



(51) International Patent Classification:

A61K 39/00 (2006.01) C12N 7/00 (2006.01)

(21) International Application Number:

PCT/US2017/042308

(22) International Filing Date:

17 July 2017 (17.07.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/363,395 18 July 2016 (18.07.2016) US

(71) Applicant: **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [US/US]; 17 Quincy Street, Cambridge, Massachusetts 02138 (US).

(72) Inventors: **WAGNER, Gerhard**; 111 Princeton Road, Chestnut Hill, Massachusetts 02467 (US). **NASR, Mahmoud**; 170 Brookline Ave., Unit 927, Boston, Massachusetts 02215 (US).

(74) Agent: **KLING, Nicole D.** et al.; Nixon Peabody LLP, 100 Summer St., Boston, Massachusetts 02110-2131 (US).

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

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

(54) Title: IMMUNOGENIC COMPOSITIONS AND METHODS RELATING THERETO

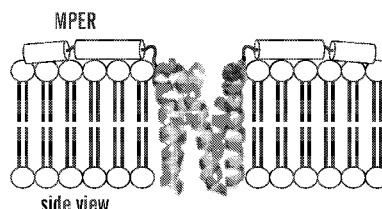
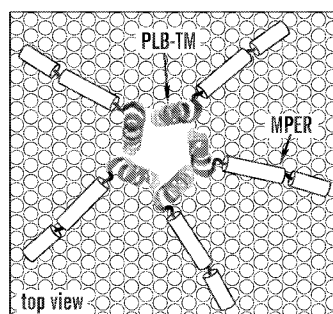


FIG. 21A

(57) Abstract: Described herein are methods and compositions relating to vaccine components, vaccines, and methods of using the same, e.g. by displaying antigens on a polypeptide scaffold embedded in, e.g., a nanodiscs.



Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

IMMUNOGENIC COMPOSITIONS AND METHODS RELATING THERETO**CROSS-REFERENCE TO RELATED APPLICATIONS**

[0001] This application claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 62/363,395 filed July 18, 2016, the contents of which are incorporated herein by reference in their entirety.

GOVERNMENT SUPPORT

[0002] This invention was made with government support under Grant Nos. F32GM113406; GM047467; GM075879; and AI037581 awarded by the National Institutes of Health. The U.S. government has certain rights in the invention.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on July 13, 2017, is named 002806-086592-PCT_SL.txt and is 205,107 bytes in size.

TECHNICAL FIELD

[0004] The technology described herein relates to nanodiscs, proteins used to assemble the nanodiscs, and methods of making nanodiscs.

BACKGROUND

[0005] Small models of the cell membrane can be used to study the structure and behavior of cell membrane proteins, e.g., as part of the development of drugs and/or vaccines. Traditionally, such models were produced by creating micelles or liposomes. Micelles are small spheres which require the use of significant levels of detergents for their creation. These detergents can destabilize and/or inhibit any associated cell membrane proteins, limiting the utility of the micelles. Liposomes do not require detergent use, but tend to aggregate and are irregular in size, making them unsuitable for analysis by crystallization and NMR.

[0006] A more recent development is nanodiscs. Nanodiscs do not contain detergents and are not prone to aggregation. A conventional nanodisc is composed of a nanometer-sized phospholipid bilayer encircled by two copies of α helical, amphipathic membrane scaffold proteins (MSPs) (Denisov et al. J Am Chem Soc. 2004;126:3477-87 and Bayburt et al. Nano Lett. 2002;2:853-6). These conventional nanodiscs are limited in size from 7-17 nm. In addition to this limited range of possible sizes, the size of each conventional nanodisc in a given preparation tends to vary significantly. This lack of uniformity is undesirable for, e.g., crystallization and causes variation in the number of membrane proteins found in each individual nanodisc.

SUMMARY

[0007] As described herein, the inventors have found that by circularizing the membrane scaffold protein, it is possible to obtain nanodiscs of up to 100 nm in diameter with extremely high consistency in the size of each individual nanodisc. Additionally, the nanodiscs made with such circularized membrane scaffold proteins can be designed to be either circular or polygonal in shape. The polygonal nanodiscs are of particular interest as they can increase the likelihood of crystallization.

[0008] In one aspect of any of the embodiments, described herein is a loopable membrane scaffold protein comprising, from N-terminus to C-terminus: i) an N-terminal circularization domain; ii) a plurality of amphipathic alpha helix domains; and iii) a C-terminal circularization domain.

