N-
20 glycosylation sites were disrupted by substituting an asparagine residue at a predicted N-glycosylation site with a glutamine (N to Q) to prevent posttranslational epitopes masking in eucaryotic cells. Sequences underwent codon optimization and GC enrichment using a proprietary algorithm to improve expression and reduce potential immunogenicity of *in vitro* transcribed RNA. Codon optimized sequences were gene synthesized by Genscript
25 with an optimized and modified IL-2 secretion signal, cloned into a proprietary *in vitro* transcription template containing an optimized T7 promoter, 3'UTR, 5'UTR and a 100-adenine tail. TcdA, TcdB and PPEP-1 nucleoside modified RNA sequences were prepared using a MegaScript transcription kit (ThermoFisher Scientific), co-transcriptionally capped using a CleanCap™ system (TriLink Biotechnologies) and
30 purified using a modified cellulose base chromatography method (M. Baiersdörfer, et al., *Mol Ther Nucleic Acids*. 2019;(15):26-35), precipitated, eluted in nuclease free water,

and quantified using the NanoDrop One system. Length and integrity were determined using an Agilent BioAnalyzer 2100 system. Endotoxin content was measured using a GenScript Toxisensor chromogenic assay, and values were below detection levels. RNA was frozen at -20°C until formulation.

5

Production and characterization of RNA-LNP vaccines

Hydrodynamic size, polydispersity index (PDI) and zeta potential of RNA-LNPs were measured using a Zetasizer Nano ZS90 (Malvern Instruments, Malvern, UK). RNA encapsulation efficiency of LNP was determined using a modified Quant-iT RiboGreen RNA assay (Invitrogen). Endotoxin levels were determined using a Limulus Amebocyte Lysate (LAL) chromogenic assay found to be <0.5 endotoxin unit (EU)/mL.

10

Phylogenomic analyses

Genomic analyses were conducted on 108 *C. difficile* isolates sourced from the work of Bushman, Frederic D., et al (F. D. Bushman, et al., *Cell Host Microbe*. 2020;(28):422-433) and 30 pairs of raw sequencing reads in this study. Initial quality control was implemented with FastQC v0.12.1 and MultiQC v1.14 (P. Ewels, et al., *Bioinformatics*. 2016;(32):3047-3048). Subsequently, de novo assembly was performed on the quality-assessed raw reads utilizing Shovill v1.1.0 (github.com/tseemann/shovill). Genome assemblies were subjected to quality evaluation with CheckM v1.1.6 and BUSCO v5.4.7 (M. Manni, et al., *Mol Biol Evol*. 2021;(38):4647-4654; D. H. Parks, et al., *Genome Res*. 2015;(25):1043-1055). Genome annotations were then performed employing bakta v1.8.1 (O. Schwengers, et al., *Microb Genom*. 2021;(7)). The core-genome alignment was carried out using Roary v3.13.0 (A. J. Page, et al., *Bioinformatics*. 2015;(31):3691-3693), which subsequently served as the input for phylogenomic tree inference through RAxML-NG v1.1 (A. M. Kozlov, et al., *Bioinformatics*. 2019;(35):4453-4455). Phylogeny was rooted at the midpoint using the Tree Of Life (iTOL) v6.8 interactive tool (I. Letunic, P. Bork, *Nucleic Acids Res*. 2021;(49):W293-W296). For subsequent visualization and annotation of phylogeny, ggplot2, and ggtree were utilized (G. Yu, *Curr Protoc Bioinformatics*. 2020;(69):e96; G. Yu, et al., *Mol Biol Evol*. 2018;(35):3041-3043; G. Yu, *Data Integration, Manipulation and Visualization of*

15

20

25

30

Phylogenetic Trees. CRC Press, 2022; S. Xu, et al., *iMeta*. 2022;(4):e56; G. Yu, et al., *Methods in Ecology and Evolution*. 2017;(8):28-36; H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. 2016).

5 *Comparative Genomic Analysis of Protein Domains*

Amino acid sequences for the TcdA CROP domain, TcdB CROP domain, and PPEP-1 (excluding the IL-2 secretion signal domain) were extracted from a VPI10463 strain. Sequences served as seed sequences for constructing a protein-type BLAST database. Subsequently, BlastX was applied to execute a comparative analysis of
 10 amino acid sequences in the database against 138 genomes included in a phylogenetic assessment. The hits with the highest bit score from BlastX were utilized to determine the percentage of coverage and identity (S. F. Altschul, et al., *J Mol Biol*. 1990;(215):403-410; S. F. Altschul, et al., *Nucleic Acids Res*. 1997;(25):3389-3402; G. M. Boratyn, et al., *Biol Direct*. 2012;(7):12; G. M. Boratyn, et al., *BMC Bioinformatics*. 2019;(20):405; C.
 15 Camacho, et al., *BMC Bioinformatics*. 2023;(24):117; C. Camacho, et al., *BMC Bioinformatics*. 2009;(10):421; A. Morgulis, et al., *Bioinformatics*. 2008;(24):1757-1764; Z. Zhang, et al., *J Comput Biol*. 2000;(7):203-214).

Identification of Toxigenic Strains via Comparative Genomic Analysis

20 Coding sequences for tcdA, tcdB, tcdC, tcdE, tcdR, cdu1, and cdd1 were sourced from the comprehensively annotated complete genome of *C. difficile* 630 (NCBI Reference Sequence: NC_009089.1). The nucleotide sequence for the PaLoc region was inferred and extracted by the locations of the flank genes cdu1 and cdd1. These nucleotide sequences served to construct an original nucleotide-specific BLAST
 25 database. To enhance the likelihood of detecting diverse toxin genes, additional sequences of toxin genes and the PaLoc region were incorporated into the database. The raw sequencing reads from 54 representative toxigenic strains were downloaded and assembled (K. E. Dingle, et al., *Genome Biol Evol*. 2014;(6):36-52). The initial round of blastn was conducted against the genomes of these 54 toxigenic strains using the original
 30 database, setting the parameters to accommodate more gaps and mismatches to boost the coverage of hits. The sequences from the highest-scoring hits were subsequently

extracted and used as seed sequences in a new BLAST database. Genomes of good quality were also used to infer and extract sequences from the PaLoc region using flanking genes *cdd1* and *cdu1*. A conserved non-coding region of 115 bp was employed as a marker for non-toxigenic strains (V. Braun, et al., *Gene*. 1996;(181):29-38). A second round of blastn was performed against the 138 genomes included in the previous phylogenetic analysis, using sequences from the new BLAST database. An arbitrary threshold of 80% coverage and 80% identity was applied to the highest-scoring hits and used as criteria for determining the presence or absence of genes. Strains were classified as non-toxigenic if they exhibited in the 115 bp conserved region and lacked toxin genes or the PaLoc region, and vice versa.

In vitro transfection of exemplary RNA LNPs and Western blot

Neuro2a cells were seeded in 12-well plates at a density of 0.2×10^6 cells per well, and transfected 24 hours post seeding at a confluence of around 80%. RNA-LNPs (1 mg/mL) were diluted (1:10) in PBS and added to each well at a final dose of 2.5 µg/well. Medium was aspirated, and cells were directly lysed 24 hours post transfection using the EZLys tissue protein extraction reagent (BioVision #8002-500) supplemented with 1X protease inhibitors (Sigma-Aldrich, #0493116001). Untreated cells were used as controls. Lysate was clarified at 15000 rpm for 5 minutes using a refrigerated centrifuge to pellet cell debris, and protein concentration in the supernatant was determined using the Pierce microBCA assay. A total of twelve (12) micrograms of cell lysate was loaded into a 4-15% precast polyacrylamide gel (Bio-Rad #4561083), transferred to a PVDF membrane using the ThermoFisher iBlot system (dry transfer). Membranes were blocked with 5% skimmed milk extract, washed and incubated for 2 hours with primary antibodies against TcdA and TcdB (Abcam 19953 and 252712) at a final concentration of 4 µg/mL. Zmp-1 (PPEP-1) was detected using a polyclonal serum from immunized mice (Boosted sera at 1:100 dilution, incubation 4 hours). Membranes were then washed three times using 1X TBST and incubated for 1 hour in the presence of an HRP conjugated donkey anti-goat IgG (Abcam 97040) to detect TcdA and TcdB. Zmp-1 (PPEP-1) was detected using an anti-mouse IgG (Abcam 97040). Proteins were visualized using ECL detection reagent (Cytiva #RPN2209) on a GE ImageQuant™

system.

Generation of recombinant TcdA, TcdB, PPEP-1 proteins

A region encoding the TcdA and TcdB CROP domain and a full length
5 PPEP-1 encoding sequence was codon optimized for expression in *E. coli*, gene
synthesized at Genscript, and cloned into the pET30a vector. An N terminal 6XHIS tag
followed by the TEV recognition/cleavage site was introduced after the start (ATG)
codon of all three constructs to allow for purification using affinity chromatography. *E.*
coli BL21 (DE3) form NEB was transformed, and positive clone expanded at a 1L scale,
10 and the recombinant proteins from the supernatant was concentrated, and purified using
the NI-IDA column on an AKTA Avant 150 system. Purified proteins were cleaved with
the TEV protease buffer exchanged into PBS and 5% sucrose (pH 7.4), filter sterilized,
quantified using the micro-BCA assay (Pierce), aliquoted and stored at -80C. Aliquots
were tested for precipitation following for multiple freeze and thaw cycles and the purity
15 was determined using densitometric analysis of a Coomassie blue stained SDS-PAGE gel
under reducing conditions. All proteins displayed more than 90% purity and had the
predicted molecular weights using western blot.

Production of fluorescently labeled proteins recombinant proteins for antigen specific B
20 *cells.*

Fluorescently labeled recombinant TcdA, TcdB and PPEP-1 protein were
prepared using the Lightning-Link R-Phycoerythrin (R-PE) and Lightning-Link (R)
Rapid Alexa Fluor 647 conjugation reagents (Novus Biologicals, 703-0010 and 336-
0005). TcdA, TcdB and PPEP-1 were diluted to 0.5 mg/mL in PBS and 50µg reacted in
25 the presence of 1:10 (v/v) of LL modifier for 3 hours at room temperature. The labeling
reaction was stopped in the presence of 1:10 (v/v) LL quencher for 30 minutes and stored
at 4°C until use.

Mice and immunization regiments

30 C57BL/6 mice were purchased from Jackson Laboratories. BALB/c mice
were obtained from Charles River and used where indicated. All animals were

maintained under specific pathogen-free conditions at either the Children's Hospital of Philadelphia or the University of Pennsylvania. Exemplary vaccines were prepared by thawing RNA-LNPs, diluting them with sterile PBS, and administering intra-muscular immunizations (50 μ L) within two hours. Mice were boosted 2 weeks after an initial
5 prime vaccination where indicated. Recombinant proteins in sucrose buffer (8% w/v) were thawed, mixed at target dose, resuspended in PBS to a final volume of 25 μ L before adding 25 μ L of Alhydrogel (Aluminum hydroxide 2% w/v). Adjuvanted recombinant vaccine were slowly mixed using a pipette and used withing two hours.

10 *Detection of antibodies in mouse serum and feces*

96-well High Bind Stripwell™ Corning 96 Well Clear Polystyrene Microplates plates were coated overnight with 1 μ g/mL of purified TcdA (CROP), TcdB (CROP), or PPEP-1. Plates were washed once with wash buffer (0.5% Tween-20 in PBS) and blocked for two hours at room temperature using a solution of heat inactivated, IgG
15 depleted, protease free bovine serum albumin (2% w/v BSA in PBS). After blocking, plates were washed three times, and mouse sera or feces was serially diluted in the blocking solution and incubated for two hours at room temperature. Plates were washed three times before the addition of horseradish peroxidase-conjugated anti-mouse secondary antibody specific to total IgG (1:10,000) and IgA (1:5,000) in blocking buffer.
20 Plates were incubated for 1.5 hours, washed three times before the addition of tetramethylbenzidine (TMB) substrate solution. The reaction was stopped by adding 50 μ L of 2N sulfuric acid, and the absorbance was measured at 450 nm using a SpectraMax™ 190 microplate reader. Antigen-specific antibody end-point dilution titer was defined as the highest dilution of serum or feces to give an OD greater than the cut-
25 off OD value determined using the Frey Method (A. Frey, et al., *J Immunol Methods*. 1998;(221):35-41)

Experimental model of C. difficile infection, C. difficile enumeration, and toxin titers

8-week-old mice were infected with *C. difficile* as previously described (J.
30 P. Zackular, et al., *Nat Med*. 2016;(22):1330-1334). Antibiotic treatment was administered by providing 0.5g/L cefoperazone in their drinking water *ad libitum* for 5

days, followed by a 2-day recovery period before *C. difficile* infection via oral gavage. Two different *C. difficile* strains were utilized where indicated: VPI10463 at a dose of 1×10^6 spores and CD196 at a dose of 1×10^5 spores. Mice were monitored daily for survival and weight loss, and mice were euthanized when weight loss exceeded 20% of their original body weight. *C. difficile* burdens were quantified by collecting fecal samples at indicated timepoints and plating on taurocholate cycloserine cefoxitin fructose agar (TCCFA). *C. difficile* toxin titers were quantified in the stool using a previously described Vero cell cytotoxicity assay (J. P. Zackular, et al., *Nat Med.* 2016;(22):1330-1334). Briefly, fecal samples were homogenized in sterile PBS, pelleted, and supernatant was filtered through a 0.2um filter. Supernatant was diluted along a tenfold series and incubated overnight with Vero monolayers. Toxin titers in stool were calculated as the reciprocal value of the highest dilution that rounded 100% of the cells and normalized per gram of feces.

15 *Determination of antigen specific CD4 and CD8 T cells, and their polyfunctionality.*

Flow cytometry analysis of T and B cells

Memory B cells: Spleens were collected, processed as single cells, filtered using a 40 µm cell strainers in complete RPMI 1640, centrifuged at 300g for 5 minutes, and red blood cells lysed with ACK (1 min), washed twice, counted and 2 million cells per sample incubated with anti-mouse CD16/32 antibody for 20 min at 4 °C. Cells were then washed with FACS buffer (1% BSA in PBS), and stained for 1 h using antibodies (Table S1). Following staining, cells were washed twice, fixed with 300 µL (1% paraformaldehyde) and samples were acquired on a BD LSR II equipped with 4 laser lines and 18 PMTs. The gating strategy, as well as the antibody list, fluorescent RBD probes, and catalog numbers are provided (Table 3).

Table 3: Gating strategy, antibody list, fluorescent RBD probes, and catalog numbers for flow cytometry analysis of B cells (GC-B cells and MBCs).

Marker	Fluorophore	Dilution (1:x)	Clone	Company	Cat#:
B220	BV421	200	RaA3-6B2	Biolegend	103239
B220	AF488	200	RaA3-6B2	Biolegend	103225

GL7	AF488	200	GL7	Thermo Fisher	53-5902-82
GL7	PECy7	200	GL7	Biolegend	144620
IgD	BV605	300	11-26c.2a	Biolegend	405727
IgM	PECF594	200	R6-60.2	BD	562565
CD4	PECy5	800	H129.19	Biolegend	130312
CD8	PECy5	400	53-6.7	BD	553034
Gr-1	PECy5	800	RB6-8C5	Biolegend	108410
F4/80	PECy5	800	BM8	Thermo Fisher	15-4801-82
CD38	AF700	100	90	Thermo Fisher	56-0381-82
CD19	BV785	100	6D5	Biolegend	115543
Zombie Aqua	V515	400	n/a	Biolegend	423102

T cells: Spleens were collected, processed, as single cells, filtered using a 70 μ m cell strainers in complete RPMI 1640, centrifuged, and red blood cells lysed in ACK lysis buffer to obtain a clear single cell suspension. To measure antigen-specific T cells, 2 million splenocytes were stimulated with 2.5 μ g/mL of TcdA CROP, or TcdB CROP, or PPEP-1 peptide pools (15 mers, 4 amino acid overlapping peptide pool) in a FACS tube for 6 h at 37 °C, 5% CO₂ with 2 mg/mL anti-CD28 (Tonbo, 40-0281-M001) providing co-stimulation. Stimulations proceeded for 1 h before adding 5 mg/mL brefeldin A (Biolegend, 420601), 2 mM monensin (Biolegend, 420701), and 5 mg/mL anti-CD107a (Biolegend, 121610) Alexa Fluor 647 for 5 h. DMSO served as a negative control and the combination of 50 mg/mL phorbol 12-myristate 13-acetate and 1 mg/mL ionomycin served as a positive control. After a total of 6 h, samples were washed with PBS, stained with Live/Dead Aqua for 5 minutes, blocked using anti-mouse CD16/32 antibody for 20 min, and stained extracellularly for 30 min using antibodies (Table 4).

Cells were washed in FACS buffer, fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences, 554714), and stained intracellularly using antibodies for 30 min (Table 4). Following intracellular staining, cells were washed twice, fixed with 300 μ L (1% paraformaldehyde) and samples were acquired on a BD LSR II equipped with 4 laser lines and 18 PMTs. The gating strategy, as well as the antibody list and catalog numbers are provided (Table 4).

Table 4: Gating strategy, antibody list, fluorescent RBD probes, and catalog numbers for flow cytometry analysis of T cells.

Marker	Fluorophore	Dilution (1:x)	Clone	Company	Cat#:
CD107a	FITC	1:100	1D4B	BD	553793
CD4	PerCP-Cy5.5	1:100	GK1.5	Biolegend	100434
CD8a	Pac Blue	1:100	53-6.7	Biolegend	100725
L/D	Aquablue	1:60	-	Thermo Fisher	L34966
CD3e	BV605	1:100	145-2C11	Biolegend	100351
IL-4	BV786		11B11	BD	564006
IL-2	APC	1:100	JES6-5H4	BD	554429
IFN-g	AF700	1:100	XMG1.2	BD	557998
IL-17	APC/Cy7	1:100	TC11-18H10.1	BD	506940
IL-5	PE	1:100	TRFK5	BD	504304
TNF-a	PE/Cy7	1:100	MP6-XT22	BD	557644

Histology and scoring

Samples were fixed in 10% neutral buffered formalin (NBF) prefilled HistoTainer™ II, dehydrated in graded ethanol series, cleared with xylene and embedded in paraffin. Sections (5 µm) were collected on Superfrost™ Plus stain slides (Fisher Scientific, Ottawa, ON, Canada), and stained with Hematoxylin and Eosin. Slides were scanned using a NanoZoomer digital slide scanner (Hamamatsu, Boston, MA, USA) and visualized using the NDP® view 2.0 software (Hamamatsu, Boston, MA, USA)

10

DNA extraction and 16S rRNA gene sequencing

Mice were co-housed for one week before experimental manipulation. Stool samples were collected at indicated timepoints for microbiota analysis. Microbial genomic DNA was extracted using DNeasy Power Soil Kit (Qiagen).

15

TcdA, TcdB and PPEP-1 RNA-LNP vaccines meet quality criteria and are expressed at high levels in vitro

The approach was based on the high potency of RNA-LNP vaccines compared to adjuvanted vaccine platforms (M. Ghattas, et al., *Vaccines (Basel)*.

2021;9(12):1490; M. G. Alameh, et al., *Immunity*. 2022;55(6):1136-1138). Receptor binding domains were targeted including combined repetitive oligopeptide (CROPs)

domains of TcdA and TcdB due to a predicted role these domains play in toxin interaction and entry into cells (G. Carter, et al., *Gut Microbes*. 2010;(1):58-64), lack of cell cytotoxic effects (no cytotoxic domain), and length (863 and 516 instead of 2710 and 2366 amino acids for the full toxins) rendering them amenable for easy RNA manufacturing. A secreted proline-proline endopeptidase 1 (PPEP-1) was selected as a vaccine target due to its predicted role in *C. difficile* adhesion, colonization, and dissemination (J. Corver, et al., *Mol Microbiol*. 2017;(105):663-673; P. J. Hensbergen, et al., *Mol Cell Proteomics*. 2014;(13):1231-1244). This zinc-metalloprotease cleaves multiple adhesion factors on a *C. difficile* cell surface, providing an opportunity to impact multiple cellular factors and processes involved in *C. difficile* fitness in the gut. To maximize humoral immune response against these immunogens, putative N-glycosylation sites were identified and replaced, sequences were codon optimized for improved intracellular expression, and an engineered secretion signal was added upstream of the open reading frames (Fig1. A). To test cross-coverage of the RNA constructs across diverse *C. difficile* strains, a comparative genomic analysis was conducted to assess the shared amino acid sequence identity of the TcdA, TcdB, and PPEP-1 sequences in the RNA constructs to sequences from 138 unique *C. difficile* strains, across five different clades from both adults and children (Fig 1B). The construct for PPEP-1 has 100% identity across the *C. difficile* strains examined, and high identity of TcdA and TcdB associated with phylogenomic clade. Resulting RNA sequences were synthesized and encapsulated into lipid nanoparticles (LNPs) using a proprietary ionizable lipid shown to be potent for vaccine application in multiple preclinical models including rodents and non-human primates (REF). The average hydrodynamic diameter (size) of TcdA, TcdB and PPEP-1 RNA -LNPs was 80 nm, and particle formulations homogenous with polydispersity index below 0.2 (Fig 1D) Encapsulation efficiency (EE), and payload integrity were determined using the Ribogreen™ reagent, and microfluidic based electrophoresis respectively. The EE of all RNAs were high (>95%), and the payload integrity remained unchanged pre and post encapsulation (Fig 1D). Transfection of TcdA, TcdB and PPEP-1 RNA-LNPs into Neuro2A cells was used to assess protein expression. Both supernatant and cell pellets were found to contain TcdA CROP, TcdB CROP, and PPEP-1 confirming that all three proteins were expressed and secreted (Fig

1F).

Immunization against C. difficile proteins supports antigen-specific cellular and humoral immune responses.

5 The present Example provides compositions and related methods that show protection against *C. difficile* infection in a subject. Moreover, the present Example demonstrates that an exemplary multivalent RNA-LNP vaccine induces a robust cellular and humoral immune response in a subject.

 Immunogenicity of RNA-LNP constructs was assessed by vaccinating
10 mice twice intramuscularly (IM) with either aTcdA/TcdB (bivalent) or a TcdA/TcdB/PPEP1 (trivalent) RNA-LNPs, a trivalent recombinant protein with alum as an adjuvant, or a luciferase RNA-LNP as a control. Two weeks after a last immunization, antigen-specific antibody titers were analyzed by endpoint ELISA. Both bivalent and trivalent RNA-LNP vaccination induced robust anti-toxin immunoglobulin (IgG)
15 responses that were modestly higher compared to those elicited by a recombinant protein + alum (Figure 2A). IgG against PPEP-1 was also detected in serum of trivalent vaccinated mice, although to a lesser extent compared to anti-toxin IgG (Figure 2A). Notably, anti-PPEP-1 IgG was significantly higher in mice vaccinated with RNA-LNP vaccine compared to a recombinant protein + alum.

20 Antibody responses in both C57BL6 (Figure 2A) and Balb/c (Figure 11A) mice were relatively high compared to Alum adjuvanted vaccines and showed dose dependent increase following a single immunization (Figure 11B). Administration of a second dose increased overall antibody titers by around a factor of 15 (Figure 12). Total IgG responses against PPEP-1 were generally very low and required at least one
25 additional vaccination to increase significantly above baseline levels indicating that PPEP-1 protein is a poor humoral immunogen. In contrast to IgG, IgA responses in circulation (serum) were significantly lower for all immunogens tested and require at least two immunizations (Figure 12C and 12D). The comparison between a monovalent, bivalent, and trivalent RNA-LNP vaccines show no reduction in antibody titers with
30 increased valency indicating absence of immune competition and corroborating previous results we, and others, obtained with multivalent influenza vaccines (C. P. Arevalo, et al.,

Science. 2022;(378):899-904; N. Pardi et al., *Nat Commun*. 2022;(13):4677).

Ability of exemplary bivalent and trivalent RNA-LNP vaccines to induce Tfh cell differentiation was tested. Mice were immunized IM with either a 1 µg (low) or 5 µg (high) dose and compared total Tfh responses to animals that received a trivalent recombinant protein plus alum (Figure 2B). Total Tfh increased significantly compared to naïve mice and mice treated with a 5 µg of luciferase RNA-LNP control, suggesting that observed Tfh responses are mainly antigen specific, and not due to spontaneous activation in draining lymph nodes (dLNs). An increase in the number of Tfh cells following RNA-LNP vaccinations and their role in supporting B cell responses in a germinal center (GC) suggests this would correlate with a strong GC B cell response in dLNs. Indeed, bivalent and trivalent RNA-LNP vaccines induced a dose-dependent, antigen-specific GC B cell response, and animals vaccinated with a bivalent or trivalent RNA-LNP presented higher numbers of antigen specific B cells compared to animals vaccinated with a luciferase RNA-LNP controls (Figure 2). Notably, a trivalent RNA-LNP vaccination induced PPEP-1 specific B cells to a lesser extent compared to anti-toxin specific responses (Figure 2C). GC B cell responses in spleen showed similar increase in numbers compared to naïve mice, and antigen specific GC B cell responses were significantly higher than alum adjuvanted recombinant protein groups (Figure 11). Antigen specific B cells showed similar result patterns as observed in dLNs with a dose dependent increase in number and frequency following vaccination with bivalent and trivalent RNA-LNP vaccines. Despite a modest increase (~3-4X) in IgG antibodies in mice treated with RNA-LNP vaccines compared to those receiving alum adjuvanted recombinant proteins, antigen specific B and GC B cells were significantly higher in RNA-LNP vaccine groups compared to alum adjuvanted recombinant vaccine groups (Figure 2C, Figure 2D).

Among other things it is an insight of the present disclosure that RNA-LNP vaccination activates all arms of adaptive immunity, including antigen specific CD4⁺ and CD8⁺ T cells. Moreover, CD4⁺ T cell responses and skewing are associated with improved outcomes to CDI. Vaccine induced T cell responses and their polyfunctionality were assessed. Mice were immunized IM with a one or two doses of either 1 µg (low) or 5 µg (high) of a bivalent or a trivalent RNA-LNP, a bivalent

recombinant protein with alum adjuvant, or a luciferase RNA-LNP as a control. Two weeks post-immunization, lymphocytes were harvested from spleen and dLN, cultured cells ex-vivo with TcdA, TcdB, or PPEP-1 peptide pools, CD4⁺ and CD8⁺ responses were assessed by flow cytometry. Antigen specific T cell responses increased with vaccine doses and were significantly higher in RNA-LNP vaccines compared to alum adjuvanted recombinant protein vaccines (Figure 2D, Figure 2D, and Figure 5A). Higher frequency of antigen specific T cell responses were observed with increasing vaccine valency. Antigen specific CD4⁺ responses against TcdA and B were higher than CD8⁺ responses and were mainly IL-2, and TNF- α +. In contrast, PPEP-1 showed very poor CD4⁺ responses but significant higher CD8⁺ responses when compared to TcdA and B. The experiment was repeated and stimulation with peptide pools versus recombinant whole proteins were compared to exclude effects of spleenocyte stimulation method. There was no difference between the two methods of stimulation (Figure 13). The study was repeated in Balb/c mice to exclude effects of MHC restriction, and similar results with TcdA and B inducing a strong antigen specific CD4⁺ T response while PPEP-1 induces a strong CD8⁺ response were observed (Figure 14).

Frequency of CD4⁺ and CD8⁺ T cells that express one, two or multiple cytokines were determined using boolean gating given that polyfunctional T cells have been associated with improved protection. Frequency of polyfunctional cells was high with double positive being the most abundant followed by triple positive cells (Figure 17). In sum, humoral and cellular immunogenicity data demonstrate that a RNA-LNP vaccine provide greater immunization relative to alum adjuvanted recombinant vaccines in multiple strains of mice, providing an improved platform for vaccination against *C. difficile*.

Accordingly, the present Example demonstrates that RNA-LNP vaccination against *C. difficile* proteins induces Tfh and GC B cell responses in accordance with robust antigen-specific antibody responses in mice.

RNA-LNP vaccines elicit protective immunity against C. difficile toxins

The present Example provides compositions and related methods that show protective immunity against *C. difficile* infection in a subject.

A lethal dose 100 (LD₁₀₀) of intraperitoneally (IP) injected recombinant TcdA and TcdB was determined to test whether anti-toxin antibodies elicited by exemplary RNA-LNP vaccines would neutralize *C. difficile* toxins *in vivo* (Figure 14). Mice were then immunized with either monovalent TcdA or TcdB RNA-LNP vaccines or phosphate buffered saline (PBS) as a control. Mice were terminally bled 2 weeks after immunization, and sera from vaccinated mice were incubated with either recombinant TcdA or TcdB. Naïve mice were then challenged with either recombinant TcdA or TcdB alone or recombinant TcdA or TcdB that was treated with sera from vaccinated mice. Mice were monitored for survival over 72 hours (Figure 5B). Mice that were challenged with toxins incubated with sera from PBS-immunized control animals behaved similarly to the mice that received recombinant toxin alone and all died within 4 hours post-challenge. However, no mortality was observed in mice treated with monovalent immunized sera and toxin.

In a second parallel strategy to test neutralization response following RNA-LNP vaccination, mice were immunized with either monovalent TcdA or TcdB RNA-LNP vaccines. As controls, mice were also immunized with TcdA or TcdB recombinant proteins with or without adjuvant. All mice were immunized twice. Two weeks after second immunization, mice were challenged with five times the LD₁₀₀ of IP injected recombinant TcdA or TcdB via IP and monitored daily for survival (Figure 5C). Mice that received TcdA or TcdB recombinant protein vaccination without adjuvant all died within 1 day-post toxin challenge. Inclusion of alum as an adjuvant improved survival only 20%; however, immunization with monovalent TcdA or TcdB RNA-LNP vaccines significantly improved survival to 100%. Collectively, these data support highly reported success of RNA-LNP vaccines as a platform for inducing high titers of circulating, neutralizing antibodies and highlight this platform's utility in targeting bacterial proteins and toxins.

Accordingly, the present Example demonstrates that a RNA-LNP vaccine described herein provides protective immunity in a subject.

RNA-LNP vaccines protect mice from severe C. difficile infection.

The present Example provides compositions and related methods that

show protection from severe *C. difficile* infection in a subject.

Mice were immunized twice prior to *C. difficile* infection (CDI) (Figure 6A). All unvaccinated control animals lost weight rapidly. By day 2 post-infection, unvaccinated mice were either deceased or moribund. All mice that received bivalent or trivalent RNA-LNP vaccines were protected from lethal infection (Figure 6B). In accordance with survival data, by day 2 post-infection, vaccinated mice presented with mild disease; they lost an average of 5% body weight and were alert and active in the cage (Figure 6C and Figure 6D). A single administration of a trivalent or bivalent vaccine was sufficient to protect mice from lethal CDI (Figure 3). Additionally, vaccination against *C. difficile* toxin was necessary for protection since immunization with PPEP-1 RNA-LNP alone did not protect mice from lethal infection (Figure 13). To define impact of vaccination on *C. difficile* virulence and dynamics, bacterial burdens of *C. difficile* and toxin titers in the stool of mice were evaluated over the course of infection. Both bivalent and trivalent vaccination led to clearance of *C. difficile* from the GI tract at similar rates (Figure 6E). However, mice vaccinated with an exemplary trivalent vaccine showed significantly reduced toxin titers in their stool on days 7-, 12-, and 14-post infection compared to mice vaccinated against toxin alone (Figure 6F). Together, these data show that immunization against PPEP-1 has an added benefit in reducing *C. difficile* virulence and demonstrates RNA-LNP vaccines have a capacity to promote rapid control of toxin-producing *C. difficile* in a gut.

Given diversity of circulating strains in a hospital and community settings, a *C. difficile* strain that is phylogenetically divergent from VPI10463 and from a clinically relevant clade was tested for cross-strain efficacy. Both exemplary bivalent and trivalent vaccines protected mice from infection with a ribotype 027 strain, CD196, demonstrated by increased survival, reduced weight loss, and improved CSS compared to unvaccinated controls (Figure 13).

Thus, the present Example demonstrates that compositions (e.g., RNA-LNP vaccines) described herein provide broad immunity across phylogeny of *C. difficile*.

RNA-LNP vaccines prompt mucosal immunity during CDI

The present Example provides compositions and related methods that

show mucosal immunity during *C. difficile* infection (CDI) in a subject.

Generation of an antigen-specific mucosal immune response remains a primary challenge for development of effective vaccine against *C. difficile*. While the Examples of the present disclosure demonstrate that immunization via IM elicits a robust systemic immune response, effective protection from infection may require local mucosal immunity. Thus, mucosal TcdA-, TcdB-, and PPEP-1-specific antibodies were assessed in stool of vaccinated mice prior to and during infection. Surprisingly, both exemplary bivalent and trivalent vaccines led to anti-toxin IgG and IgA titers in stool prior to infection (Figure 10A). IgG and IgA antibodies both slightly increased over the course of acute infection, a marked expansion of anti-toxin antibodies was observed for both classes of antibodies 14 days post-infection, suggesting a robust recall response. Notably, mice immunized with an exemplary trivalent vaccine did not demonstrate anti-PPEP-1 stool antibody titers compared to bivalent or unvaccinated controls (Figure 10A), suggesting an immune response to a PPEP-1 RNA does not reach a threshold for a measurable mucosal immune response using these methods. Together these data clearly demonstrate an exemplary RNA-LNP vaccine platform described herein is effective for generation of antigen-specific mucosal immune responses to *C. difficile* and other enteric pathogens.

Accordingly, the present Example demonstrates that an exemplary RNA-LNP vaccine described herein shows mucosal immunity during *C. difficile* infection (CDI) in a subject.

RNA-LNP vaccines against C. difficile do not impact intestinal microbiota.

The present Example provides compositions and related methods that do not impact intestinal microbiota during *C. difficile* infection (CDI) in a subject.

Thus, 16S rRNA gene sequencing was performed on fecal samples collected prior to vaccination, one week following a last immunization, and two days after a cessation of antibiotic treatment (Figure 9A). Mice vaccinated with either exemplary trivalent or bivalent RNA-LNP vaccines showed no difference in microbial diversity (α -diversity) and no significant shift in community structure when compared to pre-vaccination timepoints and unvaccinated controls (Figure 9B- Figure 9D).

Additionally, vaccination did not affect relative abundances of beneficial taxa known to antagonize *C. difficile*, including members affiliated with the Clostridiales family (Figure 5). Vaccination also had no effect on microbial diversity and composition following antibiotic treatment (Figure 4), demonstrating a vaccination did not enhance antibiotic-mediated perturbation or impact stability of the microbial community.

Thus, the present Example demonstrates that RNA-LNPs described herein do not impact intestinal microbiota during *C. difficile* infection (CDI) in a subject.

Example 2: A multivalent RNA-LNP vaccine protects against *C. difficile* infection

Clostridioides difficile causes a wide range of gastrointestinal disorders, ranging from mild diarrhea to toxic megacolon and death (H. Dudukgian et al., *J. Gastrointest. Surg.* 2010;(14):315–322; C. P. Kelly et al., *Annu. Rev. Med.* 1998;(49):375–390; V. K. Viswanathan et al., *Gut Microbes.* 2010;(1):234–242). In the USA, *C. difficile* is one of the most reported nosocomial pathogens; worldwide, *C. difficile* infection (CDI) is a major public health threat (F. C. Lessa et al., *N. Engl. J. Med.* 2015;(372):825–834). Previous vaccine trials for CDI prevention focused on adjuvant-supplemented toxoid or recombinant vaccines targeting the combined repetitive oligopeptide (CROP) domains or receptor binding domains (RBDs) of two potent *C. difficile* toxins, TcdA and TcdB, to inhibit toxin binding to intestinal epithelial cells (M. Henderson et al., *Vaccines.* 2017;(5):25). A recent Pfizer phase 3 clinical trial showed promise in reducing the severity and duration of CDI but failed to prevent initial infection, highlighting the challenges in developing effective *C. difficile* vaccines (NCT03090191). This setback, along with Sanofi discontinuing its *C. difficile* vaccine (NCT01887912), underscores the need to explore alternative vaccination strategies (T. V. Riley et al., *Vaccine.* 2019;(37):7300–7306).

Nucleoside-modified messenger RNA (mRNA) vaccines have emerged as a leading platform against a myriad of infectious pathogens, including lethal bacterial pathogens (M. G. Alameh et al., *Curr. Top. Microbiol. Immunol.* 2020;(440):111–145; M. D. Buschmann et al., *Vaccines.* 2021;(9):65; E. Kon et al., *Sci. Adv.* 2023;(9):eadg1036). Therefore, a multivalent RNA strategy targeting multiple virulence factors would prevent *C. difficile*-associated disease. The initial vaccine strategy included

RBDs and CROP domains of both toxins, and the metalloprotease virulence factor Pro-Pro endopeptidase 1 (PPEP-1/Zmp1) (Figure 1H). Since PPEP-1 is a highly conserved factor that modulates pathogen motility and adhesion through the cleavage of multiple factors on the *C. difficile* cell surface, targeting a fitness and virulence regulator would promote decolonization after disease onset (P. J. Hensbergen et al., *FEBS Lett.* 2015;(589)3952–3958).

Genetically diverse strains of *C. difficile* circulate in both hospital and community settings (M. J. Mansfield et al., *PLOS Pathog.* 2020;(16):e1009181). To explore the conservation of the vaccine targets, a comparative genomic analysis of shared amino acid sequence identity was conducted between the TcdA, TcdB, and PPEP-1 sequences in the RNA constructs and 137 representative *C. difficile* strains (Figure 1A) (F. D. Bushman et al., *Cell Host Microbe.* 2020;(28):422–433.e7). The construct for PPEP-1 had minimum 97.4%, median 99.5%, and average 99.4% amino acid similarity across the *C. difficile* strains examined (Figure 1A). Among toxigenic strains, TcdA had minimum 39.4%, median 97.1%, and average 90.9% identity, and TcdB had minimum 87.0%, median 100%, and average 97.8% amino acid identity (Figure 1A). Together, these constructs cover the diversity of clinically relevant and potentially emergent strains and represent strong candidates for further development. Resulting RNA sequences were synthesized and encapsulated into lipid nanoparticles (LNPs) (M. G. Alameh et al., *Immunity.* 2022;(55):1136–1138; K. Musunuru et al., *Nature.* 2021;(593):429–434; D. Weissman et al., *Cell Host Microbe.* 2021;(29):23–31.e4). All RNA-LNPs passed quality control parameters (Figure 1I and Figure 1J) and all three proteins were expressed and secreted following transfection of Neuro2A cells (Figure 1K). PPEP-1 expression was considerably reduced compared with that of the toxins (Figure 1K). In vitro expression of TcdA and TcdB RNA-LNPs did not reduce cell viability, demonstrating the safety of engineered immunogens for intramuscular administration in animals (Figure 1L).

RNA-LNP vaccination induces robust immune responses

Multivalent RNA-LNP vaccines elicit high antibody titers against multiple antigens simultaneously (C. P. Arevalo et al., *Science.* 2022;(378):899–904. N. Pardi et al., *Nat. Commun.* 2022;(13)4677). To examine the immunogenicity of the vaccines,

mice were vaccinated with TcdA/TcdB (bivalent) or TcdA/TcdB/PPEP-1 (trivalent) RNA-LNPs, trivalent recombinant protein with alum as an adjuvant, or luciferase RNA-LNP, and serum antigen-specific antibody titers were measured by endpoint ELISA (Figure 1B). RNA-LNP vaccines induced ~2 to 4 times higher anti-toxin immunoglobulin G (IgG) responses compared with that of recombinant protein + alum (Figure 1B). IgG against PPEP-1 was also detected at lower levels compared with anti-toxin IgG (Figure 1B). Notably, anti-PPEP-1 IgG was significantly higher with the RNA-LNP vaccine compared with recombinant protein + alum (Figure 1B). A modest dose dependency was observed (Figure 12A), and IgG titers significantly increased following a second vaccination (Figure 12A-Figure 12C). IgA responses in circulation were significantly lower for all immunogens and required at least two immunizations (Figure 12D and Figure 12E). Monovalent, bivalent, and trivalent RNA-LNP vaccines showed no reduction in antibody titers with increased valency, corroborating previous results with multivalent influenza vaccines (Figure 11C) (C. P. Arevalo et al., *Science*. 2022;(378):899–904. N. Pardi et al., *Nat. Commun.* 2022;(13)4677). Finally, antibody responses were independent of mouse genetic background (Figure 11A and Figure 11B). Together, these data demonstrate the robust capacity of RNA-LNP vaccines to elicit antigen-specific immune responses to *C. difficile* virulence factors.

T follicular helper (Tfh) cells drive germinal center (GC) reactions and are important for potent antibody induction (S. Crotty, *Immunity*. 2019;(50):1132–1148; C. G. Vinuesa et al., *Annu. Rev. Immunol.* 2016;(34)335–368). Vaccination with RNA-LNP vaccines led to a significant increase in total Tfh cells (Figure 1C and Figure 4A) and, given their role in supporting B cell responses in the GC, this may correlate with strong GC B cell responses in the spleen and the draining lymph nodes (Figure 1D, Figure 1E, Figure 4 and Figure 3). Indeed, RNA-LNP vaccines induced a dose-dependent, antigen-specific GC B cell response (Figure 1D, Figure 1E, and Figure 3). In accordance with antibody titers, trivalent RNA-LNP vaccine induced PPEP-1-specific B cells to a lesser extent compared with anti-toxin-specific B cells (Figure 1E and Figure 3B). Despite modest increases in IgG antibodies compared with alum-adjuvanted recombinant proteins, GC B cells and antigen-specific B cells were markedly higher in the RNA-LNP groups (Figure 1D, Figure 1E, and Figure 3). These data demonstrate that RNA-LNP

vaccination against *C. difficile* virulence factors induces robust Tfh and GC B cell responses that are associated with potent and long-lived responses (S. Crotty, *Immunity*. 2019;(50):1132–1148; C. G. Vinuesa et al., *Annu. Rev. Immunol.* 2016;(34)335–368).

Additionally, immunogenicity studies in hamsters showed that the vaccine
5 was well-tolerated, with no observed weight loss or adverse effects (Figure 22A and Figure 22B). A single immunization induced robust anti-toxin systemic IgG levels (Figure 22C). A second immunization was necessary to elicit anti-PPEP-1 IgG response (Figure 22B), supporting observations in mice (Figure 12C). By contrast, a single immunization with recombinant protein + alum was insufficient to induce anti-TcdB IgG
10 response (Figure 22C). Figure 29 depicts the results of example experiments wherein hamsters were injected with 2µg/antigen at days 0 and 21. Serum was collected at days 21 and 36. Endpoint ELISA titers were measured on serum samples.

These data demonstrate the capacity of RNA-LNP vaccines to drive robust antigen-specific immune responses to *C. difficile* virulence factors in multiple
15 relevant animal models.

RNA-LNP vaccination activates all arms of adaptive immunity, including antigen-specific CD4+ and CD8+ T cells. CD4+ T cell responses and skewing are essential for highly neutralizing antigen-specific antibody responses (M. M. Painter et al., *Immunity*. 2021;(54):2133–2142.e3). To assess T cell responses, splenocytes were
20 stimulated with overlapping peptide pools covering TcdA, TcdB, or PPEP-1 (Figure 13 and Figure 18). T cell responses increased with vaccine dose and were significantly higher with RNA-LNP vaccines compared with alum-adjuvanted recombinant protein vaccines (Figure 13B and Figure 13C). Notably, higher frequencies of antigen-responsive T cells with increasing vaccine valency were observed (Figure 13B and Figure 13C).
25 TcdA and TcdB peptide-responsive CD4+ T cells were higher than CD8+ responses, characterized by increased IL-2 and TNFα expression (Figure 13B and Figure 13C). By contrast, PPEP-1 showed poor CD4+ responses but higher CD8+ responses (Figure 13B and Figure 13C). Similar results were obtained with full recombinant proteins (Figure 13D and Figure 13E). Polyfunctional T cells are associated with improved protection
30 during infection and can support a more robust immune response (J. R. Almeida et al., *J. Exp. Med.* 2007;(204):2473–2485; P. A. Darrah et al., *Nat. Med.* 2007;(13):843–850; M.

L. Precopio et al., *J. Exp. Med.* 2007;(204):1405–1416). Following multivalent RNA-LNP vaccination, the frequency of polyfunctional T cells was high, with double positive being the most abundant followed by triple positive cells (Figure 18). These data show that RNA-LNP vaccines against *C. difficile* virulence factors provide improved immune responses compared with alum-adjuvanted recombinant vaccines, further supporting their viability for clinical development.

RNA-LNP vaccines protect against C. difficile toxins

C. difficile–associated disease is primarily driven by the effects of the toxins (M. C. Abt et al., *Nat. Rev. Microbiol.* 2016;(14):609–620; G. P. Carter et al., *mBio.* 2015;(6):e00551). To assess whether anti-toxin antibodies elicited by RNA-LNP vaccines would effectively neutralize *C. difficile* toxins in vivo, mice were immunized with monovalent vaccines and challenged them with five times the LD₁₀₀ of intraperitoneal recombinant TcdA or TcdB and monitored survival (Figure 1F, Figure 1G, Figure 14A and Figure 14B). Mice vaccinated with recombinant TcdA or TcdB, without adjuvant, all died within one day of the toxin challenge (Figure 1F and Figure 1G). The inclusion of alum as an adjuvant improved survival by only 20%; however, immunization with monovalent TcdA or TcdB RNA-LNP vaccines improved survival to 100% (Figure 1F and Figure 1G). Notably, serum transfer from mice vaccinated with RNA-LNPs was sufficient to neutralize toxin and protect from lethal challenge (Figure 14C and Figure 14D). Collectively, these data show the efficacy and potency of RNA-LNP vaccines to elicit, neutralizing antibodies and highlight this platform's value in targeting bacterial proteins and toxins.

RNA-LNP vaccines do not compromise intestinal microbiota

The gut microbiota is important for health and provides resistance to invading pathogens, including *C. difficile*, which makes antibiotic use that compromises the microbiota a primary risk factor for CDI. Therefore, it is critical that vaccines against *C. difficile* do not have unintended effects on this microbial community, as disruptions could impact susceptibility to infection. Additionally, little is known about the effect of RNA-LNP vaccines on the microbiota. Thus the composition of the gut microbiota was

examined both after vaccination and following antibiotic treatment (Figure 6A). α and β diversity of the fecal microbiota were not affected, showing that immunization does not enhance antibiotic-mediated perturbation or affect microbiome stability (Figure 6, Figure 5, and Figure 8). Vaccination did not lead to substantial shifts in relative abundances of any taxa, including those known to antagonize *C. difficile*, such as the Clostridiales family (Figure 5). These data show the safety of our approach and vaccine targets in a mouse model. Additional testing in dirty mice and other models will confirm findings and recapitulate the immune responses reported.

10 *RNA-LNP vaccines protect from severe CDI*

Whether RNA-LNP vaccination was protective in a murine model of infection was next investigated (Figure 9A) (A. B. Smith et al., *Nature*. 2022;(611):780–786; J. P. Zackular et al., *Nat. Med.* 2016;(22):1330–1334; J. Soto Ocaña et al., *Sci. Adv.* 2023;(9):eadh5552; C. M. Theriot et al., *Gut Microbes*. 2011;(2):326–334). Following infection with 20 times the lethal dose of *C. difficile* (strain VPI 10463), vaccinated mice were fully protected from mortality while all unvaccinated mice were moribund by day 2 following infection (Figure 9B). Vaccinated mice presented with mild disease, lost an average of 5% body weight, and were all alert and active following infection (Figure 9C and Figure 9D). Notably, a single low-dose vaccination was sufficient to protect mice from lethal CDI (Figure 15A to Figure 15D). Additionally, vaccination against *C. difficile* toxin was necessary for protection, as immunization with PPEP-1 RNA-LNP alone did not protect the mice (Figure 15E and Figure 15F). Disease pathology was assessed and found that CDI induces edema, inflammation, and epithelial damage independent of vaccination status, suggesting that although vaccination reduces the effects of acute infection it is insufficient to protect the intestinal epithelium from the local cytotoxic effects of TcdA and TcdB in this mouse model (Figure 9E). Furthermore, this suggests that translocation and systemic effects of the toxins may play a key role in lethal disease and that vaccine-mediated systemic immunity is important for protection against lethality during CDI.

To define the effects of vaccination on *C. difficile* persistence and virulence, we evaluated pathogen burden and toxin titers in the stools of mice during

infection (Figure 9F and Figure 9G). All vaccinated mice cleared *C. difficile* from the GI tract at similar rates, as measured by colony forming units (CFUs) (Figure 9F). However, trivalent vaccination resulted in enhanced clearance of toxin titers in the stool compared with bivalent vaccination (Figure 9G). This indicates that despite low expression (Figure 5 1K) and low immunogenicity (Figure 1B and Figure 1E), immunization against PPEP-1 has an added benefit in reducing *C. difficile* pathogenesis (Figure 9G). Together, these data show that a multivalent strategy can promote enhanced clearance from the gut, and we postulate that this response can be improved with greater immunogenicity of PPEP-1. We also observed that bivalent and trivalent vaccines protected mice with a
10 phylogenetically divergent ribotype 027 strain of *C. difficile* (Figure 23), further validating high conservation of our vaccine targets and broad protection conferred by vaccination.

RNA-LNP vaccines against C. difficile prompt mucosal immunity

15 Generation of antigen-specific mucosal immunity remains a primary challenge for developing an effective *C. difficile* vaccine. Recent reports have shown that TcdB aids *C. difficile* in subverting the immune system (K. M. Norman et al., *Cell Rep.* 2024;(43);114245). Indeed, naïve mice mount a poor natural humoral immune response to *C. difficile* toxins (Cd only) at day 14 post-infection (Figure 9H, Figure 9I, and Figure
20 23), whereas vaccination elicited significant anti-toxin IgG and IgA titers in stool 14 days post-boost (vaccine only) (Figure 9H, Figure 9I, and Figure 22D). Notably, vaccination alone was sufficient to mount robust anti-toxin mucosal antibodies but insufficient to elicit anti-PPEP-1 antibodies (Figure 9H, Figure 9I, and Figure 22D). However, when antibody titers were examined in the stool of vaccinated mice 14 days post-infection
25 (vaccination + Cd) we observed a four- to tenfold increase in anti-toxin antibodies compared with vaccination alone and observed anti-PPEP-1 IgA (Figure 9H and Figure 9I), suggesting that hybrid immunity elicited by vaccination and infection enhances mucosal immunity against PPEP-1.

 Individuals with prior CDI are at an elevated risk of relapsing or recurrent
30 infections, making such populations ideal targets for vaccination. To determine the ability of RNA-LNP vaccines to elicit mucosal antibodies when administered after infection,

mice were vaccinated after recovery from a sublethal infection (Cd + vaccination) (Figure 9H and Figure 9I). At day 14 post-boost, these mice had a ~100- to 300-fold increase in anti-toxin IgA compared with infection alone, and a one to fourfold increase compared to vaccination alone (Figure 9H). Infection prior to vaccination also led to a 60-fold increase in anti-PPEP-1 IgA in stools (Figure 9H). Similar trends were also observed in mucosal anti-PPEP-1 IgG (Figure 9I), and vaccination also elicited antigen-specific mucosal antibody responses in hamsters (Figure 22D). Together, these data show that RNA-LNP vaccination leads to mucosal immunity to *C. difficile* toxins and that this vaccine strategy primes the immune response for mucosal antibodies against surface proteins like PPEP-1 during a subsequent infection. Importantly, these data also establish that vaccination can overcome dampened immunity to natural infection and lead to strong mucosal immunity in patients with recent prior CDIs. Notably, mucosal antibody titers are reduced compared to systemic titers and we hypothesize that induction of a potent mucosal IgA response against each immunogen would improve the efficacy of vaccines against *C. difficile* significantly.

Durable immunity protects against CDI

Long-term durability of vaccine immune responses are essential for successful protection against acute as well as relapsing and recurrent infections. IgG titers in the serum of immunized mice remained stable seven weeks post-boost (Figure 20A). Mucosal TcdA and TcdB antibodies also remained stable over time whereas PPEP-1 antibodies were not detected (Figure 20B). Vaccination induced similar levels of TcdA- and TcdB-specific splenic memory B cells (MBCs), irrespective of vaccine valency (Figure 20C). PPEP-1 specific MBCs were significantly higher compared with anti-toxin MBCs (Figure 20C) but with lower overall antibody responses, indicating a lack of correlation between MBCs and circulating antibodies for PPEP-1, a phenomenon that has been described for other antigens (H. Leyendeckers et al., *Eur. J. Immunol.* 1999;(29);1406–1417). Whether long-lived vaccine responses could protect mice from severe CDI was next examined using a delayed infection model (Figure 20D). Long-lived immunity protected mice from lethal CDI (Figure 20E). Mice lost 6% of their original body weight during acute infection and presented with mild disease (Figure 20F and

Figure 20G).

The effect of RNA-LNP vaccination on recurrent infection was examined using a long-term reinfection model. More than 6 months after vaccination and a primary infection, we rechallenged immunized mice who had recovered and completely cleared
5 *C. difficile* with a second CDI (Figure 20H). Similar to primary infection, vaccination protected mice from mortality (Figure 20I). All mice presented with mild disease, lost an average of 8% body weight during secondary infection, and remained alert and active (Figure 20J and Figure 20K). These data show the capacity of the RNA-LNP vaccine platform to induce long-lived memory responses that confer durable protection against
10 CDI.

Targeting C. difficile spores reduces pathogenesis

So far, it has been shown that the RNA-LNP vaccine platform induces robust anti-toxin immune responses that prevent morbidity and mortality associated with
15 severe CDI. Moreover, by targeting non-toxin virulence factors decolonization of toxin-producing *C. difficile* can be achieved. However, *C. difficile* is a highly complex organism that alternates between two main phases in its life cycle: the vegetative cell and spore. *C. difficile* spores—the primary mode of transmission for this pathogen—are highly resistant to traditional chemical disinfectants, phagocytosis, and antimicrobials (A.
20 N. Edwards et al., *Front. Microbiol.* 2016;(7):1698; D. Paredes-Sabja et al., *PLOS ONE*. 2012;(7):e43635), presenting a major challenge in treating CDI. It was postulated that addition of a spore-specific immunogen could further improve the vaccine. The *C. difficile* exosporium morphogenic protein (CdeM), a cysteine-rich protein expressed on the outer spore coat that is specific to *C. difficile* among other spore-formers in the gut
25 was selected (P. Calderón-Romero et al., *PLOS Pathog.* 2018;(14):e1007199). Comparative genomic analysis of the CdeM sequence in the RNA construct showed high conservation across *C. difficile* phylogeny (Figure 21A and Figure 24A) (F. D. Bushman et al., *Cell Host Microbe*. 2020;(28):422–433.e7).

Immunization with a tetravalent vaccine elicited antigen-specific systemic
30 and mucosal antibodies against all immunogens without compromised immunogenicity (Figure 21B and Figure 24B). To experimentally determine whether a multivalent

approach targeting *C. difficile* spore and vegetative cells influences acute CDI, mice were immunized with either CdeM/ PPEP-1/TcdA/TcdB (tetravalent), CdeM/ TcdA/TcdB (spore trivalent), PPEP-1/TcdA/TcdB (vegetative trivalent), or CdeM monovalent RNA-LNPs. CdeM alone was insufficient to protect mice from lethal infection whereas mice immunized with multivalent vaccines presented with mild disease and were fully protected (Figure 21C- Figure 21E). Notably, inclusion of CdeM (tetravalent and spore trivalent) immunization resulted in lower *C. difficile* CFUs and toxin in stools on day 1 post-infection compared with vegetative trivalent immunization (Figure 6F and Figure 6G), suggesting that vaccination against spores can limit initial colonization even when challenged with a high spore inoculum. At day 7 post-infection, tetravalent immunization resulted in substantially reduced toxin titers in stools compared with spore or vegetative trivalent RNA-LNPs vaccines (Figure 6H and Figure 6I). Together these data indicate that a multivalent approach can protect from disease, reduce colonization, and promote decolonization of toxigenic *C. difficile* from the GI tract.

Finally, to model human immune responses to vaccination, an immunogenicity study was performed in relatively aged nonhuman primates. Although a single immunization was sufficient to induce anti-TcdA and anti-CdeM IgG, two immunizations were necessary for antigen-specific IgG against all immunogens (Figure 21J), further highlighting the capacity of RNA-LNP vaccines to elicit robust immunity against bacterial toxins and virulence factors.

This study provides evidence for the value of the mRNA-LNP vaccine platform for the treatment and prevention of CDI. It has been shown that clinically relevant doses of RNA-LNP vaccines elicit robust systemic and mucosal immunity against several *C. difficile* virulence factors simultaneously in multiple clinically relevant animal models. Through a combination of vaccine targets, it has been demonstrated that toxin-specific antibodies protect mice from lethal toxin challenges and CDI whereas immunization against vegetative cell and spore proteins reduce *C. difficile* colonization and promote decolonization during severe infection (Figure 19). This suggests that clearance depends on a vaccine-induced humoral immune response, and further experiments to optimize expression and immunogenicity of vaccine targets including PPEP-1 and to improve mucosal responses to vaccination will be conducted. This study

lays the groundwork for future investigation and provides a framework for the next generation of *C. difficile* RNA-based therapeutics.

Materials and Methods:

5

RNA design and production

Amino acid sequences of a TcdA and TcdB receptor binding domain (RBD) containing a combined repetitive oligopeptides (CROP) motifs were obtained from GenBank accession numbers P16154.2 and P18177.3 respectively. The amino acid sequences of PPEP-1 and CdeM were obtained from GenBank accession number Q183R7, and WP_009893169.1, respectively. Putative N-glycosylation sites were disrupted by substituting an asparagine residue at a predicted N-glycosylation site with a glutamine (N to Q) to prevent posttranslational epitopes masking in eucaryotic cells. Sequences underwent codon optimization and GC enrichment using a proprietary algorithm to improve expression and reduce potential immunogenicity of *in vitro* transcribed RNA. Codon optimized sequences were gene synthesized by Genscript with an optimized and modified IL-2 secretion signal, cloned into a proprietary *in vitro* transcription template containing an optimized T7 promoter, 3'UTR, 5'UTR and a 100-adenine tail. TcdA, TcdB and PPEP-1, and CdeM nucleoside modified RNA sequences were prepared using a MegaScript transcription kit (ThermoFisher Scientific), co-transcriptionally capped using a CleanCap™ system (TriLink Biotechnologies) and purified using a modified cellulose base chromatography method (M. Baierdörfer, et al., *Mol Ther Nucleic Acids*. 2019;(15):26-35), precipitated, eluted in nuclease free water, and quantified using the NanoDrop One system. Length and integrity were determined using an Agilent BioAnalyzer 2100 system. Endotoxin content was measured using a GenScript Toxisensor chromogenic assay, and values were below detection levels. RNA was frozen at -20°C until formulation.

Production and characterization of RNA-LNP vaccines

Hydrodynamic size, polydispersity index (PDI) and zeta potential of RNA-LNPs were measured using a Zetasizer Nano ZS90 (Malvern Instruments, Malvern,

UK). RNA encapsulation efficiency of LNP was determined using a modified Quant-iT RiboGreen RNA assay (Invitrogen). Endotoxin levels were determined using a Limulus Amebocyte Lysate (LAL) chromogenic assay found to be <0.5 endotoxin unit (EU)/mL.

5 *Phylogenomic analyses*

Genomic analyses were conducted on 137 *C. difficile* isolates sourced from the work of Bushman, Frederic D., et al (F. D. Bushman, et al., *Cell Host Microbe*. 2020;(28):422-433) and 30 pairs of raw sequencing reads in this study. Initial quality control was implemented with FastQC v0.12.1 and MultiQC v1.14 (P. Ewels, et al., *Bioinformatics*. 2016;(32):3047-3048). Subsequently, de novo assembly was performed on the quality-assessed raw reads utilizing Shovill v1.1.0 (github.com/tseemann/shovill). Genome assemblies were subjected to quality evaluation with CheckM v1.1.6 and BUSCO v5.4.7 (M. Manni, et al., *Mol Biol Evol*. 2021;(38):4647-4654; D. H. Parks, et al., *Genome Res*. 2015;(25):1043-1055). Genome annotations were then performed employing bakta v1.8.1 (O. Schwengers, et al., *Microb Genom*. 2021;(7)). The core-genome alignment was carried out using Roary v3.13.0 (A. J. Page, et al., *Bioinformatics*. 2015;(31):3691-3693), which subsequently served as the input for phylogenomic tree inference through RAxML-NG v1.1 (A. M. Kozlov, et al., *Bioinformatics*. 2019;(35):4453-4455) using the general time-reversible (GTR) substitution model (C. Lanave, et al. *J. Mol. Evol*. 1984;(20):86–93) accounting for among-site rate heterogeneity using the Γ distribution and four rate categories (Z. Yang, *J. Mol. Evol*. 1994;(39):306–314) for 10 individual searches with maximum parsimony random-addition starting trees and random topology starting tree. Node support was evaluated with 100 nonparametric bootstrap pseudoreplicates (J. Felsenstein, *Evolution*. 1985;(39):783–791). Phylogeny was rooted at the midpoint using the Tree Of Life (iTOL) v6.8 interactive tool (I. Letunic, P. Bork, *Nucleic Acids Res*. 2021;(49):W293-W296). For subsequent visualization and annotation of phylogeny, ggplot2, and ggtree were utilized (G. Yu, *Curr Protoc Bioinformatics*. 2020;(69):e96; G. Yu, et al., *Mol Biol Evol*. 2018;(35):3041-3043; G. Yu, *Data Integration, Manipulation and Visualization of Phylogenetic Trees*. CRC Press, 2022; S. Xu, et al., *iMeta*. 2022;(4):e56; G. Yu, et al., *Methods in Ecology and Evolution*. 2017;(8):28-36; H. Wickham, *ggplot2: Elegant*

Graphics for Data Analysis. Springer-Verlag New York. 2016).

Comparative Genomic Analysis of Protein Domains

Amino acid sequences for the TcdA CROP domain, TcdB CROP domain,
 5 PPEP-1, and CdeM (excluding the IL-2 secretion signal domain) were extracted from
 UniProt Knowledgebase with primary accession numbers of P16154 from position 1840
 to position 2710, P18177 from position 1851 to position 2366, Q183R7 from position 27
 to position 220, WP_009893169.1 from position 2 to position 164, respectively..
 Sequences served as seed sequences for constructing a protein-type BLAST database.
 10 Subsequently, BlastX was applied to execute a comparative analysis of amino acid
 sequences in the database against 137 genomes included in a phylogenetic assessment.
 The hits with the highest bit score from BlastX were utilized to determine the percentage
 of coverage and identity (S. F. Altschul, et al., *J Mol Biol*. 1990;(215):403-410; S. F.
 Altschul, et al., *Nucleic Acids Res*. 1997;(25):3389-3402; G. M. Boratyn, et al., *Biol*
 15 *Direct*. 2012;(7):12; G. M. Boratyn, et al., *BMC Bioinformatics*. 2019;(20):405; C.
 Camacho, et al., *BMC Bioinformatics*. 2023;(24):117; C. Camacho, et al., *BMC*
Bioinformatics. 2009;(10):421; A. Morgulis, et al., *Bioinformatics*. 2008;(24):1757-1764;
 Z. Zhang, et al., *J Comput Biol*. 2000;(7):203-214).

Identification of Toxigenic Strains via Comparative Genomic Analysis

Coding sequences for tcdA, tcdB, tcdC, tcdE, tcdR, cdu1, and cdd1 were
 sourced from the comprehensively annotated complete genome of *C. difficile* 630 (NCBI
 Reference Sequence: NC_009089.1). The nucleotide sequence for the PaLoc region was
 inferred and extracted by the locations of the flank genes cdu1 and cdd1. These
 25 nucleotide sequences served to construct an original nucleotide-specific BLAST
 database. To enhance the likelihood of detecting diverse toxin genes, additional
 sequences of toxin genes and the PaLoc region were incorporated into the database. The
 raw sequencing reads from 54 representative toxigenic strains were downloaded and
 assembled (K. E. Dingle, et al., *Genome Biol Evol*. 2014;(6):36-52). The initial round of
 30 blastn was conducted against the genomes of these 54 toxigenic strains using the original
 database, setting the parameters to accommodate more gaps and mismatches to boost the

coverage of hits. The sequences from the highest-scoring hits were subsequently extracted and used as seed sequences in a new BLAST database. Genomes of good quality were also used to infer and extract sequences from the PaLoc region using flanking genes *cdd1* and *cdu1*. A conserved non-coding region of 115 bp was employed as a marker for non-toxigenic strains (V. Braun, et al., *Gene*. 1996;(181):29-38). A second round of blastn was performed against the 137 genomes included in the previous phylogenetic analysis, using sequences from the new BLAST database. An arbitrary threshold of 80% coverage and 80% identity was applied to the highest-scoring hits and used as criteria for determining the presence or absence of genes. Strains were classified as non-toxigenic if they exhibited in the 115 bp conserved region and lacked toxin genes or the PaLoc region, and vice versa.

In vitro transfection of exemplary RNA LNPs and Western blot

Neuro2a cells were seeded in 12-well plates at a density of 0.2×10^6 cells per well, and transfected 24 hours post seeding at a confluence of around 80%. RNA-LNPs (1 mg/mL) were diluted (1:10) in PBS and added to each well at a final dose of 2.5 µg/well. Medium was aspirated, and cells were directly lysed 24 hours post transfection using the EZLys tissue protein extraction reagent (BioVision #8002-500) supplemented with 1X protease inhibitors (Sigma-Aldrich, #0493116001). Untreated cells were used as controls. Lysate was clarified at 15000 rpm for 5 minutes using a refrigerated centrifuge to pellet cell debris, and protein concentration in the supernatant was determined using the Pierce microBCA assay. A total of twelve (12) micrograms of cell lysate was loaded into a 4-15% precast polyacrylamide gel (Bio-Rad #4561083), transferred to a PVDF membrane using the ThermoFisher iBlot system (dry transfer). Membranes were blocked with 5% skimmed milk extract, washed and incubated for 2 hours with primary antibodies against TcdA and TcdB (Abcam 19953 and 252712) at a final concentration of 4 µg/mL. PPEP-1 was detected using a polyclonal serum from immunized mice (Boosted sera at 1:100 dilution, incubation 4 hours). Membranes were then washed three times using 1X TBST and incubated for 1 hour in the presence of an HRP conjugated donkey anti-goat IgG (Abcam 97040) to detect TcdA and TcdB. PPEP-1 was detected using an anti-mouse IgG (Abcam 97040). Proteins were visualized using

ECL detection reagent (Cytiva #RPN2209) on a GE ImageQuant™ system.

Cell viability measurements

Cells that are relevant for toxicity assessment and representative of
5 different tissues that come in to contact with the expressed immunogens encoded on the
RNA-LNPs were used for in vitro transfection and assessment of viability post-
transfection. HUVEC (CRL-1730), Caco-2 (HTB37), and human primary skeletal muscle
cells (PCS-950-010) were cultured as per ATCC using F12K Medium (ATCC 30-2004),
EMEM (ATCC 30-2003), and Mesenchymal Stem Cell Basal medium (ATCC PCS-500-
10 030) respectively. The muscle cells were cultured in the presence of the Primary Skeletal
Muscle Growth Kit (ATCC PCS-950-040). Cells were seeded in 96-well plates at a
density of 10,000 cells per well and transfected with 1 or 3 µg TcdA or TcdB RNA-LNP.
Full-length recombinant TcdA (8619-GT-020) or TcdB (6246-GT-020), recombinant 5
Luciferase RNA-LNP (Luc LNP), and untreated cells were used as controls. After
15 incubation for 24 or 48 hours, 50 µL XTT cell proliferation reagent (Roche
11465015001) was added to each well, and plates were incubated for an additional 5
hours before measuring absorbance at 492 nm using the Thermofisher Varioskan plate
reader. Cell viability was calculated relative to the untreated cells.

20 *Generation of recombinant TcdA, TcdB, PPEP-1, and CdeM proteins*

A region encoding the TcdA and TcdB receptor binding proteins
(including CROP domain), the full length PPEP-1 coding sequence, and the full length
CdeM coding sequence was codon optimized for expression in *E. coli*, gene synthesized
at Genscript, and cloned into the pET30a vector. An N terminal 6XHIS tag followed by
25 the TEV recognition/cleavage site was introduced after the start (ATG) codon of all three
constructs to allow for purification using affinity chromatography. *E. coli* BL21 (DE3)
form NEB was transformed, and positive clones were used to inoculate 1 L cultures.
Recombinant proteins from the supernatant were concentrated, and purified using the NI-
IDA column on an AKTA Avant 150 system. Purified proteins were cleaved with the
30 TEV protease buffer exchanged into PBS and 5% sucrose (pH 7.4), filter sterilized,
quantified using the micro-BCA assay (Pierce), aliquoted and stored at -80C. Aliquots

were tested for precipitation following for multiple freeze and thaw cycles and the purity was determined using densitometric analysis of a Coomassie blue stained SDS-PAGE gel under reducing conditions. All proteins displayed more than 90% purity and had the predicted molecular weights using western blot.

5

Spleen and lymph node harvests

Spleens were collected, processed as single cells, filtered using 70 μ m cell strainers in complete RPMI 1640, centrifuged, and red blood cells lysed in ACK lysis buffer to obtain a clear single cell suspension. Splenocytes were resuspended in 1 mL complete RPMI 1640 media, counted, and used immediately for studies. Draining lymph nodes (inguinal and popliteal) were collected and processed as described above (without ACK). Cells from the dLNs were resuspended in 250 μ L RPMI 1640 media, counted, and used immediately.

15 *Production of fluorescently labeled proteins recombinant proteins for antigen specific B cells.*

Fluorescently labeled recombinant TcdA, TcdB and PPEP-1 protein were prepared using the Lightning-Link R-Phycoerythrin (R-PE) and Lightning-Link (R) Rapid Alexa Fluor 647 conjugation reagents (Novus Biologicals, 703-0010 and 336-0005). TcdA, TcdB and PPEP-1 were diluted to 0.5 mg/mL in PBS and 50 μ g reacted in the presence of 1:10 (v/v) of LL modifier for 3 hours at room temperature. The labeling reaction was stopped in the presence of 1:10 (v/v) LL quencher for 30 minutes and stored at 4°C until use.

25 *Flow cytometry analysis of T and B cells*

Tfh cells: 2 million cells were stained with anti-mouse CD16/32 antibody for 20 minutes and stained with an anti-CXCR5- biotin for 30 minutes on ice, washed twice and incubated with Streptavidin BV-421 in the presence of surface antibodies (Table 3) for 30 minutes, washed, fixed and permeabilized with the FoxP3/Transcription Factor Staining Kit (eBioScience) according to the manufacturer instructions and stained for Bcl6. Following intracellular staining, cells were washed twice, fixed in 300 μ L 1%

paraformaldehyde for acquisition. The gating strategy, as well as the antibody list and catalog numbers are provided (Table 5; Figure 4A).

Table 5: Antibodies used for immunophenotyping of Tfh cells

Marker	Fluorophore	Dilution (1:x)	Clone	Company	Cat#:
TCR beta chain	BV421	200	H57-597	Biolegend	109230
B220	AF488	200	RA3-6B2	Biolegend	103225
CXCR5	PE	100	L138D7	Biolegend	145504
CD4	PE/Dazzle 594	400	RM4-5	Biolegend	100565
CD8a	PECy5	400	53-6.7	BD Biosciences	553034
CD62L	AF700	100	MEL-14	eBioscience	56-0621-82
PD-1	PECy7	200	RMP1-30	Biolegend	109110

5

Germinal Center (GC) and Memory B cells: 2 million cells per sample incubated with antimouse CD16/32 antibody for 20 min at 4°C. Cells were then washed with FACS buffer (2% FBS in PBS) and stained for 1 h using antibodies (Table 3). Following staining, cells were washed twice, and fixed in 300 µL 1% paraformaldehyde for acquisition. The gating strategy, as well as the antibody list, fluorescent TcdA and TcdB RBD probes, and catalog numbers are provided (Table 3, Figure 4B-D).

10 T cells: 2 million splenocytes were stimulated with 2.5 µg/mL of TcdA CROP, or TcdB CROP, or PPEP-1 peptide pools (15 mers, 4 amino acid overlapping peptide pool) in a FACS tube for 6 hours at 37°C, 5% CO₂ with 2 mg/mL anti-CD28 (Tonbo, 40-0281-M001) providing costimulation. Stimulations proceeded for 1 hour before adding 5 mg/mL brefeldin A (Biolegend, 420601), 2 mM monensin (Biolegend, 420701), and 5 mg/mL anti-CD107a (Biolegend, 121610) Alexa Fluor 647 for 5 hours. DMSO served as a negative control and the combination of 50 mg/mL phorbol 12-myristate 13-acetate and 1 mg/mL ionomycin served as a positive control. After a total of

15

6 hours, samples were washed with PBS, stained with Live/Dead Aqua for 5 minutes, blocked using anti-mouse CD16/32 antibody for 20 minutes, and stained extracellularly for 30 minutes using antibodies (Table 4). Cells were washed in FACS buffer, fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences, 554714), and stained
5 intracellularly using antibodies for 30 min (Table 4). Following intracellular staining, cells were washed twice, and fixed in 300 μ L 1% paraformaldehyde for acquisition. The gating strategy, as well as the antibody list and catalog numbers are provided (Table 4; Figure 13A).

All samples were acquired on a BD LSR II equipped with 4 laser lines and
10 18 PMTs and data were analyzed in FlowJo v10.

In vivo toxin neutralization assays

To determine the LD100 of intraperitoneally injected TcdA and TcdB, recombinant toxins were prepared at doses of 5, 10, 25, 50, 100, and 125ng in 200 μ L of
15 sterile PBS. Naïve 7-week-old C57BL/6J mice were randomly assigned into groups of 5. Recombinant proteins were injected i.p. and mice were monitored for behavior, body condition, and mortality every 4 hours post-injection. Doses were selected for use in subsequent studies by their ability to kill 100% of animals within 4 hours after injection. To determine neutralization capacity of vaccine-elicited immune response, mice were
20 immunized twice i.m. with 1 or 5 μ g of TcdA or TcdB monovalent RNA-LNP or monovalent recombinant protein vaccine with or without alum adjuvant (details on vaccine preparation below). Two weeks after last immunization, immunized mice were either challenged i.p. with five times the previously determined LD100 (625ng rTcdA, 125ng rTcdB) and monitored for behavior, body condition, and mortality every 4 hours
25 post-injection or terminally bled. Serum from immunized mice was diluted 1:20 in PBS and incubated with rTcdA or rTcdB LD100 (150 or 25 ng respectively) for 1 hour at 37 $^{\circ}$ C. Naïve mice were then challenged i.p. with recombinant toxin alone, monovalent immunized serum + recombinant toxin, or control (unvaccinated) serum + recombinant toxin and monitored for behavior, body condition, and mortality every 4 hours
30 postinjection.

Mouse immunization and C. difficile infection

5-week-old male and female C57BL/6J mice (Jackson Laboratories, strain no. 000664) were used in most studies and BALB/c (Jackson Laboratories, strain no. 000651) were used where indicated. All mice were maintained under specific pathogen-free conditions at either the Children's Hospital of Philadelphia or the University of Pennsylvania. Multivalent RNA-LNPs (1-5 µg) were mixed at 1:1 w/w ratio, diluted up to 50µL in PBS, and administered intra-muscularly into the hind leg within two hours of thaw. Recombinant protein vaccines were mixed 1:1 with alhydrogel (aluminum hydroxide 2% w/v). 28 days later, mice were either euthanized for analysis of vaccine-induced humoral and cellular immune responses after single immunization or they were boosted with the same dose. 14 days later, mice were either euthanized for analysis of vaccine-induced immune responses after second immunization or they were infected with *C. difficile* as previously described (J. Zackular et al., *Nat. Med.* 2016;(22):1330–1334). Antibiotic treatment was administered by providing 0.5g/L cefoperazone in their drinking water ad libitum for 5 days, followed by a 2-day recovery period before *C. difficile* infection via oral gavage. Two different *C. difficile* strains were utilized where indicated: VPI 10463 at a dose of 1×10^6 spores and CD196 at a dose of 1×10^5 spores. Mice were monitored daily for survival and weight loss, and mice were euthanized when weight loss exceeded 20% of their original body weight. For studies to analyze immune responses vaccination after *C. difficile* infection, 5-week-old C57BL/6J mice were infected with 1×10^5 spores *C. difficile* CD196 as described above. 14 days post-infection, mice were immunized with 1µg RNA-LNPs. 28 days later, mice received a second immunization. 14 days later, serum and feces were collected for analysis of vaccine-induced humoral immune responses. For long-term studies, C57BL/6J mice were immunized as described above. 40 days after last immunization, mice were either euthanized for analysis of vaccine-induced immune responses or infected with *C. difficile* as described above. For re-challenge studies, C57BL/6J mice were re-infected with *C. difficile* 200+ days after primary infection as described above.

Hamster studies

Syrian hamsters (HsdHan: AURA, Envigo, catalog no. 8901M) were

maintained at University of Texas Medical Branch. 4-5-week-old male golden Syrian hamsters were pre-bled and vaccinated intra-muscularly with trivalent RNA-LNP vaccine (1:1:1 w/w ratio), trivalent recombinant protein + alhydrogel, or empty lipid nanoparticles at week 0 and 3. Body weight and clinical signs were collected at days 0, 1, 3, 5, 7, and 14 post-prime and at day 1, 2, 3, 7, and 14 post-boost. Feces were collected 14 days after each immunization. Serum was collected at days 0, 21, 28, and 45 post-prime. Hamsters were sacrificed as per university protocols and observed for gross abnormalities.

10 *Detection of antibodies in mouse serum and feces*

96-well High Bind Stripwell™ Corning 96 Well Clear Polystyrene Microplates plates were coated overnight with 1 µg/mL of purified TcdA (CROP), TcdB (CROP), PPEP-1, or CdeM. Plates were washed once with wash buffer (0.5% Tween-20 in PBS) and blocked for two hours at room temperature using a solution of heat inactivated, IgG depleted, protease free bovine serum albumin (2% w/v BSA in PBS). After blocking, plates were washed three times with wash buffer. Feces were homogenized in 1 mL PBS and centrifuged for 10 minutes at 10,000 x g. Sera or fecal supernatants were serially diluted in the blocking solution, added to plates, and incubated for two hours at room temperature. Plates were washed three times before the addition of horseradish peroxidase-conjugated anti-mouse secondary antibody specific to total mouse IgG (1:10,000) and IgA (1:5,000), anti-rhesus IgG (1:20,000), or anti-hamster IgG (1:8,000) in blocking buffer. Plates were incubated for 1.5 hours, washed three times before the addition of tetramethylbenzidine (TMB) substrate solution. The reaction was stopped by adding 2 N sulfuric acid, and the absorbance was measured at 450 nm using a SpectraMax™ 190 microplate reader. Antigen-specific antibody end-point dilution titer was defined as the highest dilution of serum or feces to give an OD greater than the cut-off OD value determined using the Frey Method (A. Frey, et al., *J Immunol Methods*. 1998;(221):35-41). Fecal antibody titers were normalized to gram of feces.

30 *C. difficile* enumeration, and toxin titers

C. difficile burdens were quantified by collecting fecal samples at

indicated timepoints and plating on taurocholate cycloserine cefoxitin fructose agar (TCCFA). *C. difficile* toxin titers were quantified in the stool using a previously described Vero cell cytotoxicity assay (J. P. Zackular, et al., *Nat Med.* 2016;(22):1330-1334). Briefly, fecal samples were homogenized in sterile PBS, pelleted, and supernatant was filtered through a 0.2µm filter. Supernatant was diluted along a tenfold series and incubated overnight with Vero monolayers. Toxin titers in stool were calculated as the reciprocal value of the highest dilution that rounded 100% of the cells and normalized per gram of feces.

10 *Histology and scoring*

Samples were fixed in 10% neutral buffered formalin (NBF) prefilled HistoTainer™ II, dehydrated in graded ethanol series, cleared with xylene and embedded in paraffin. Sections (5 µm) were collected on Superfrost™ Plus stain slides (Fisher Scientific, Ottawa, ON, Canada), and stained with Hematoxylin and Eosin. Slides were scanned using a NanoZoomer digital slide scanner (Hamamatsu, Boston, MA, USA) and visualized using the NDP® view 2.0 software (Hamamatsu, Boston, MA, USA). Sections were scored in a blind manner by a pathologist based on previously described criteria (C. M. Theriot et al., *Gut Microbes.* 2011;(2):326–334). Histological scores were reported as a cumulative score of three independent criteria: inflammation, edema, and epithelial cell damage.

DNA extraction from Stool

Mice were co-housed for one week before experimental manipulation. Stool samples were collected at indicated timepoints for microbiota analysis. Microbial genomic DNA was extracted using DNeasy Power Soil Kit (Qiagen).

16S rRNA gene library prep

Barcoded PCR primers annealing to the V4 region of the 16S rRNA gene were used for library generation. PCR reactions were carried out in duplicate using Q5 High-Fidelity DNA Polymerase (NEB, Ipswich, MA). Each PCR reaction contained 0.5µM of each primer, 0.34 U Q5 Pol, 1X Buffer, 0.2 mM dNTPs, and 5.0µl DNA in a

total volume of 50µl. Cycling conditions were as follows: 1 cycle of 98°C for 1 minute; 20 cycles of 98°C for 10 seconds, 56°C for 20 seconds, and 72°C for 20 seconds; and 1 cycle of 72°C for 8 minutes. After amplification, duplicate PCR reactions were pooled and then purified using a 1:1 volume of SPRI beads. DNA in each sample was then
5 quantified using PicoGreen and pooled in equal molar amounts. The resulting library was sequenced on the Illumina MiSeq using 2x250 bp chemistry. Extraction blanks and 9 DNA-free water were subjected to the same amplification and purification procedure to allow for empirical assessment of environmental and reagent contamination. Positive controls, consisting of five artificial 16S gene fragments synthesized in gene blocks and
10 combined in known abundances, were also included.

Bioinformatics processing and statistical analysis

Sequence data were processed using QIIME2 (E. Bolyen et al., *Nat. Biotechnol.* 2019;(37):852–857). Read pairs were processed to identify amplicon
15 sequence variants with DADA2 (B. Callahan et al., *Nat. Methods.* 2016;(13):581–583). Taxonomic assignments were generated by comparison to the Silva reference database version 132 (C. Quast et al., *Nucleic Acids Res.* 2013;(41):D590–D596), using the naïve Bayes classifier implemented in scikit-bio (N. Bokulich et al., *Microbiome.* 2018;(6):90). A phylogenetic tree was inferred from the sequence data using MAFFT (K. Katoh et al.,
20 *Mol. Biol. Evol.* 2013;(30):772–780). Similarity between samples were assessed by weighted and unweighted UniFrac distance (C. Lozupone et al., *Appl. Environ. Microbiol.* 2005;(71):8228–8235; C. Lozupone et al., *Appl. Environ. Microbiol.* 2007;(73):1576–1585). Data files from QIIME were analyzed in R environment for statistical computing. Linear mixed effects models were used at each time point to
25 estimate the mean difference between study groups. Cage information was added to the models as the random effect to account for the coprophagic nature of mice. Relative abundances of bacteria were log10 transformed and modeled as the outcome. Only the bacteria with at least 1% mean relative abundance across samples were tested. Community-level differences between sample groups were assessed using the
30 PERMANOVA test (M. Abt., *Nat. Rev. Microbiol.* 2016;(14):609–620; M. Anderson, *Austral Ecol.* 2001;(26):32–46). When multiple tests were done, p-values were corrected

for false discovery rate using Benjamini-Hochberg method.

Non-human primate studies

An 18-year-old male rhesus macaque (N = 1) was injected i.m. in the right
5 deltoid with 200 µg of tetravalent RNA-LNP vaccine (TcdA, TcdB, PEPP-1, and CdeM)
in 250µL total volume (1:1:1:1 w/w or 50µg/immunogen). A second age-matched male
rhesus monkey (N = 1) was i.m. injected in the right deltoid with 200 µg of an eDHFR-
tagged SARS-CoV-2 S2P RNA-LNP vaccine. NHPs received two immunizations
separated by 21 days, and blood was collected at day 1, 8, 21, and 35 post injection.
10 Plasma and PBMCs were separated and stored until use.

Statistical analysis

Data are presented as mean ± SD or mean ± SEM where indicated. Graph
schematics and illustrations were created using iTOL v6.8 (I. Letunic et al., *Nucleic Acids*
15 *Res.* 2021;(49):W293–W296), SPICE 6 (M. Roederer et al., *Cytometry A*.
2011(79A):167-174) and Graphpad prism V10.0.2

Example 3: Maternal antibody transmission after RNA-LNP vaccination.

20 8 weeks old female BL6 mice were immunized once with a dose of 1µg of
bivalent TcdA and TcdB CROP RNA-LNP vaccines two weeks before breeding. Mice
were left to breed and transferred into a new cage once pregnancies were confirmed.

Neonates were challenged with *C.difficile* VPI10463 at day 8 after birth
(Figure 26 and Figure 27). Matching controls were not challenged and used as controls
25 (NT). The total stomach content or feces from Pups (neonates) were collected 2 weeks
after birth (day 14) and analyzed using an Endpoint ELISA. The stomach content is
considered the best correlate to measure antibodies from the dam milk.

Figure 25 depicts the anti-toxin IgG and IgA titers measured on 8 weeks
old female BL6 mice immunized with of 1µg of bi-valent TcdA and TcdB CROP RNA-
30 LNP vaccines two weeks before breeding. Serum samples were collected at 14 days post
delivery of the first litter.

Figure 26 and Figure 27 depict Anti-toxin IgG and IgA titers in stomach content and feces respectively measured on pups 2 weeks after birth. The dotted line shows the antibody titers in serum of dams at the day of stomach content and feces collection.

5 Figure 28 depicts Anti-toxin IgG and IgA titers measured on serum from control animals (dams) and in the stomach and feces of pups from control animals. As expected, no antibodies were detected in any of the samples.

 These example experiments show that maternal antibodies are transferred to pups as demonstrated by the anti-toxin specific antibodies in stomach contents and
10 feces. Pups from unvaccinated mice did not have anti-toxin specific antibodies.

 The disclosures of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entirety. While this has been disclosed with reference to specific embodiments, it is apparent that
15 some embodiments and variations of this may be devised by others skilled in the art without departing from the true spirit and scope of the. The appended claims are intended to be construed to include all such embodiments and equivalent variations.

CLAIMS

What is claimed is:

1. A composition comprising at least one RNA molecule encoding at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, or CspC, a fragment or variant thereof, or any combination thereof.
2. The composition of claim 1, wherein the composition comprises a combination of RNA molecules encoding at least two *C. difficile* antigens, wherein the composition comprises at least two of TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, or fragments or variants thereof.
3. The composition of claim 2, comprising a combination of RNA molecules encoding TcdA and TcdB, or fragments or variants thereof.
4. The composition of claim 2, comprising a combination of RNA molecules encoding TcdA, TcdB, and PPEP-1, or fragments or variants thereof.
5. The composition of claim 1, wherein the composition comprises a fragment of TcdA or TcdB comprising a receptor binding domain (RBD) including a combined repetitive oligopeptide (CROPs) domain.
6. The composition of claim 1, wherein the composition comprises an RNA molecule encoding a variant of at least one *C. difficile* antigen comprising an amino acid substitution which disrupts an N-glycosylation site.
7. The composition of claim 1, wherein the RNA molecule comprises a nucleic acid sequence encoding an amino acid sequence having at least 80% identity to an amino acid sequence of:
 - a) a TcdA antigen comprising an amino acid sequence of SEQ ID NO:1, SEQ ID

NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) a TcdB antigen comprising an amino acid sequence of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) a PPEP-1 antigen comprising an amino acid sequence of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) a CdeM antigen comprising an amino acid sequence of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) a CWP84 antigen comprising an amino acid sequence of SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) a CWP66 antigen comprising an amino acid sequence of SEQ ID NO:28, or a fragment or variant thereof;

g) a ZupT antigen comprising an amino acid sequence of SEQ ID NO:29, or a fragment or variant thereof; or

h) a CspC antigen comprising an amino acid sequence of SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

8. The composition of claim 1, comprising at least one of:

a) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, or a fragment or variant thereof, encoding a TcdA antigen;

b) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, or SEQ ID NO:55, or a fragment or variant thereof, encoding a TcdB antigen;

c) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62,

SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, or SEQ ID NO:66, or a fragment or variant thereof, encoding a PPEP-1 antigen;

d) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, or SEQ ID NO:74, or a fragment or variant thereof, encoding a CdeM antigen;

e) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77 or SEQ ID NO:78, or a fragment or variant thereof, encoding a CWP84 antigen;

f) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:79 or SEQ ID NO:80, or a fragment or variant thereof encoding a CWP66 antigen;

g) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:81 or SEQ ID NO:82, or a fragment or variant thereof, encoding a ZupT antigen; or

h) an RNA molecule comprising a nucleotide sequence of SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87 or SEQ ID NO:88, encoding a CspC antigen.

9. The composition of claim 1, comprising a combination of RNA molecules, wherein the combination of RNA molecules comprises at least one selected from:

a) an RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) an RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) an RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) an RNA molecule encoding SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) an RNA molecule encoding SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) an RNA molecule encoding SEQ ID NO:28, or a fragment or variant thereof;

g) an RNA molecule encoding SEQ ID NO:29, or a fragment or variant thereof;
and

h) an RNA molecule encoding SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

10. The composition of claim 9, comprising a combination of RNA molecules encoding at least two amino acid sequences selected from:

a) an RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) an RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) an RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) an RNA molecule encoding SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) an RNA molecule encoding SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) an RNA molecule encoding SEQ ID NO:28, or a fragment or variant thereof;

g) an RNA molecule encoding SEQ ID NO:29, or a fragment or variant thereof;
and

h) an RNA molecule encoding SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

11. The composition of claim 9, comprising a combination of RNA molecules encoding at least three amino acid sequences selected from:

a) an RNA molecule encoding SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) an RNA molecule encoding SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant

thereof;

c) an RNA molecule encoding SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) an RNA molecule encoding SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) an RNA molecule encoding SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) an RNA molecule encoding SEQ ID NO:28, or a fragment or variant thereof;

g) an RNA molecule encoding SEQ ID NO:29, or a fragment or variant thereof;

and

h) an RNA molecule encoding SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

12. The composition of any one of claims 1 to 11 wherein the RNA molecule is mRNA, self-replicating RNA, self-amplifying RNA, or circular RNA.

13. The composition of any one of claims 1 to 12, wherein at least one RNA molecule is a nucleoside modified RNA molecule comprising at least one modified nucleoside.

14. The composition of claim 13, wherein at least one modified nucleoside comprises pseudouridine, 1-methyl pseudouridine, 2'-O-methylpseudouridine, 5-methyl-2'-O-methylpseudouridine, 3-methylpseudouridine, 5-methyl-uridine, methoxyuridine, dihydrouridine, 5,2'-O-dimethyluridine, 4-thiouridine, 5-methyl-2-thiouridine, 2-thio-2'-O-methyluridine, (3-(3-amino-3-carboxypropyl)uridine); 5-hydroxyuridine; 5-methoxyuridine; uridine 5-oxyacetic acid; uridine 5-oxyacetic acid methyl ester; 5-(carboxyhydroxymethyl)uridine; 5-(carboxyhydroxymethyl)uridine methyl ester; 5-methoxycarbonylmethyluridine; 5-methoxycarbonylmethyl-2'-O-methyluridine; 5-methoxycarbonylmethyl-2-thiouridine; 5-aminomethyl-2-thiouridine; 5-methylaminomethyluridine; 5-methylaminomethyl-2-thiouridine; 5-methylaminomethyl-2-selenouridine; 5-carbamoylmethyluridine; 5-carbamoylmethyl-2'-O-methyluridine; 5-

carboxymethylaminomethyluridine; 5-carboxymethylaminomethyl-2'-O-methyluridine; or 5-carboxymethylaminomethyl-2-thiouridine.

15. The composition of any one of claims 1-14, wherein the RNA molecule comprises a nucleic acid sequence encoding a leader sequence, a tag or signal peptide.

16. The composition of any one of claims 1-15, wherein the RNA molecule further comprises a linker.

17. The composition of any one of claims 1-16, wherein the composition comprises one or more lipid nanoparticle (LNP), wherein the one or more RNA molecules are partially or fully encapsulated within the one or more LNP.

18. A pharmaceutical composition comprising the composition of any one of claims 1-17 and at least one pharmaceutically acceptable excipient.

19. A combination comprising:

(i) a first pharmaceutical composition comprising a first RNA molecule, wherein the first RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof; and

(ii) a second pharmaceutical composition comprising a second RNA molecule, wherein the second RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof,

wherein the first RNA molecule and the second RNA molecule are not the same, and wherein the RNA molecules are partially or fully encapsulated within one or more LNPs.

20. A combination comprising:

(i) a first pharmaceutical composition comprising a first RNA molecule, wherein the first RNA molecule encodes at least one *C. difficile* antigen, wherein the

at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof;

(ii) a second pharmaceutical composition comprising a second RNA molecule, wherein the second RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof; and

(iii) a third pharmaceutical composition comprising a third RNA molecule, wherein the third RNA molecule encodes at least one *C. difficile* antigen, wherein the at least one *C. difficile* antigen comprises TcdA, TcdB, PPEP-1, CdeM, CWP84, CWP66, ZupT, CspC, a fragment thereof, or a variant thereof;

wherein the first RNA molecule, the second RNA molecule, and the third RNA molecule are not the same, and wherein the RNA molecules are partially or fully encapsulated within one or more LNPs.

21. A method of inducing an immune response against at least one strain of *C. difficile* infection in a subject, comprising administering a composition according to any one of claims 1-17, or a combination according to claim 19 or 20, or a pharmaceutical composition according to claim 18.

22. A method of treating or preventing a disease or disorder associated with *C. difficile* infection in a subject, comprising administering a composition according to any one of claims 1-17, or a combination according to claim 19 or 20 or a pharmaceutical composition according to claim 18.

23. The method of claim 21 or 22, wherein the composition is administered by a delivery route selected from the group consisting of intradermal, subcutaneous, inhalation, intranasal, and intramuscular.

24. The method of any one of claims 21-23, wherein the method comprises a single administration of the composition.

25. The method of any one of claims 21-23, wherein the method comprises multiple administrations of the composition.

26. The method of claim 21, wherein the composition induces a broad immune response against multiple strains of *C. difficile* in a cell, tissue or subject.

27. The method of claim 21, wherein the composition induces a protective immune response in the subject.

28. The method of claim 21, wherein the composition challenges and/or hinders *C. difficile* colonization in a subject.

29. The method of claim 21, wherein the disease or disorder is colitis or diarrhea.

30. The method of claim 21, wherein administration does not affect intestinal microbiota in the subject.

31. The composition of any one of claims 1-17 for use in the treatment or prevention of a *C. difficile* infection in a subject comprising administering one or more doses of the composition to a subject.

32. The pharmaceutical composition of claim 18 for use in the treatment or prevention of a *C. difficile* infection in a subject comprising administering one or more doses of the composition to a subject.

33. The combination of claim 19 or 20 for use in the treatment or prevention of a *C. difficile* infection in a subject comprising administering one or more doses of the composition to a subject.

34. The composition of any one of claims 1-16, or the combination of any one

of claims 19 or 20, wherein the composition further comprises a delivery vehicle.

35. The composition of claim 34, wherein the delivery vehicle is selected from poly(ethylenimine) PEI, chitosan, Janus dendrimers, dendrimers, lipopolymers, and a one component system.

36. The composition of any one of claims 1-18, or the combination of any one of claims 19 or 20, wherein the composition further comprises an adjuvant.

37. The composition of claim 36, wherein the adjuvant is selected from the group consisting of alum, ASO3, squalene, liposomes, chitosan, Matrix M, a cytokine, a nucleic acid molecule encoding a cytokine, or any combination thereof.

38. The composition of claim 37, wherein the cytokine is IL-12, IL-21, or APRIL.

39. The composition of any one of claims 1-18, or the combination of any one of claims 19 or 20, wherein the sequence encoding the antigen is operably linked to nucleotide sequence encoding an oligomerization domain.

40. The composition of claim 39, wherein the oligomerization domain comprises ferritin, or a fragment or variant thereof.

41. A composition comprising a fusion molecule comprising a fusion of a *C. difficile* antigen to an oligomerization domain.

42. The composition of claim 41, wherein the oligomerization domain comprises ferritin, or a fragment or variant thereof.

43. The composition of claim 41, wherein the *C. difficile* antigen comprises an amino acid sequence having at least 80% identity to an amino acid sequence of:

a) a TcdA antigen comprising an amino acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7, or a fragment or variant thereof;

b) a TcdB antigen comprising an amino acid sequence of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 or SEQ ID NO:14, or a fragment or variant thereof;

c) a PPEP-1 antigen comprising an amino acid sequence of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:21, or a fragment or variant thereof;

d) a CdeM antigen comprising an amino acid sequence of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:25, or a fragment or variant thereof;

e) a CWP84 antigen comprising an amino acid sequence of SEQ ID NO:26 or SEQ ID NO:27, or a fragment or variant thereof;

f) a CWP66 antigen comprising an amino acid sequence of SEQ ID NO:28, or a fragment or variant thereof;

g) a ZupT antigen comprising an amino acid sequence of SEQ ID NO:29, or a fragment or variant thereof; or

h) a CspC antigen comprising an amino acid sequence of SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32, or a fragment or variant thereof.

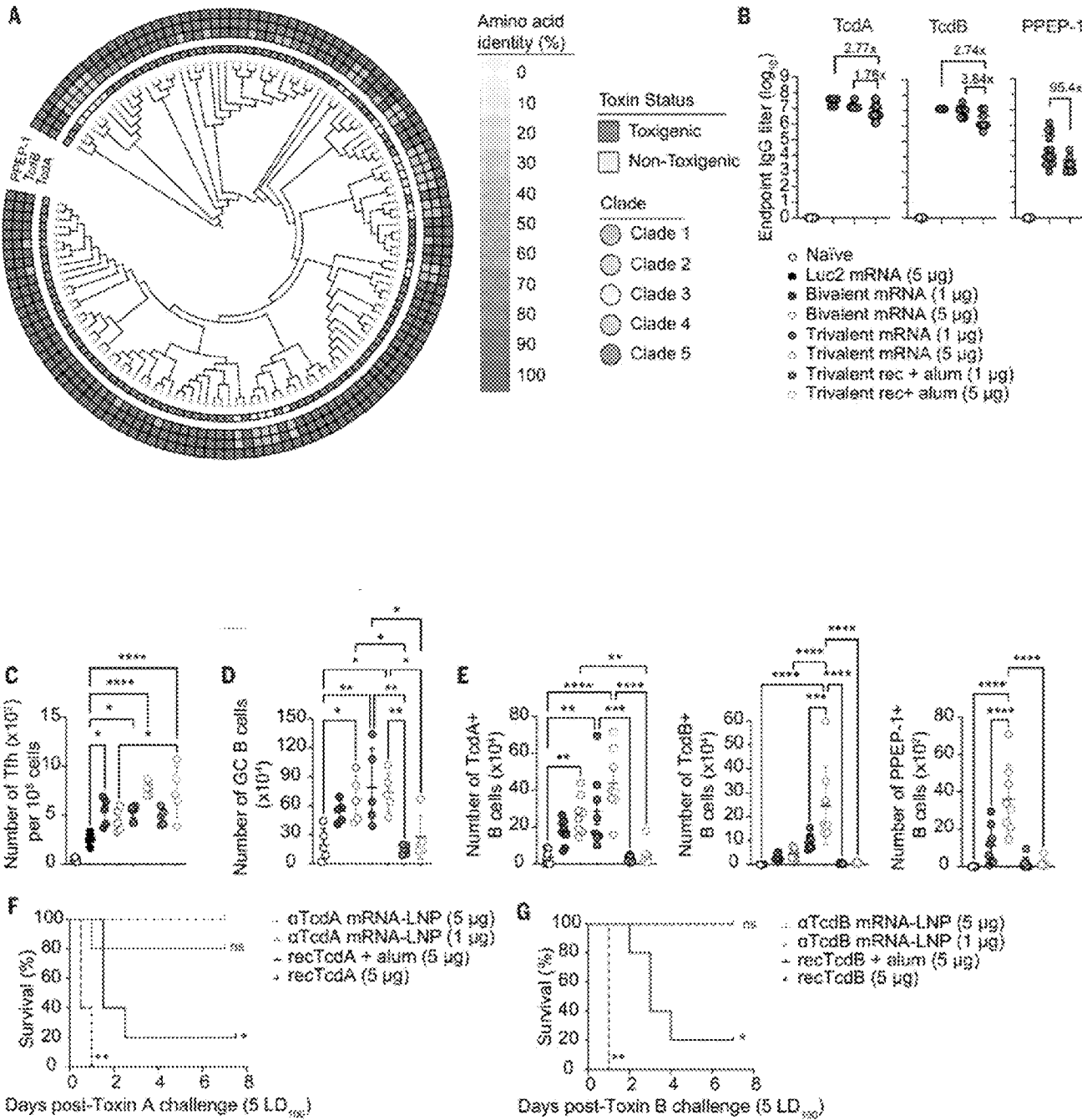
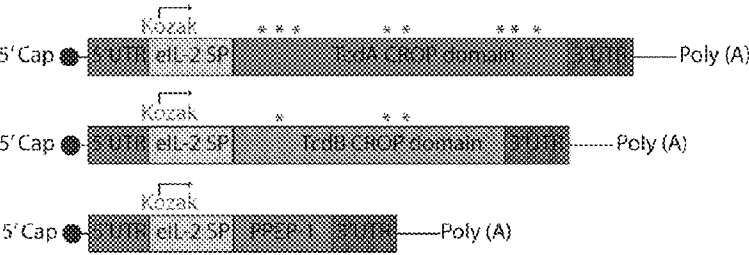


Figure 1

H



I

Composition	Molar ratio (%)	N/P ratio
Ionizable lipid	48	Ionizable lipid to mRNA ratio of 6
DSPC	10	
Cholesterol	38.5	
PEG-lipid	1.5	

J

mRNA-LNP	Batch	Size (nm)	Surface charge (mV)	EE (%)	RNA integrity (DS)	RNA integrity (DP)
TcdA	1	80	-4	95	>90%	>90%
TcdA	2	82	-5	98	>90%	>90%
TcdB	1	85	-4	95	>90%	>90%
TcdB	2	84	-3	95	>90%	>90%
PPEP-1	1	78	-5	97	>90%	>90%
PPEP-1	2	75	-5	96	>90%	>90%
CdcM	1	82	-3	97	>90%	>90%

K

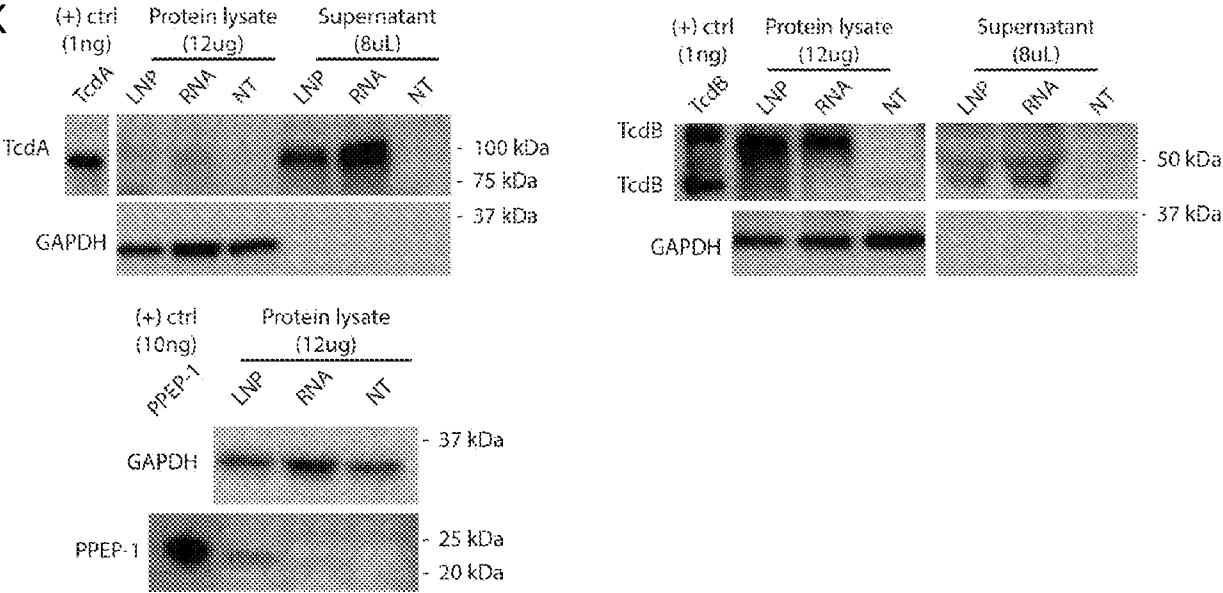


Figure 1 cont.

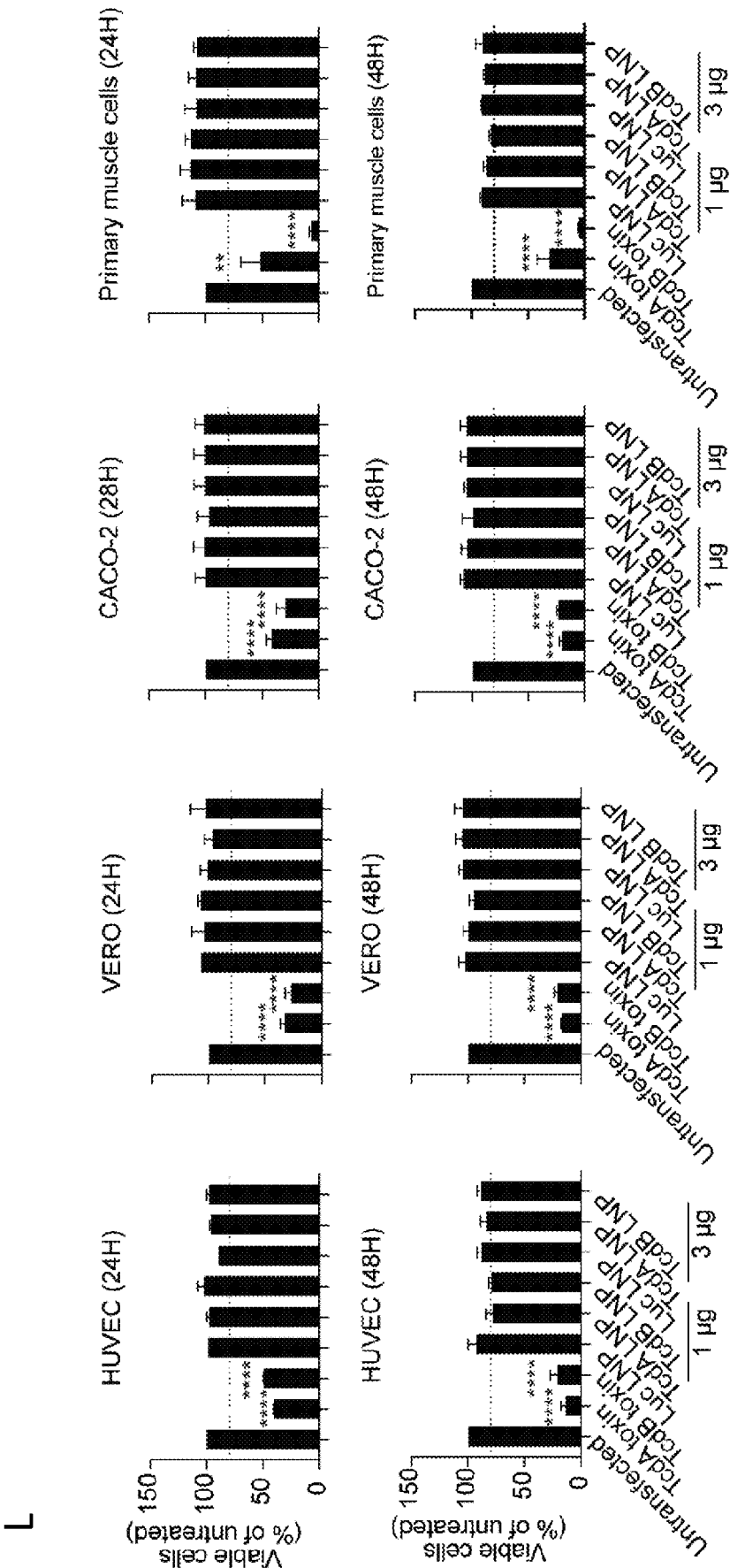


Figure 1 cont.

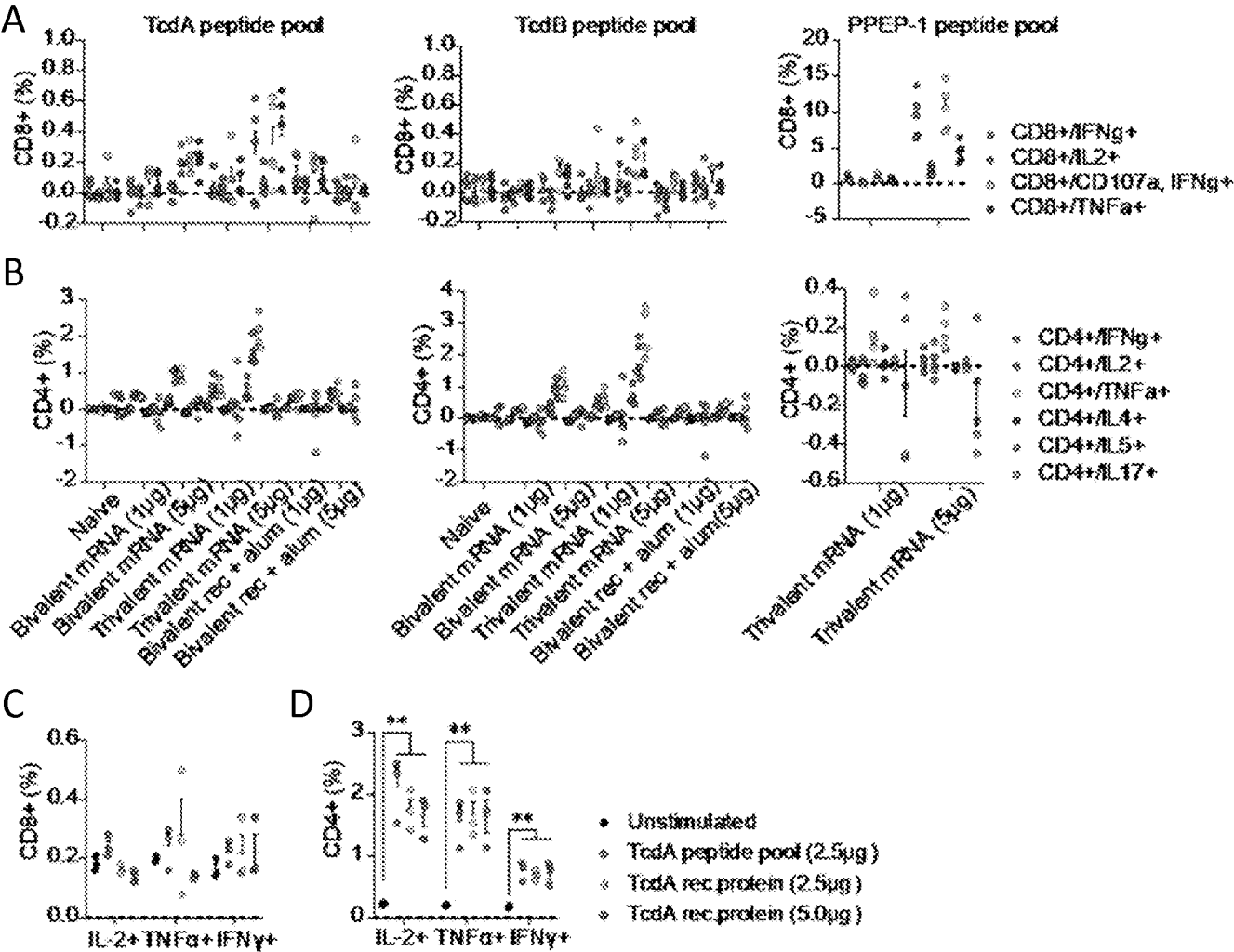


Figure 2

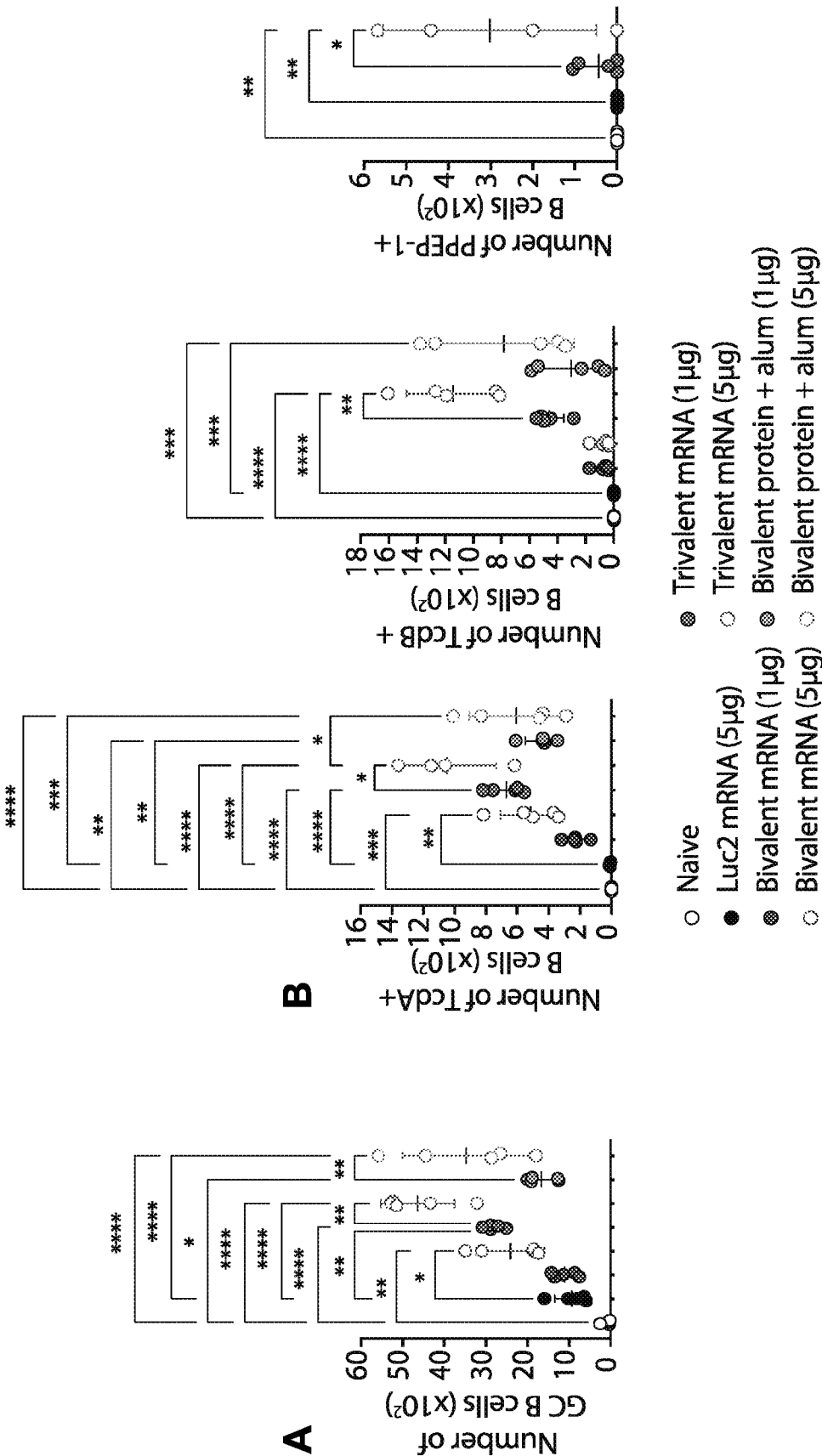


Figure 3

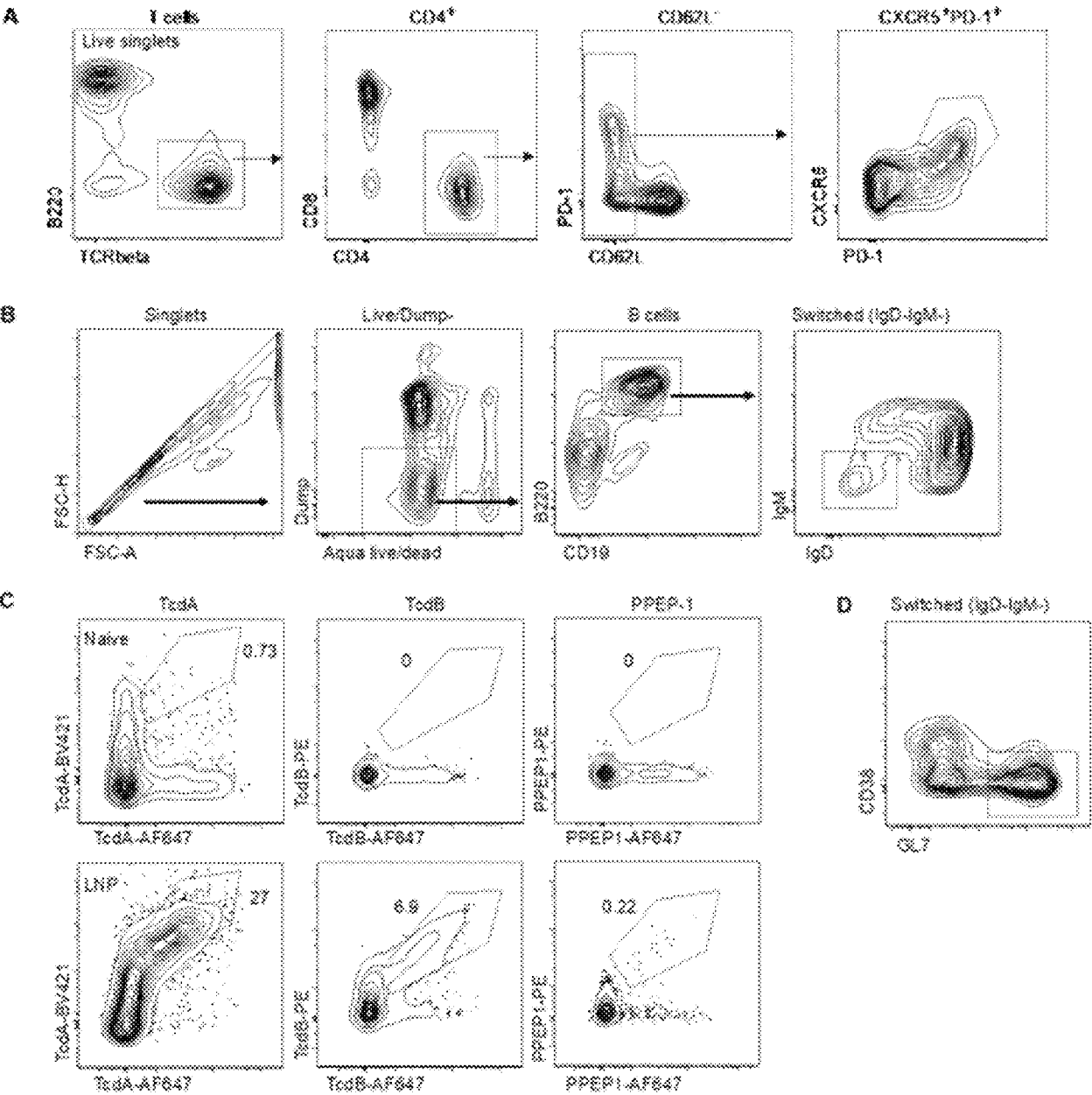


Figure 4

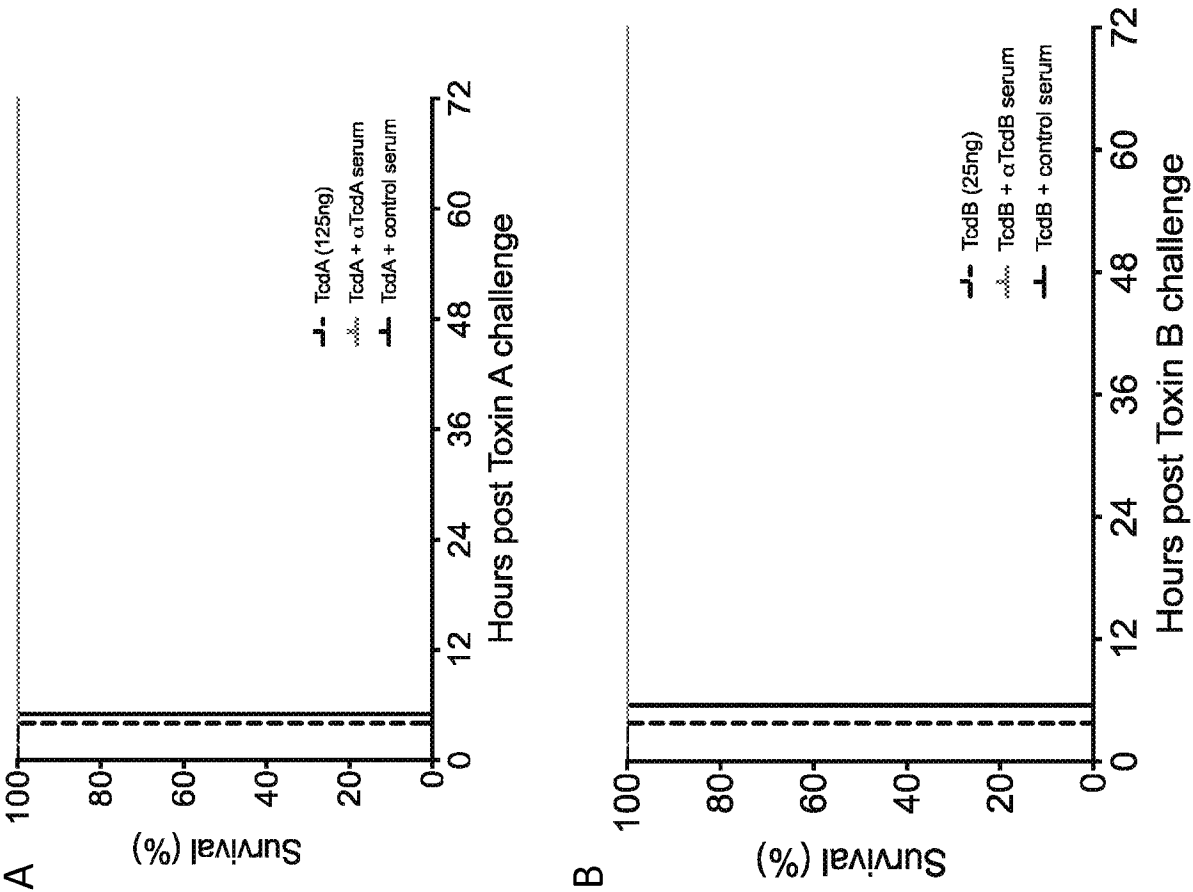


Figure 5

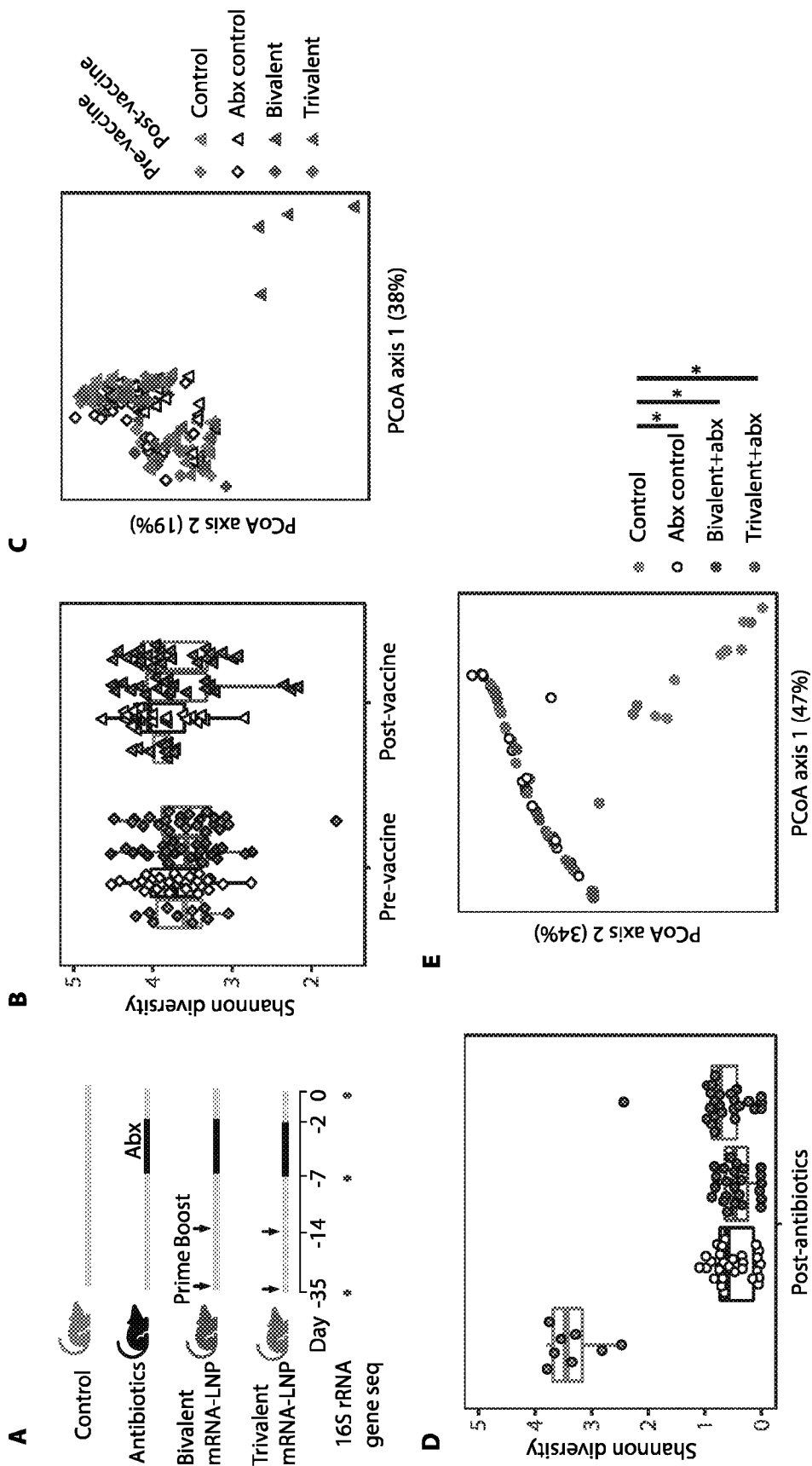


Figure 6

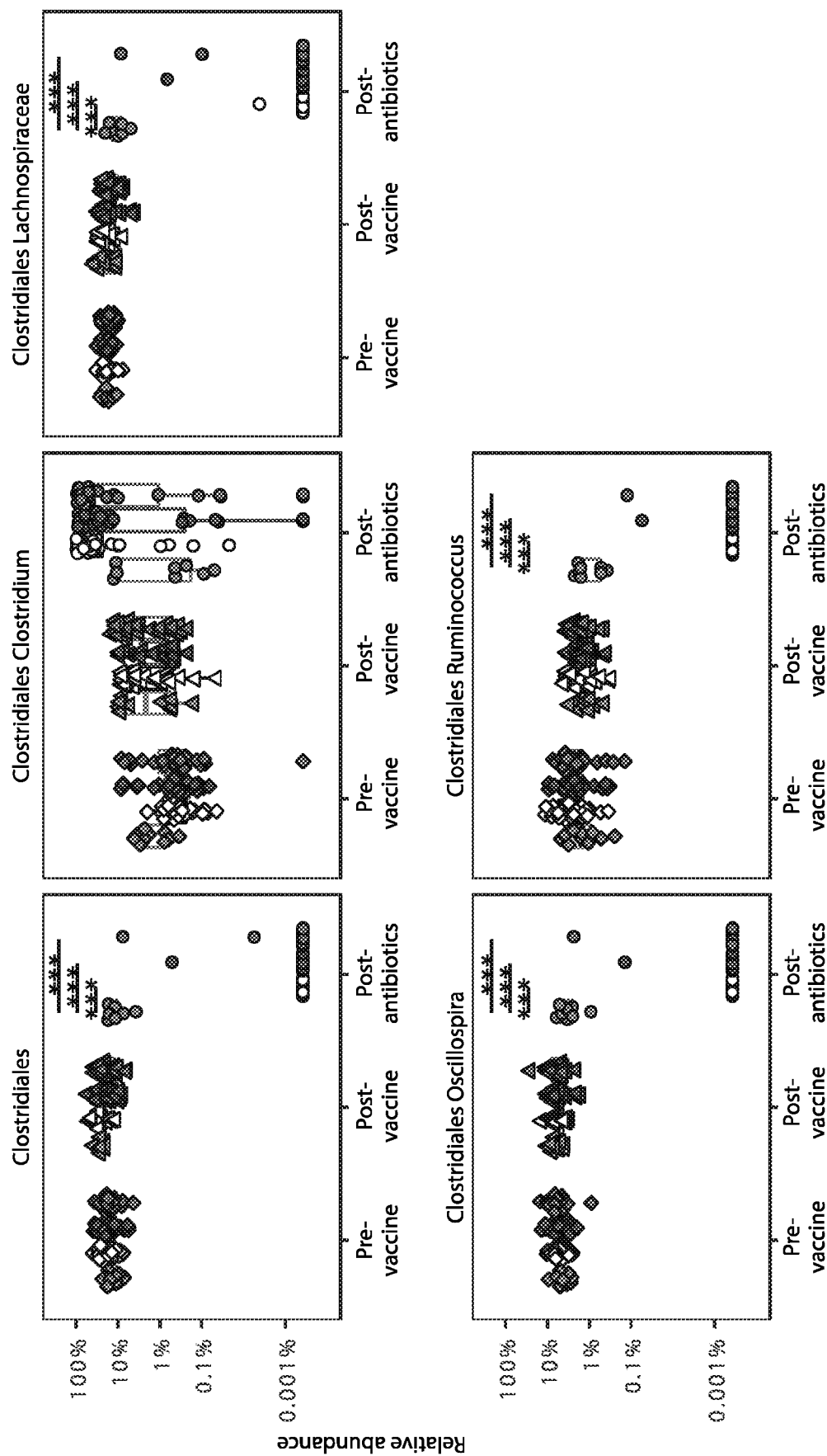


Figure 7

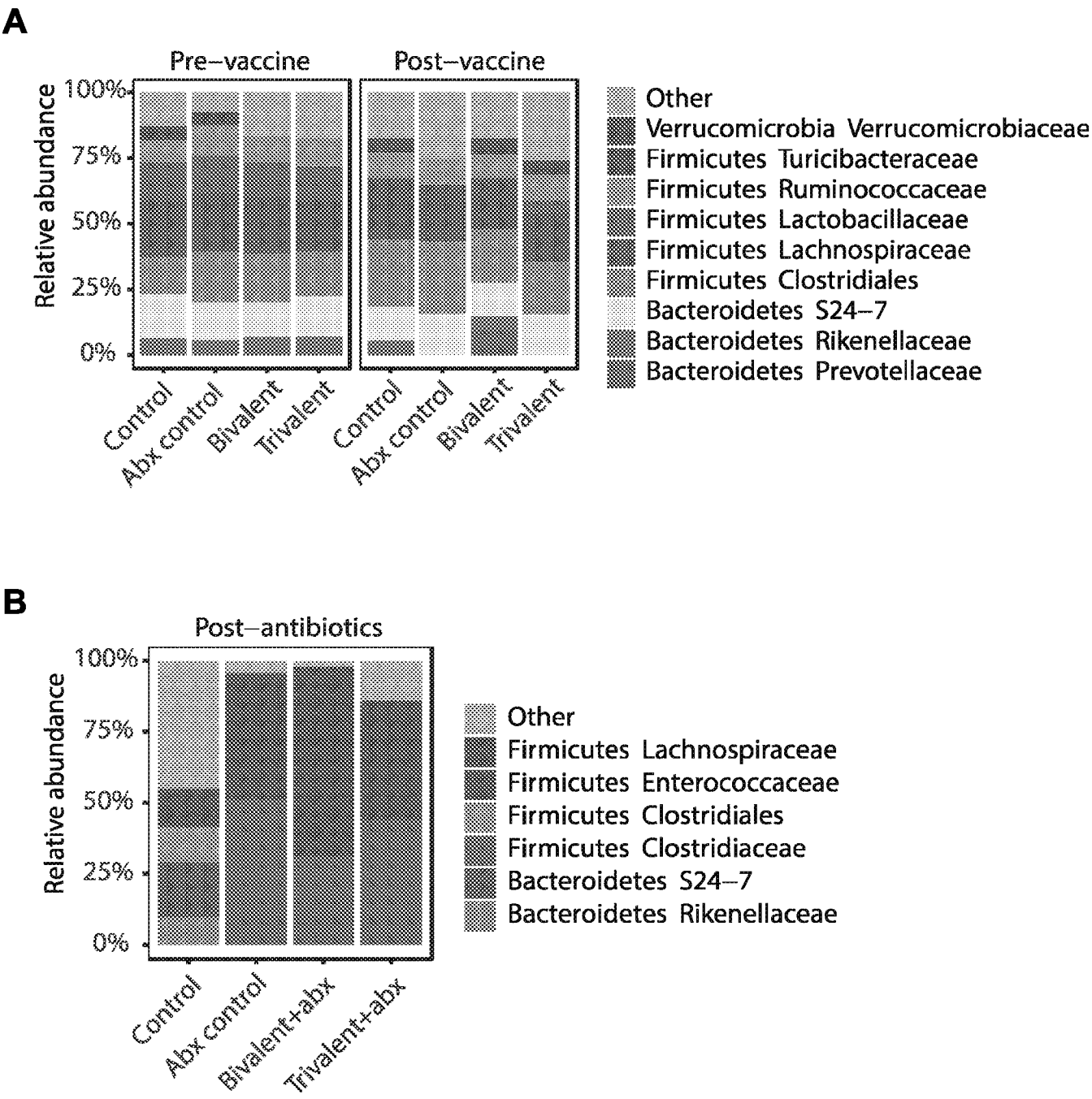


Figure 8

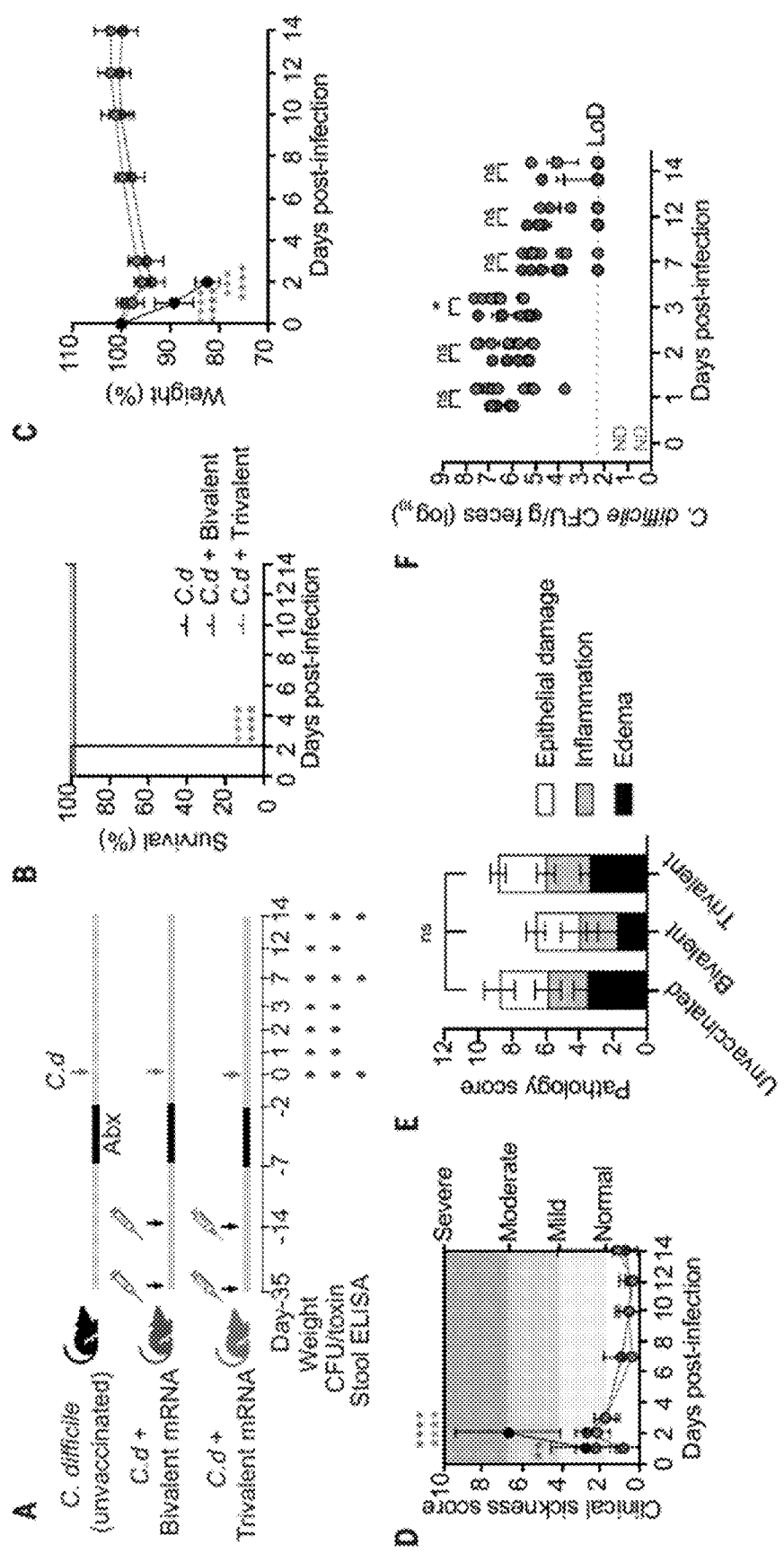


Figure 9

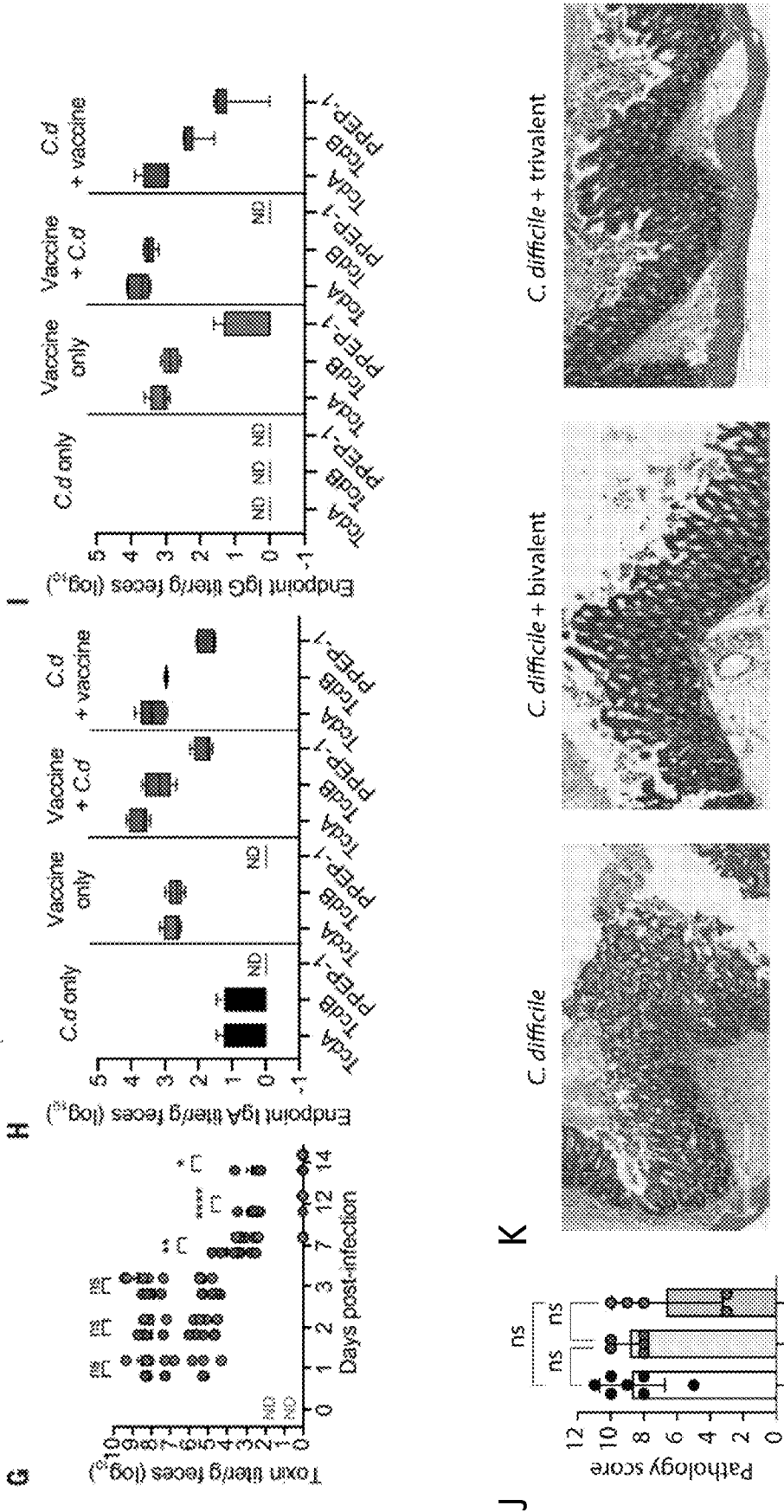


Figure 9 cont.

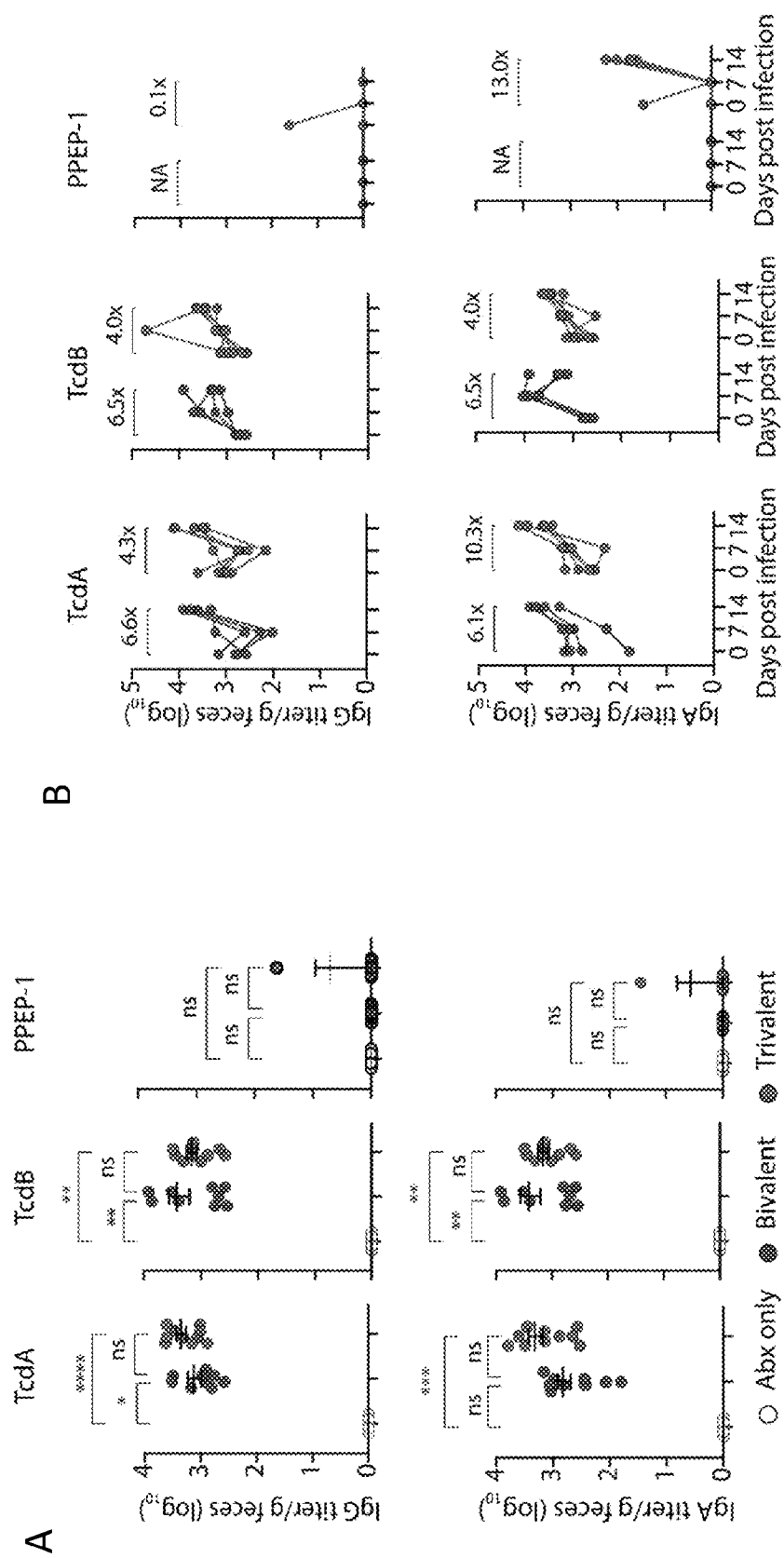


Figure 10

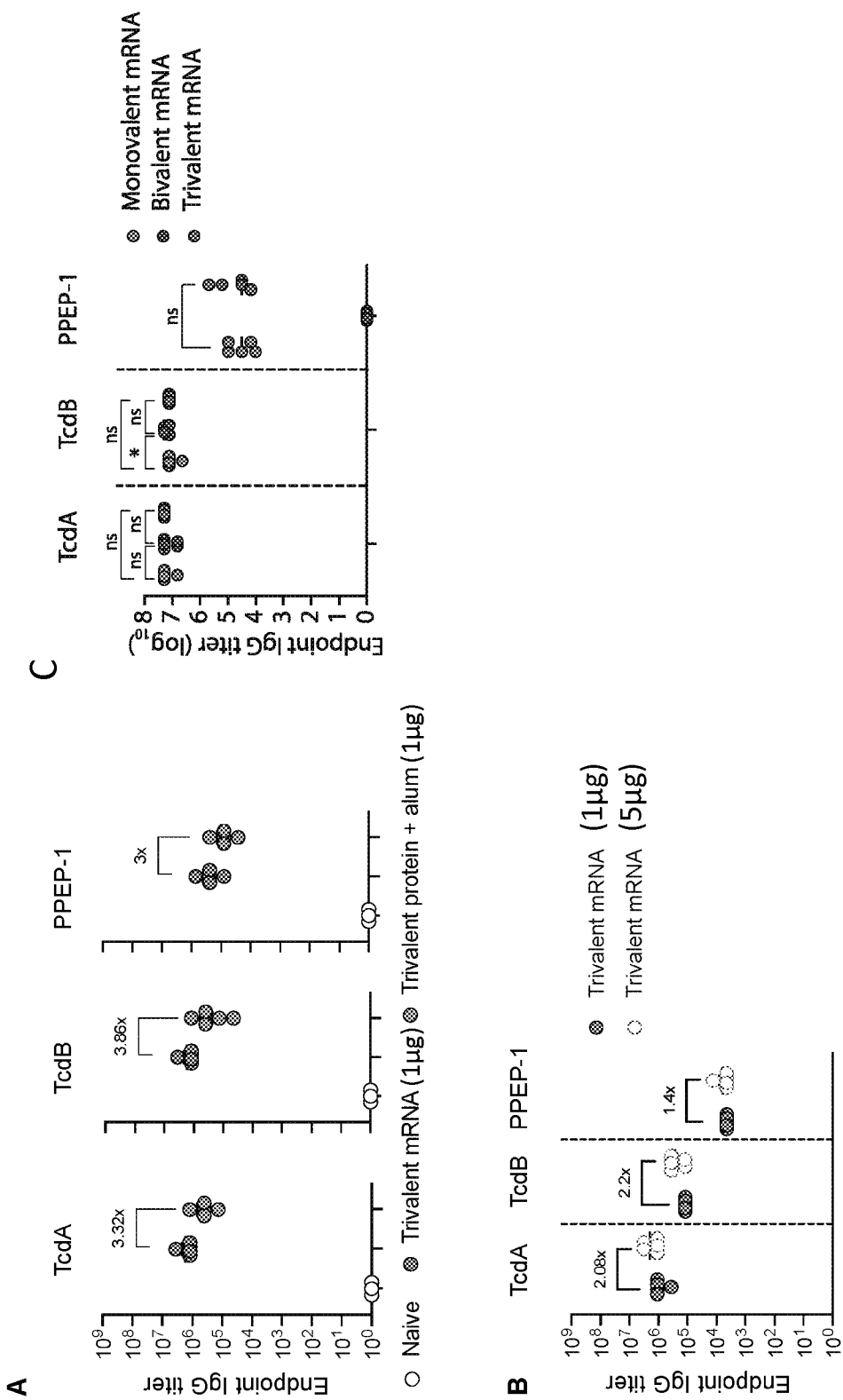


Figure 11

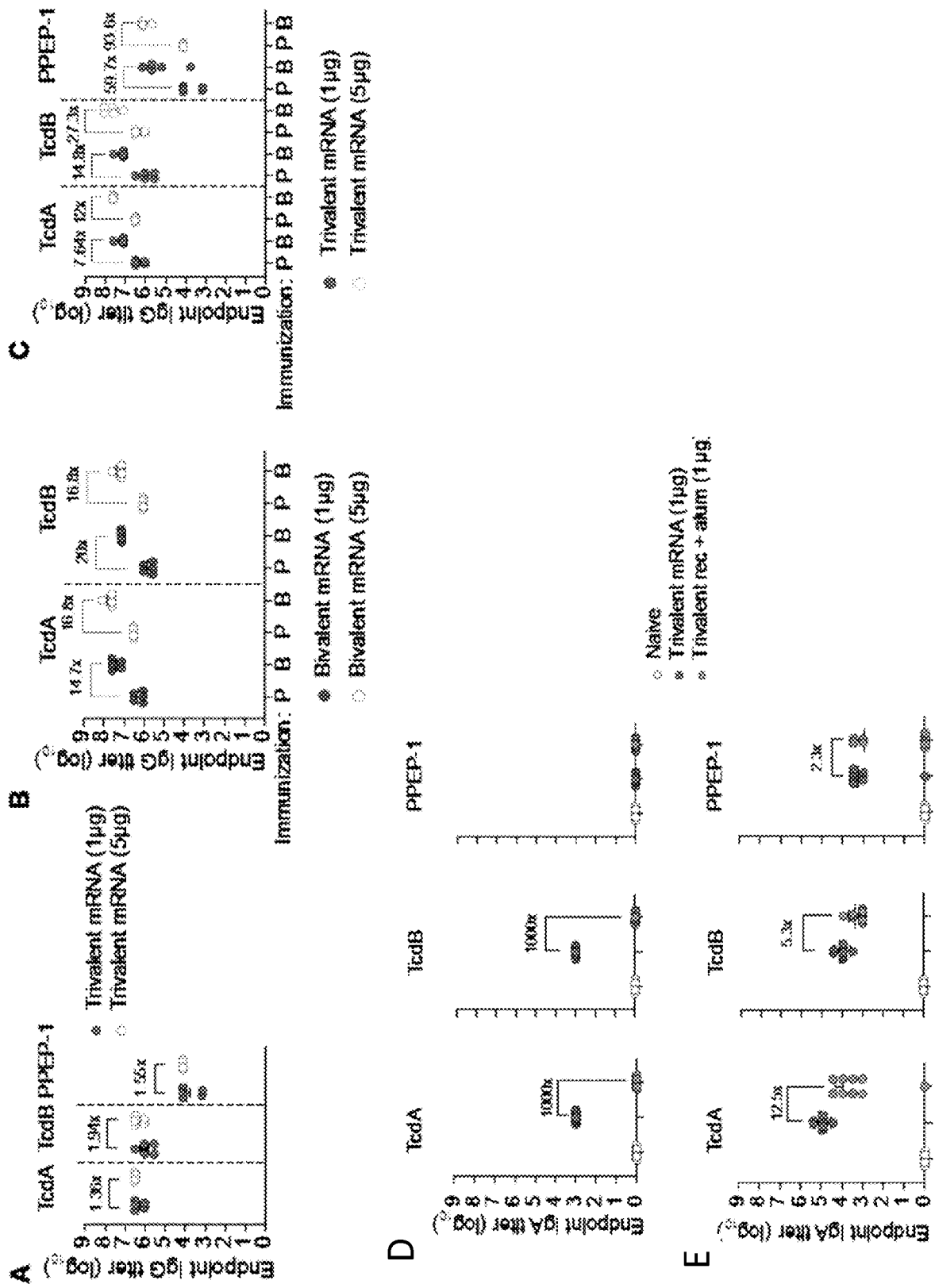


Figure 12

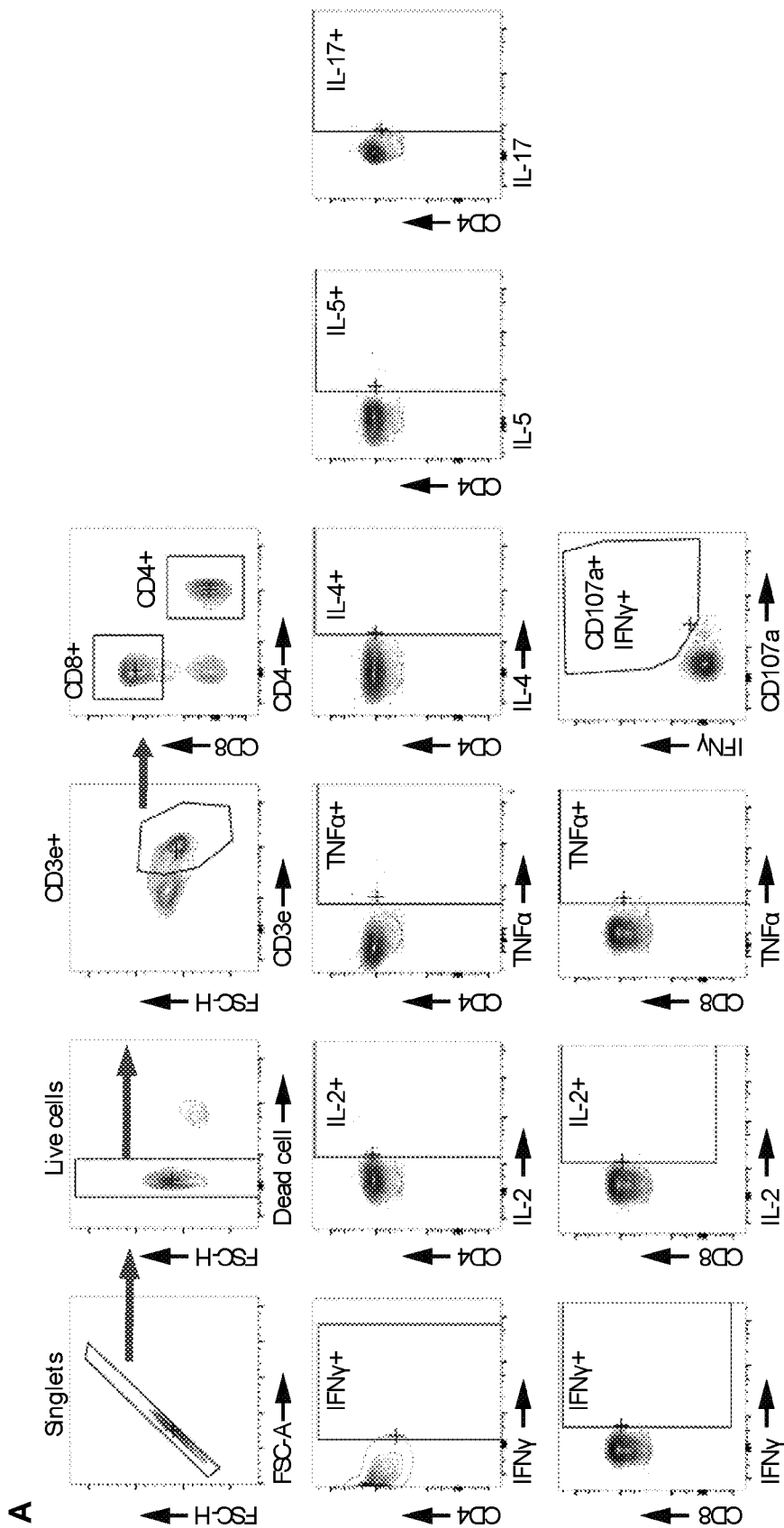


Figure 13

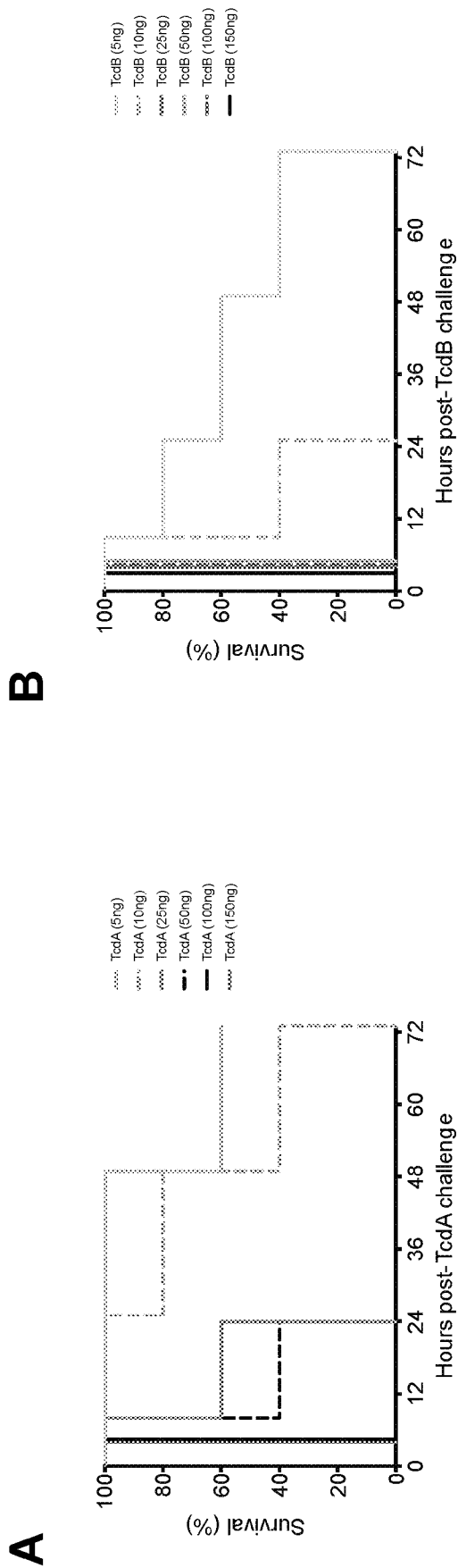


Figure 14

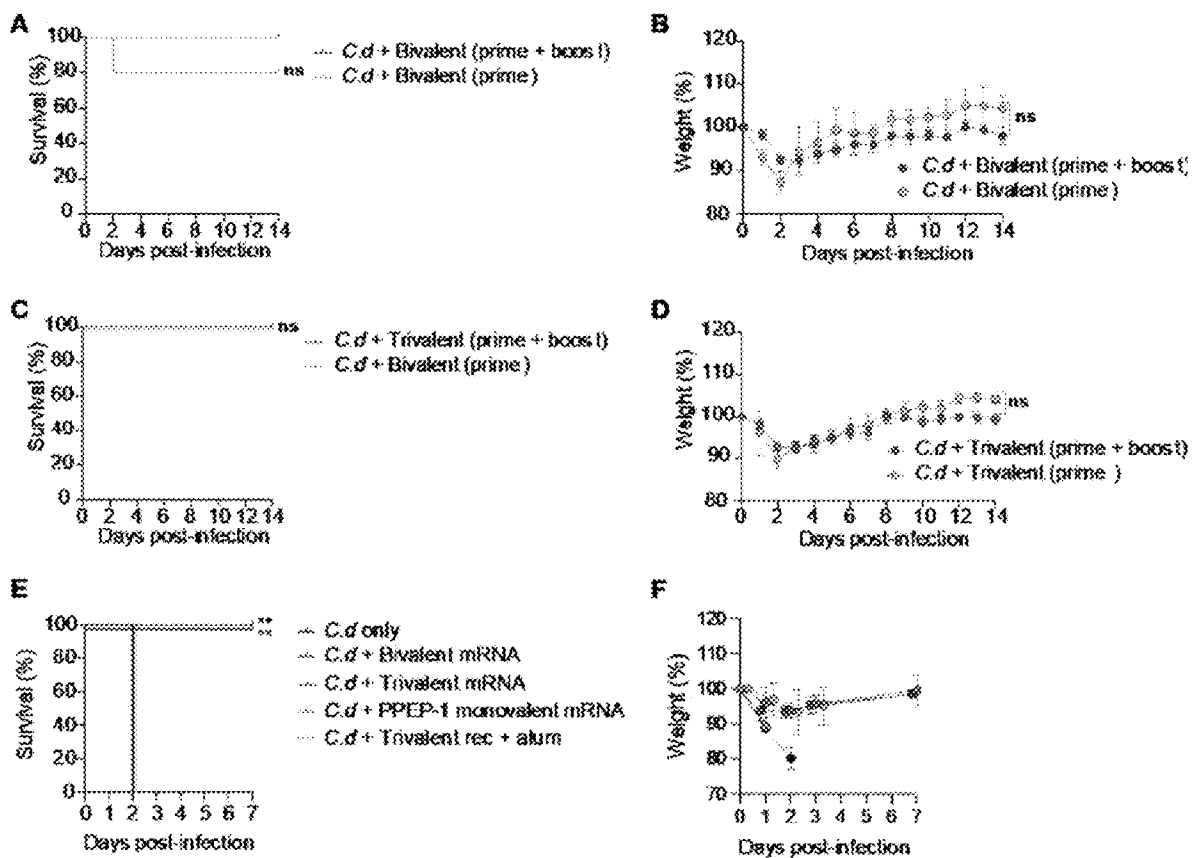


Figure 15

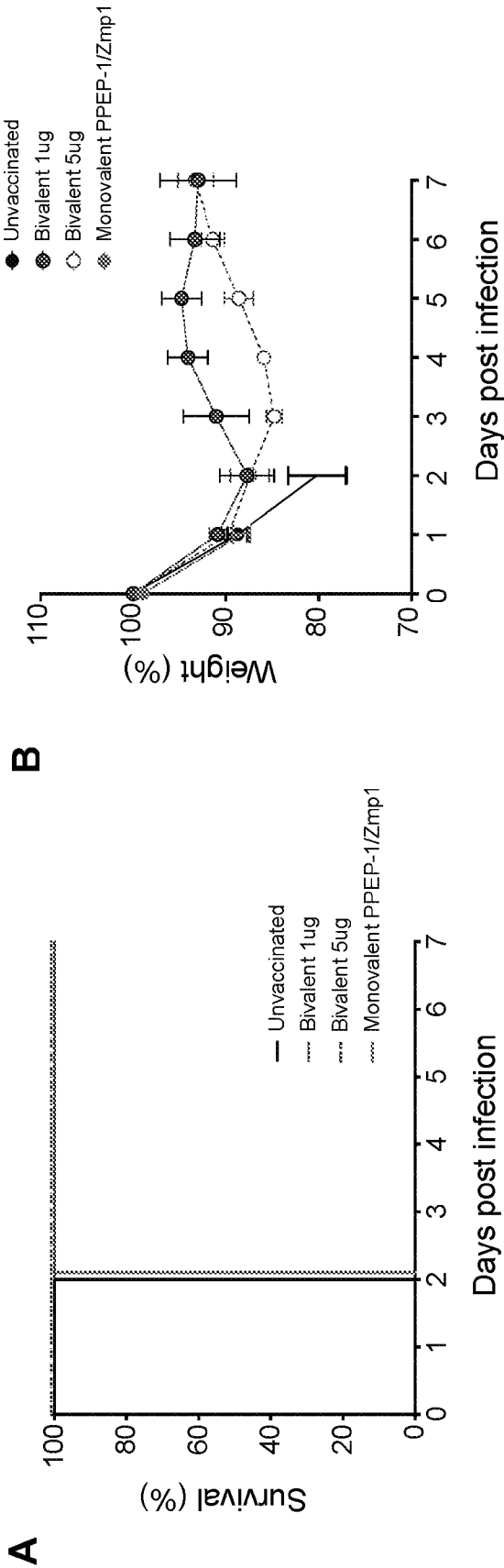


Figure 16

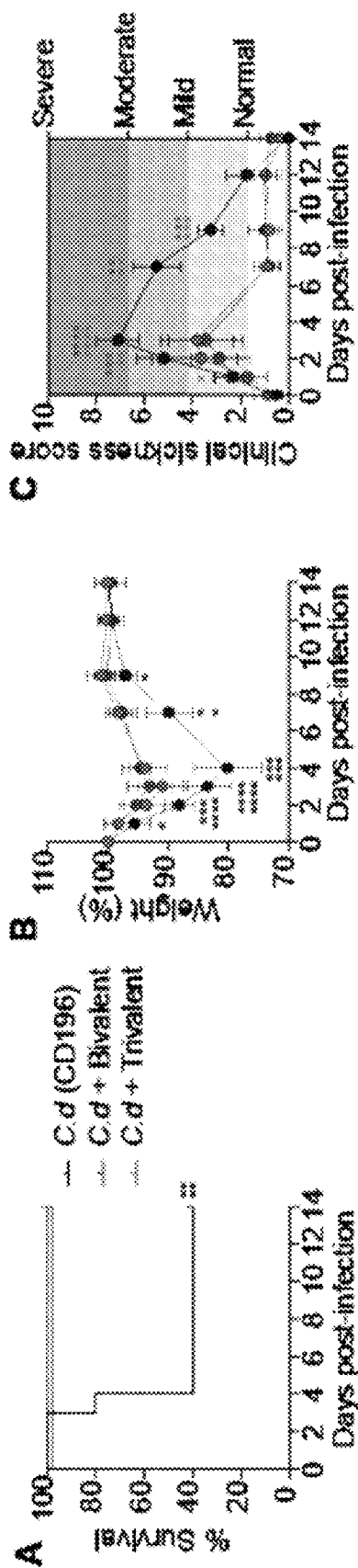


Figure 17

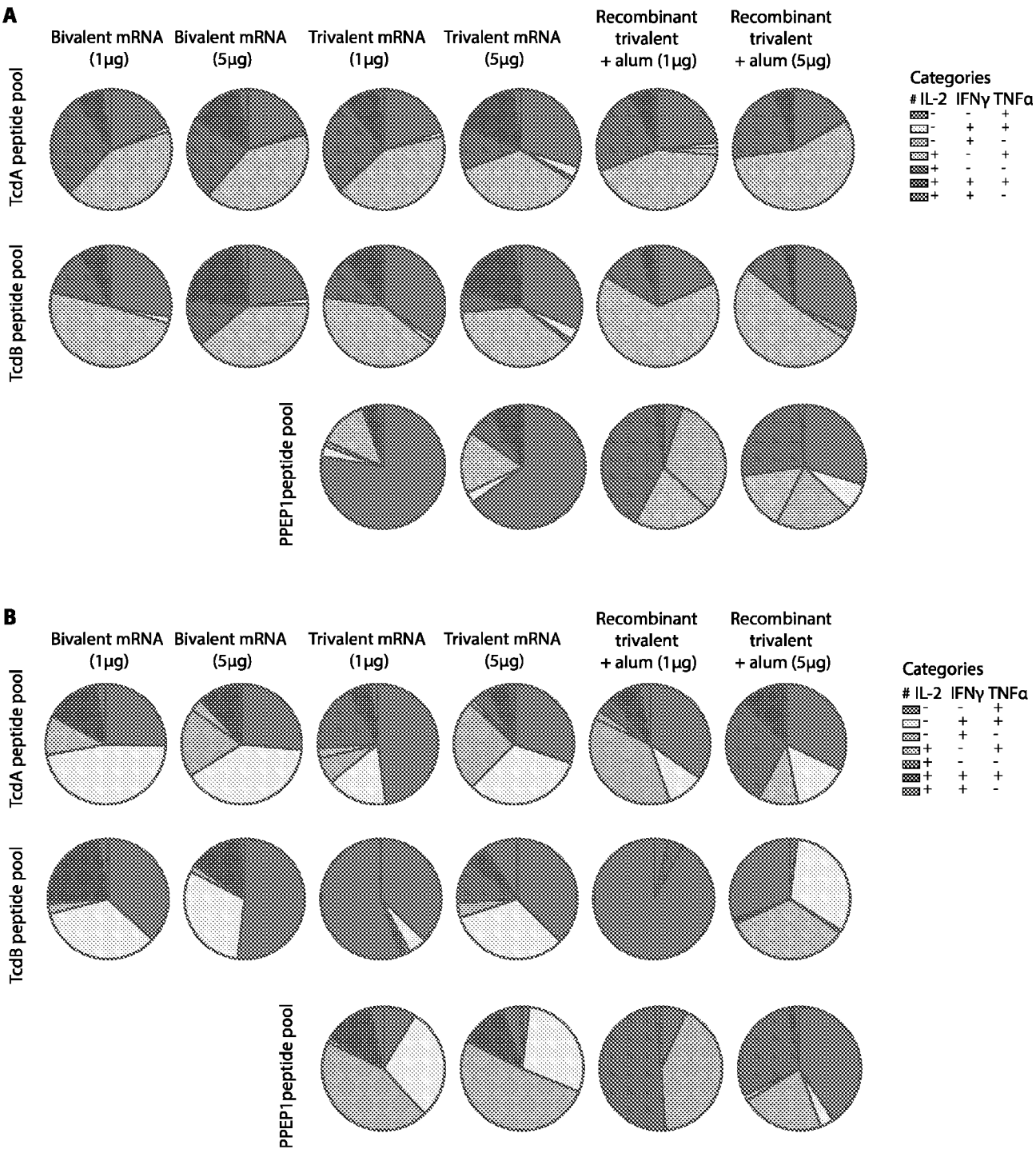


Figure 18

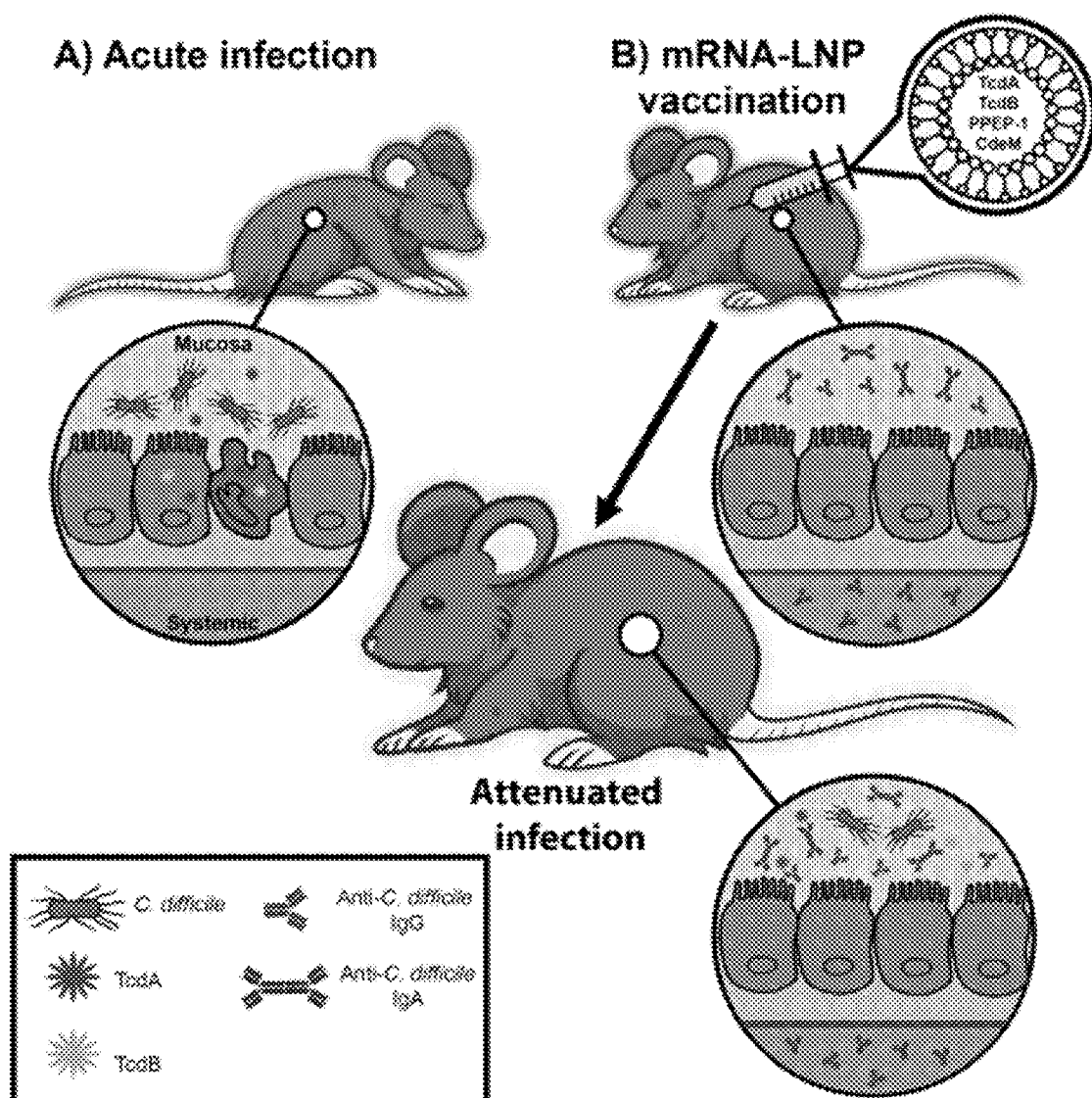


Figure 19

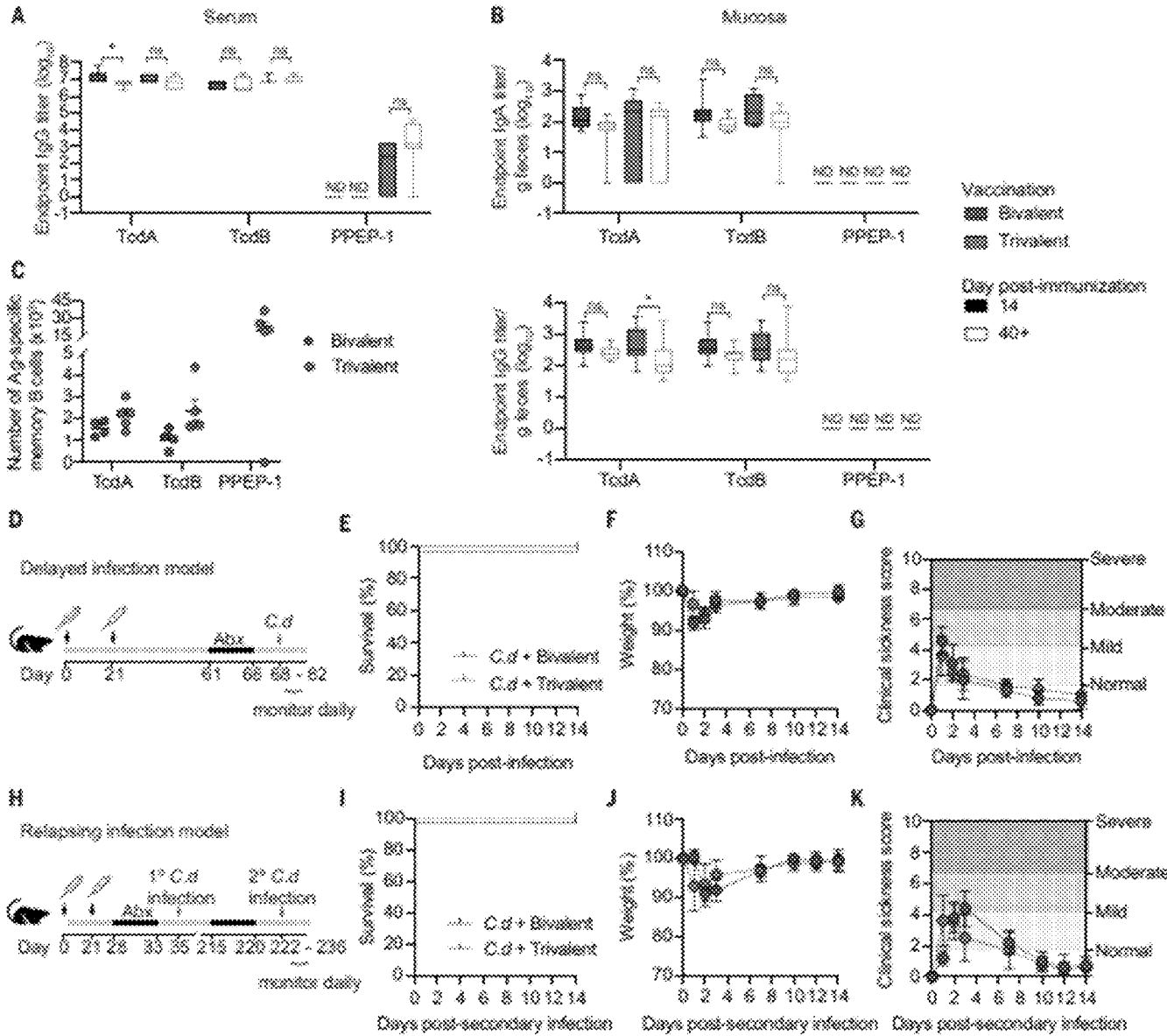


Figure 20

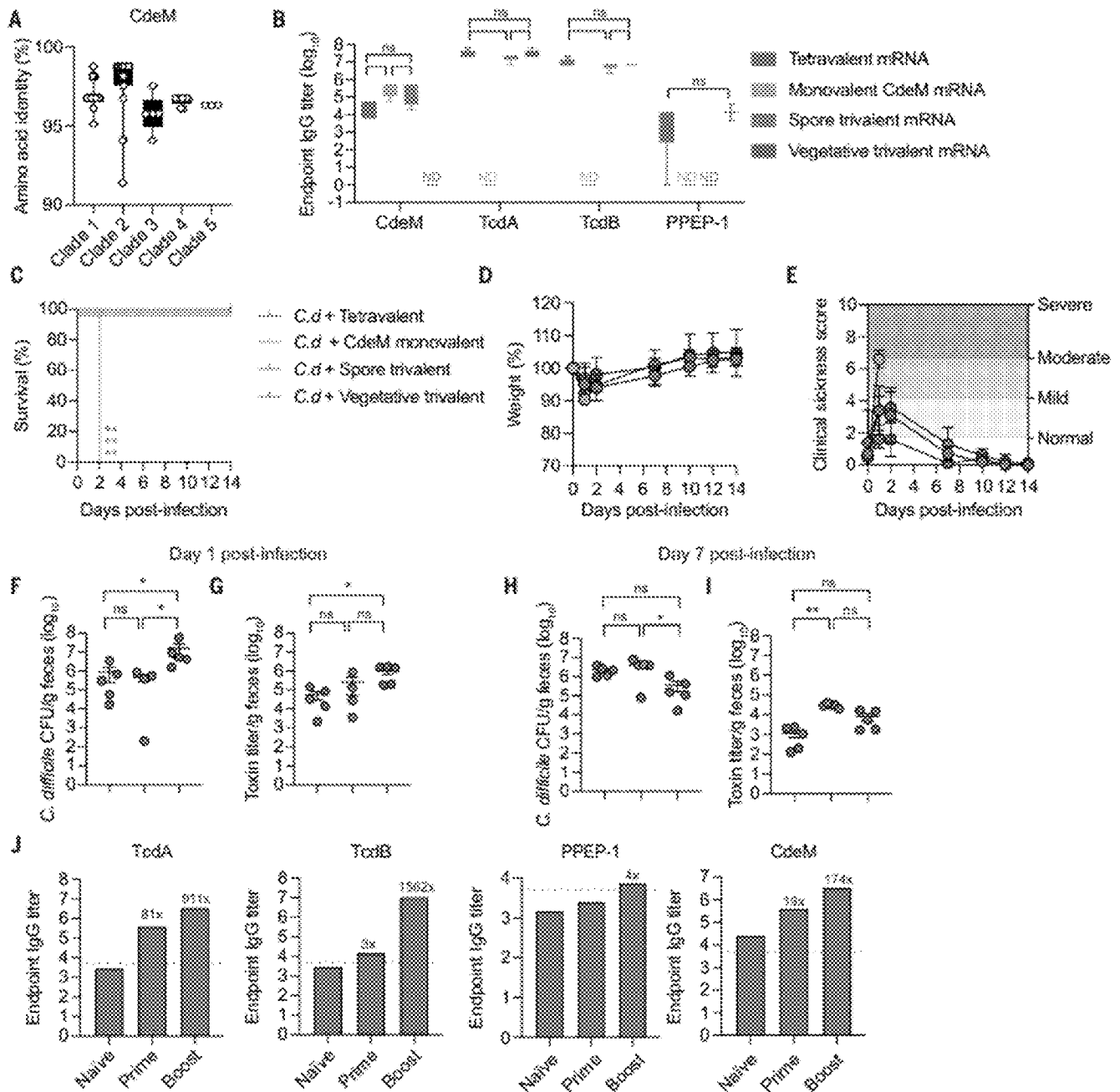


Figure 21

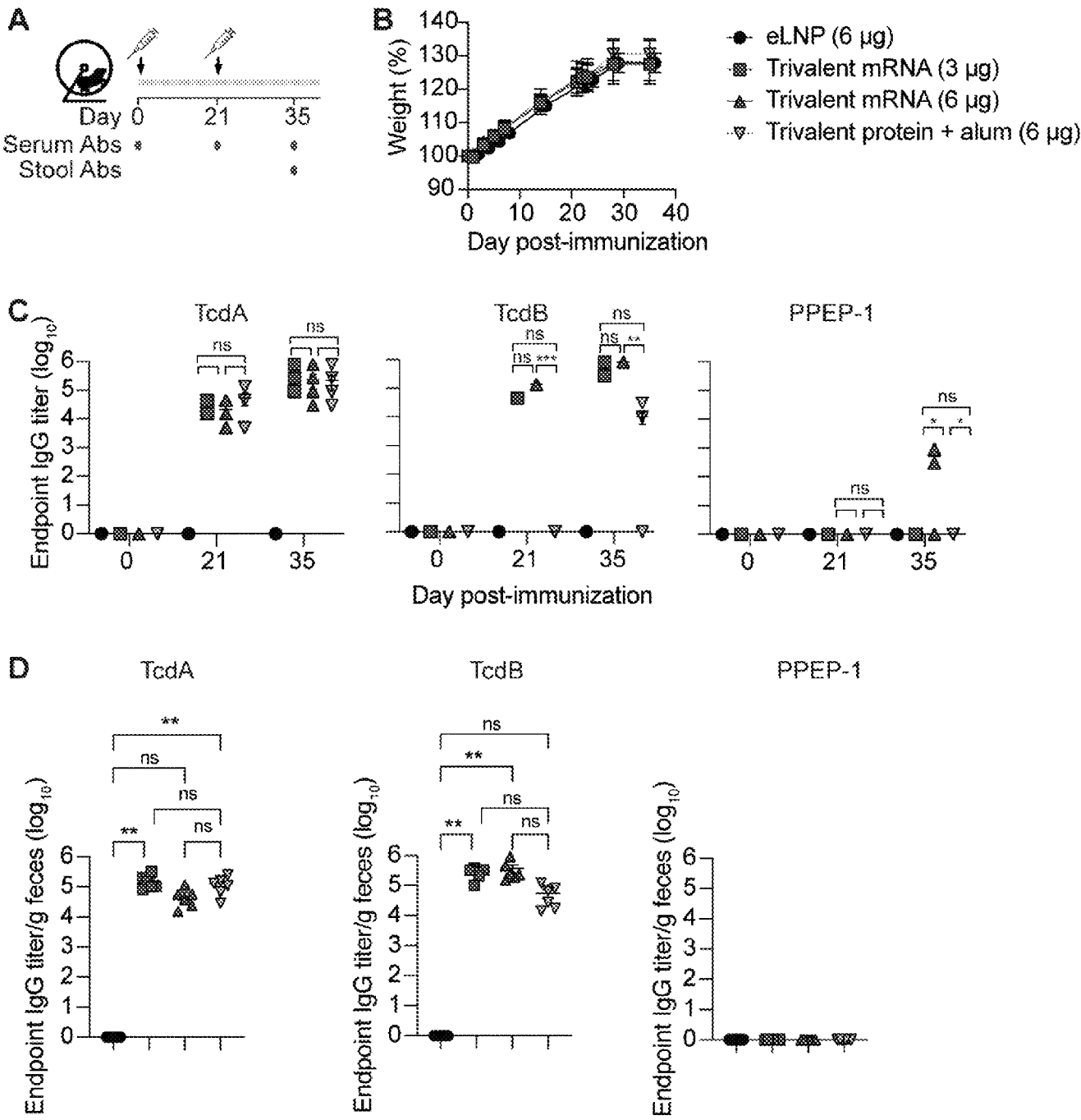


Figure 22

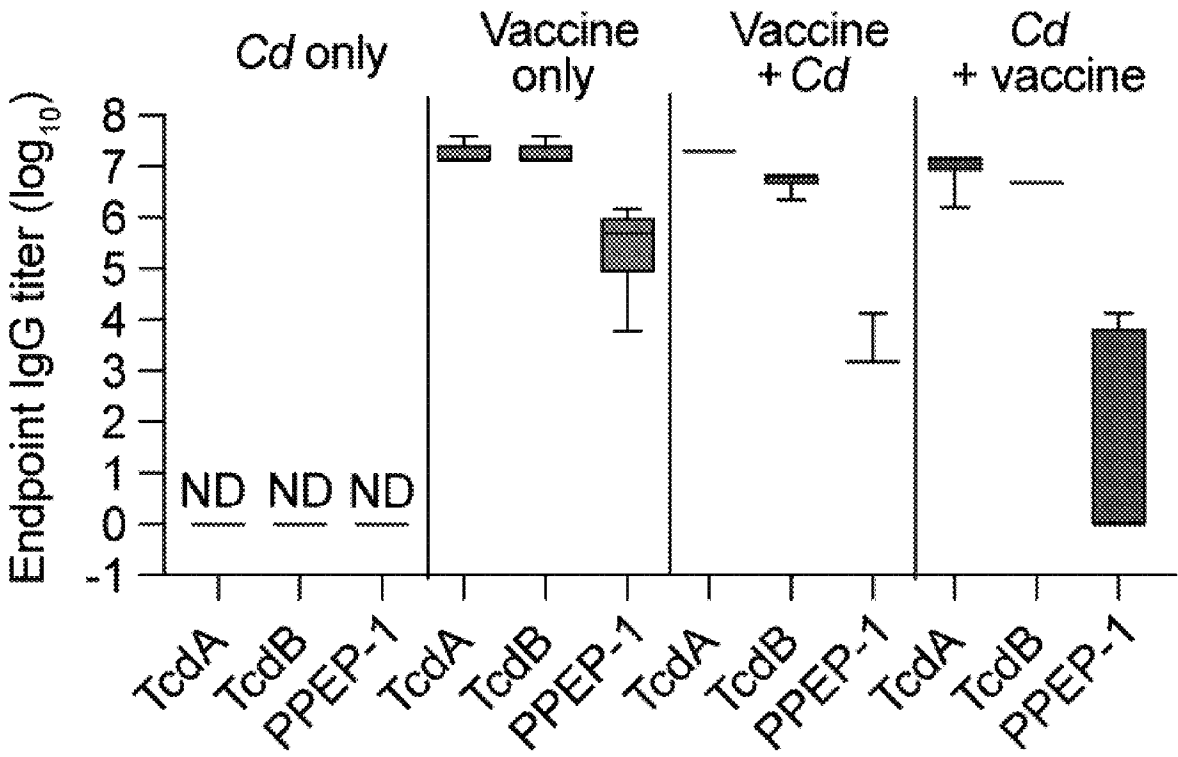


Figure 23

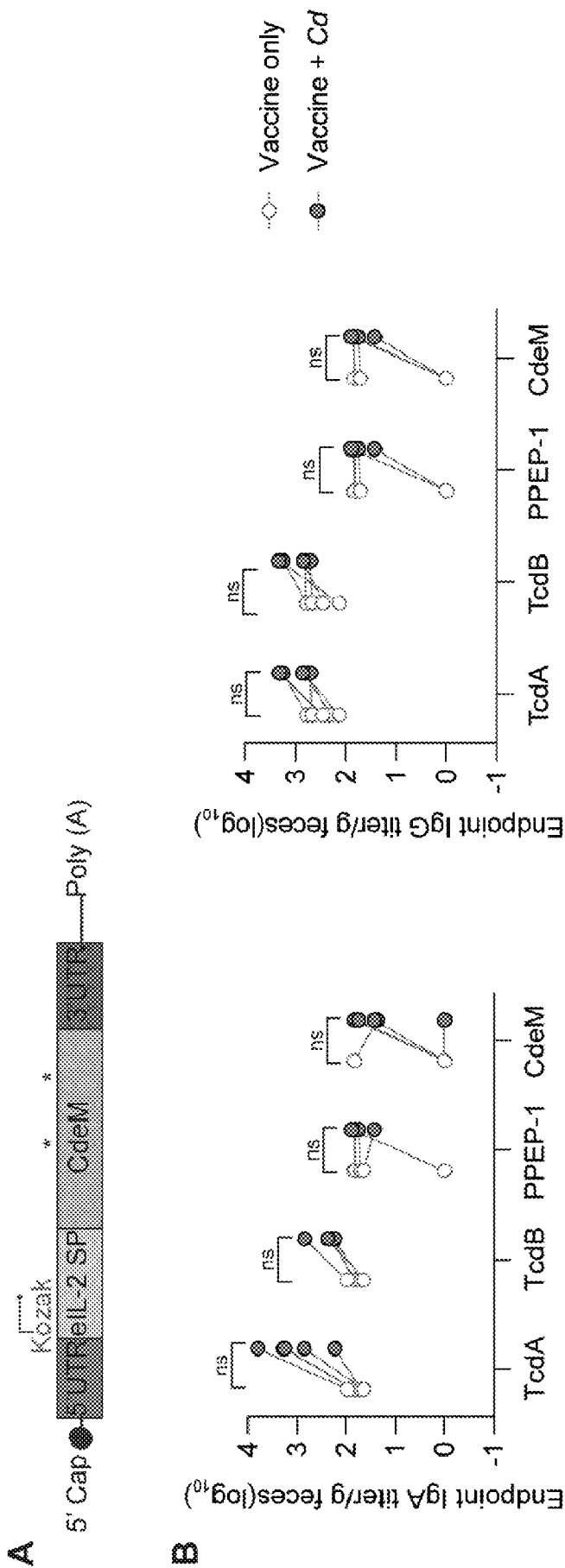


Figure 24

Vaccinated dam Serum

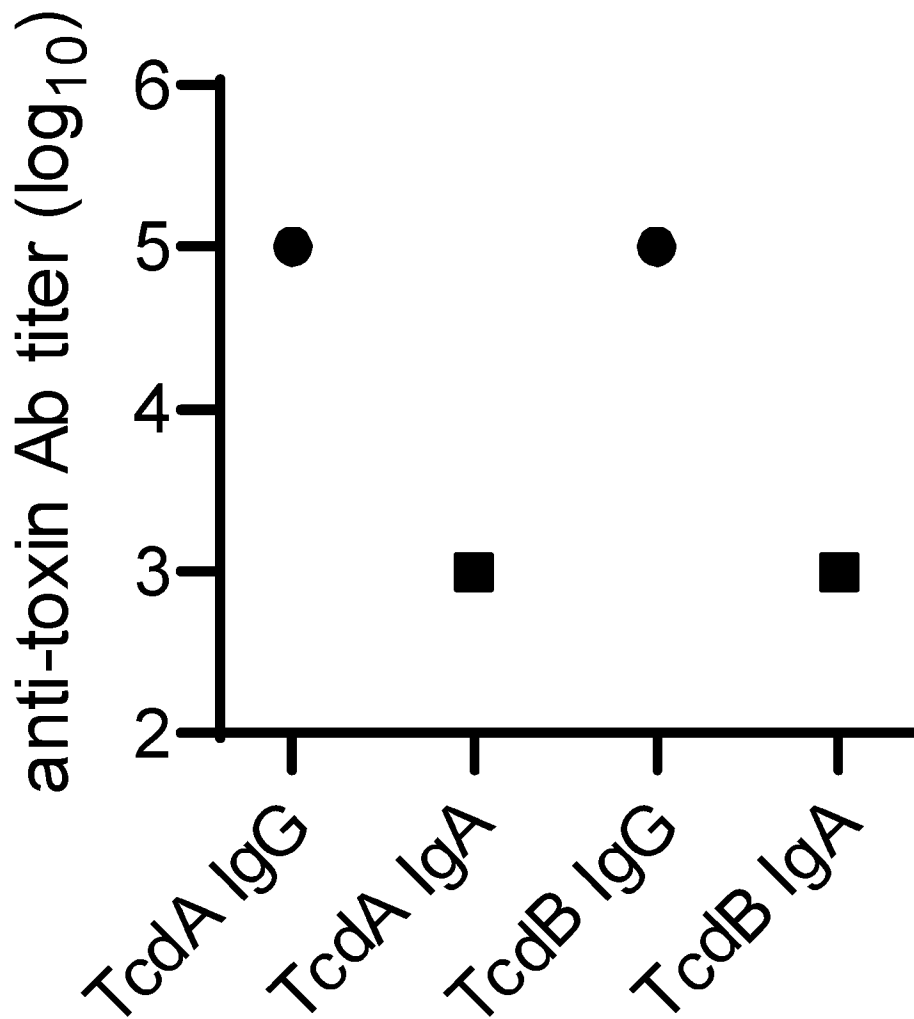


Figure 25

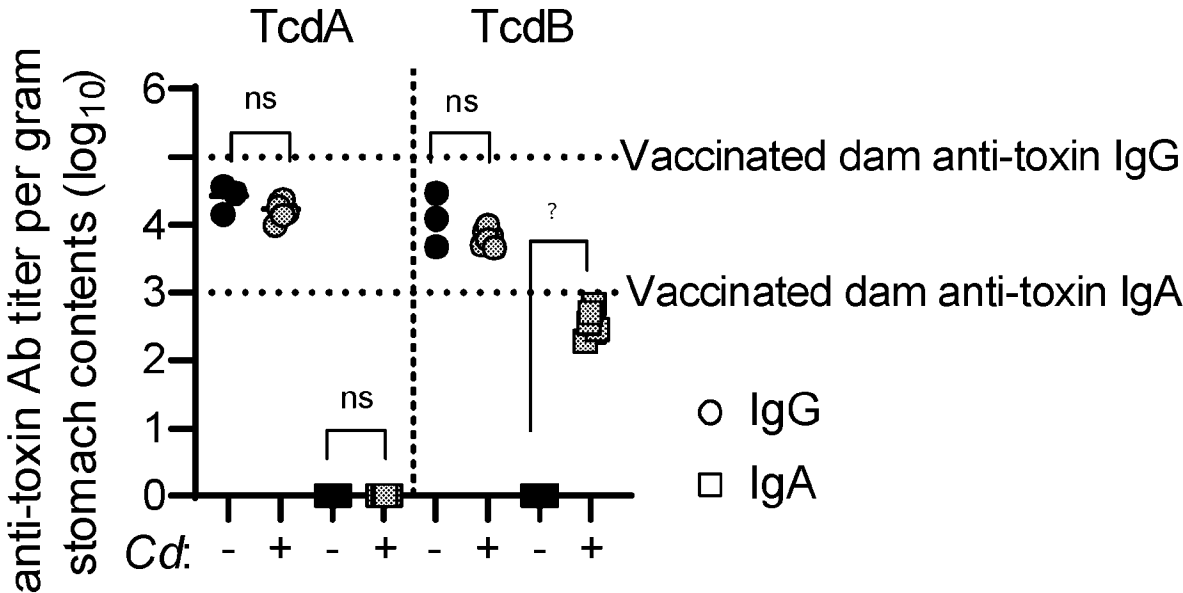


Figure 26

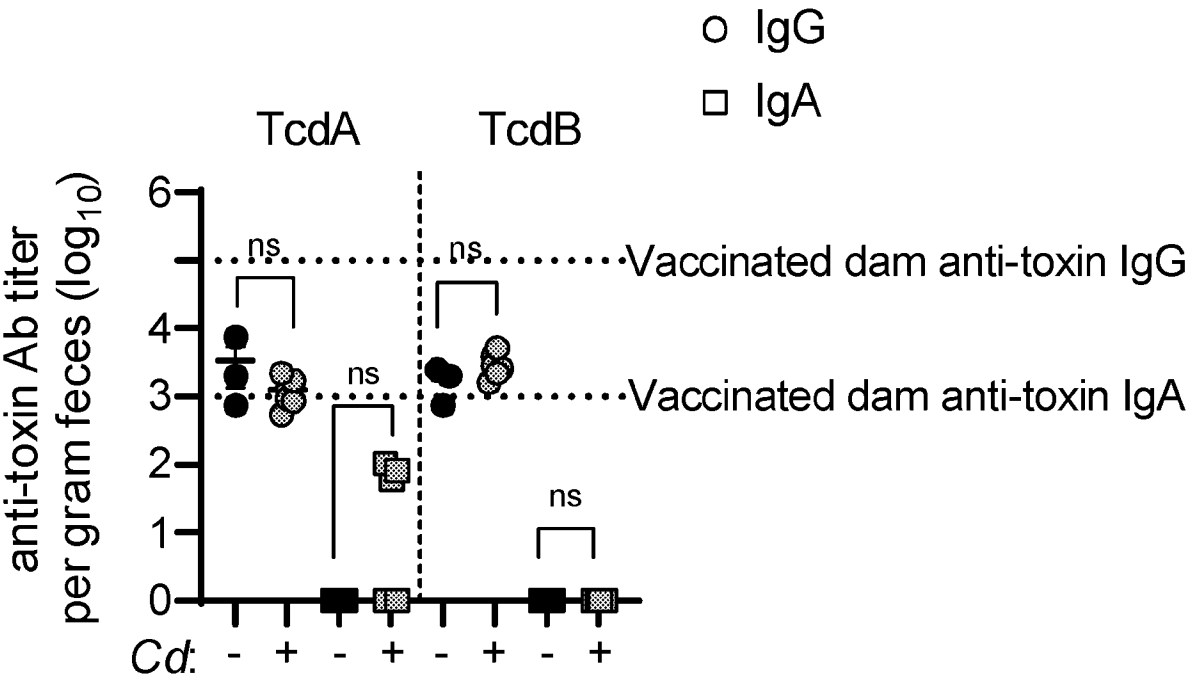


Figure 27

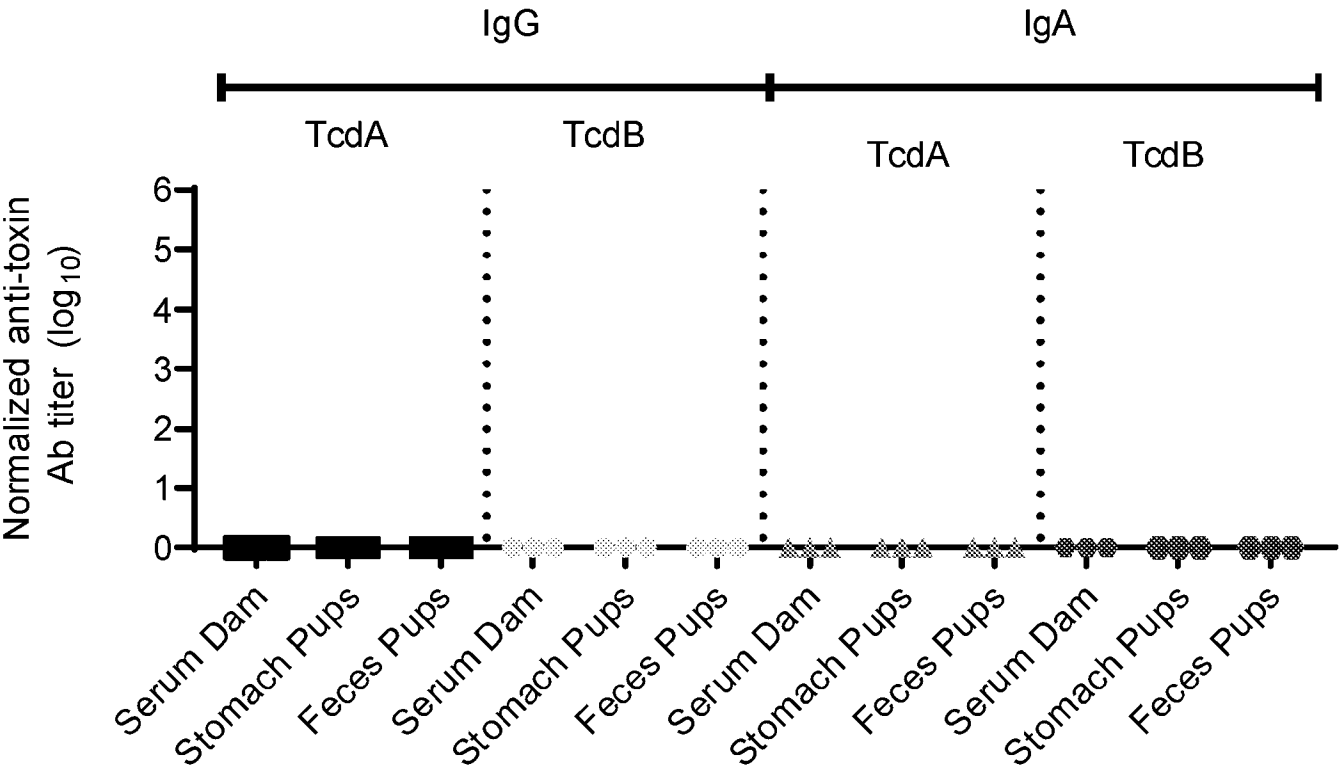


Figure 28

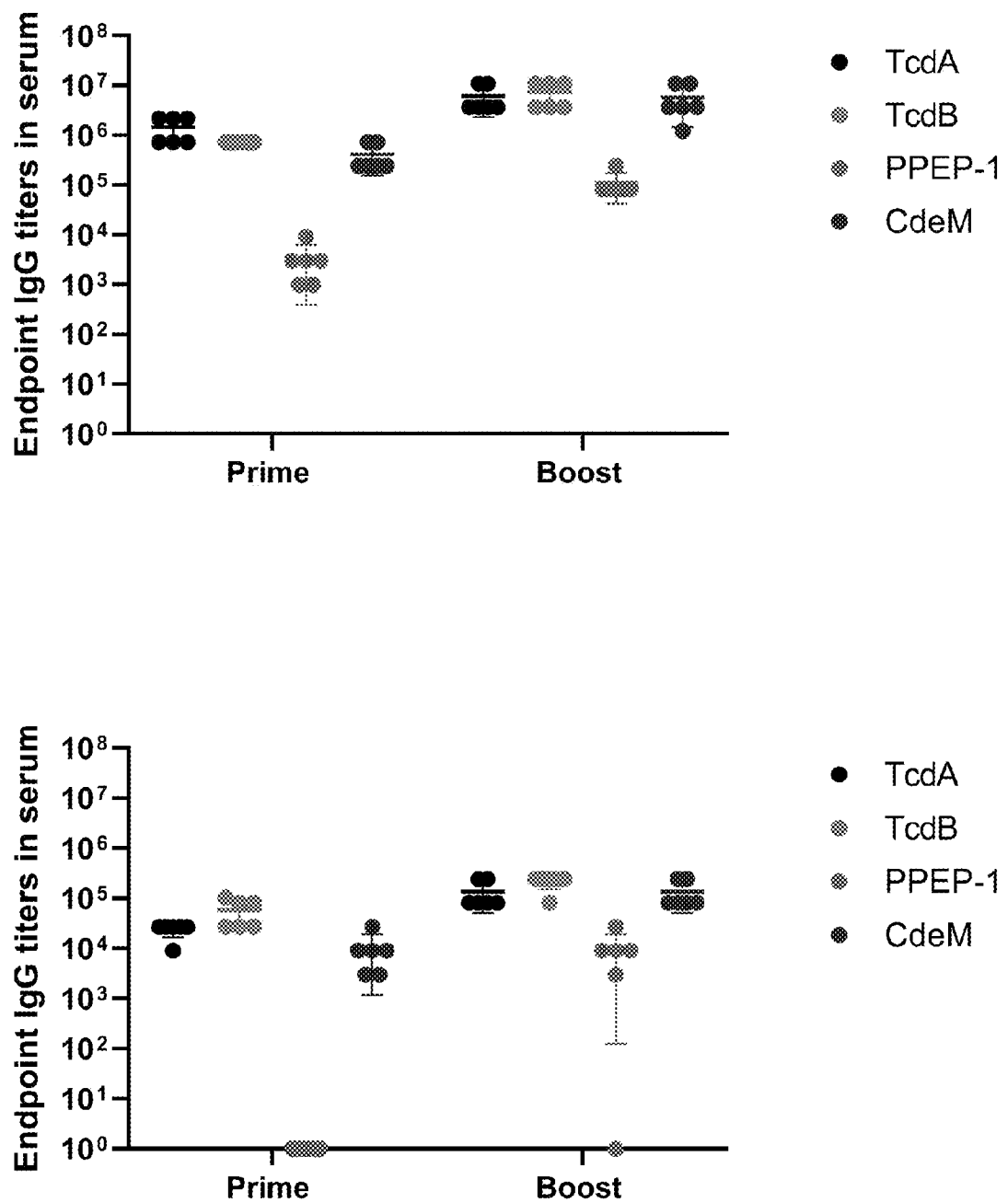


Figure 29

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/060394

A. CLASSIFICATION OF SUBJECT MATTERIPC: **A61K 39/08** (2025.01); **C07K 14/33** (2025.01); **C12N 15/63** (2025.01); **A61P 31/04** (2025.01)CPC: **A61K 39/08**; **C07K 14/33**; **C12N 15/63**; **A61P 31/04**; A61K 2039/53; A61K 2039/55544

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History Document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2023/020992 A1 (GLAXOSMITHKLINE BIOLOGICALS) 23 February 2023 (23.02.2023) pg 96, ln 27-28; pg 98, ln 7-9; pg 101, ln 31-33; pg 105, ln 33-34; pg 106, ln 12-15; pg 114, ln 28-30; pg 120, ln 34; pg 124, ln 5-6; pg 357, ln 30-12	1, 5, 7, 9, 12/ (1, 5, 7, 9, 12)
Y	pg 96, ln 27-28; pg 98, ln 7-9; pg 101, ln 31-33; pg 105, ln 33-34; pg 106, ln 12-15; pg 114, ln 28-30; pg 120, ln 34; pg 124, ln 5-6; pg 357, ln 30-12	6, 12/(6)
A	pg 96, ln 27-28; pg 98, ln 7-9; pg 101, ln 31-33; pg 105, ln 33-34; pg 106, ln 12-15; pg 114, ln 28-30; pg 120, ln 34; pg 124, ln 5-6; pg 357, ln 30-12	8, 12/(8)
Y	US 9,446,112 B2 (KUTZLER et al.) 20 September 2016 (20.09.2016) abstract; col 5, ln 27-29; col 34, ln 24-26	6, 12/(6)
A	US 8,852,600 B2 (GARDINER et al.) 07 October 2014 (07.10.2014) abstract; col 2, ln 16-24	8, 12/(8)
A	US 11,149,286 B1 (MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH) 19 October 2021 (19.10.2021) abstract; col 10, ln 54-55	8, 12/(8)

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

24 March 2025 (24.03.2025)

Date of mailing of the international search report

01 April 2025 (01.04.2025)

Name and mailing address of the ISA/US

COMMISSIONER FOR PATENTS
MAIL STOP PCT, ATTN: ISA/US
P.O. Box 1450
Alexandria, VA 22313-1450
UNITED STATES OF AMERICA

Authorized officer

KARI RODRIQUEZ

Facsimile No. 571-273-8300

Telephone No. PCT Help Desk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/060394

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/060394

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: **13-18, 21-40**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1.

Group I+: Claims 1-12 and 19-20, directed to a composition comprising at least one RNA molecule encoding at least one C. difficile antigen. The composition will be searched to the extent that the composition encompasses TcdA and SEQ ID NOs: 1 and 34. The first named invention was determined based on first claim antigen (claims 1, 7). This first named invention has been selected based on the guidance set forth in section 10.54 of the PCT International Search and Preliminary Examination Guidelines. It is believed that claims 1, 5-9, and 12(in part), encompass this first named invention, and thus these claims will be searched without fee to the extent that they encompass a composition comprising at least one RNA molecule encoding at least one C. difficile antigen comprising TcdA and SEQ ID NO: 1 and 34. Additional compositions comprising additional antigens will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected compositions. Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be a composition comprising at least one RNA molecule encoding at least one C. difficile antigen comprising TcdB and SEQ ID NOs: 8 and 45 (claims 1, 5-9, 12(in part)).

Group II+: Claims 41-43, directed to a composition comprising a fusion molecule comprising a fusion of a C. difficile antigen to an oligomerization domain. Group II+ will be searched upon payment of additional fees. The composition may be searched for an additional fee and election as such, for example, to the extent that the composition encompasses TcdA and SEQ ID NO: 1 (claims 41-43). Additional compositions will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected compositions. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. Another exemplary election would be a composition comprising a fusion molecule comprising a fusion of a C. difficile antigen to an oligomerization domain comprising TcdB and SEQ ID NO: 8 (claims 41-43).

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/060394

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **1, 5-9, and 12(in part), limited to TcdA and SEQ ID NO: 1 and 34**

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.