[0009] In some embodiments of any of the aspects, one circularization domain comprises: a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and the other circularization domain comprises: a sequence LPXTG/A; wherein X represents any amino acid. In some embodiments of any of the aspects, the N-terminal circularization domain comprises: a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and the C-terminal circularization domain comprises: a sequence LPXTG/A; wherein X represents any amino acid. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to 10 amino acids in length and comprises amino acids selected from the group consisting of: glycine and alanine. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 glycine residues. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 alanine residues. In some embodiments of any of the aspects, the sequence LPXTG/A is covalently bound to the at least one glycine or alanine residue of (i). In some embodiments of any of the aspects, the sequence LPXTG/A is LPGTG/A (SEQ ID NO: 24); LPSTG/A (SEQ ID NO: 25); LPETG/A (SEQ ID NO: 26).

[0010] In some embodiments of any of the aspects, one circularization domain comprises an intein separated from the domain of (ii) by a Cys; and the other circularization domain comprises an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein. In some embodiments of any of the aspects, the N-terminal circularization domain comprises an intein separated from the domain of (ii) by a Cys; and the C-terminal circularization domain comprises an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein. In some embodiments of any of the aspects, at least one intein is flanked by a chitin binding domain (CBD).

[0011] In some embodiments of any of the aspects, each circularization domain comprises a Cys.

[0012] In some embodiments of any of the aspects, the amphipathic alpha helix domain comprises: a) an ApoA-I polypeptide; or b) at least one alpha helical polypeptide with secondary amphipathicity or a

repeat thereof; c) at least one amphipathic proline-rich polypeptide or a repeat thereof; or d) at least one repeat of the sequence XZXPPP; wherein X is any hydrophobic residue and Z is any hydrophilic residue. In some embodiments of any of the aspects, the alpha helical polypeptide has a sequence selected from the group consisting of: DWFKAFYDKLAEKFKEAF (SEQ ID NO: 27); DFLKAFYDKVAEKFKEAAPDWFKAFYDKVAEKFKEAF (SEQ ID NO: 28); KLALRLALKAFKAALKLA (SEQ ID NO: 29); RLALDLALRAFKAAWKLA (SEQ ID NO: 30); ELAWDLAFEALDAELKLD (SEQ ID NO: 31); FLKLLKKFLKLFKKLLKLF (SEQ ID NO: 32); WLKLLKKWLKLWKKLLKL (SEQ ID NO: 33); LLKFLKRLLKLLKDLWKLL (SEQ ID NO: 34); WEAAFAEALAEAWAEHLAEALAEAVEALAA (SEQ ID NO: 35); WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 36); RAFARALARALKALARALKALAR (SEQ ID NO: 37); GLFEALLELLESLWELLLEA (SEQ ID NO: 38); KLLKLLLKLWKKLLKLLK (SEQ ID NO: 39); DWLKAFYDKVAEKLKEAF (SEQ ID NO: 40); and DWFKAFYDKVAEKFKEAF (SEQ ID NO: 41).

[0013] In some embodiments of any of the aspects, the protein further comprises a cap sequence N-terminal of the N-terminal circularization domain that can be cleaved to expose the N-terminal circularization domain as the N-terminus of the protein. In some embodiments of any of the aspects, the cap sequence comprises a sequence selected from the group consisting of: a) Leu-Val-Pro-Arg-Gly-Ser (SEQ ID NO: 42); and b) Glu-X-Leu-Tyr- Φ -Gln- ϕ where X is any residue, Φ is any large or medium hydrophobic residue and ϕ is glycine or alanine (SEQ ID NO: 43). In some embodiments of any of the aspects, the cap sequence further comprises an affinity purification tag N-terminal of the cleavage site. In some embodiments of any of the aspects, the protein further comprises a His tag flanking one of the circularization domains. In some embodiments of any of the aspects, the His tag is located C-terminal of the C-terminal circularization domain. In some embodiments of any of the aspects, the protein further comprises a linker flanking one of the circularization domains. In some embodiments of any of the aspects, the linker is located C-terminal of the C-terminal circularization domain.

[0014] In some embodiments of any of the aspects, the plurality of amphipathic alpha helix domains are separated by flexible linker sequences. In some embodiments of any of the aspects, the flexible linker sequence is selected from the group consisting of: LPGTGS (SEQ ID NO: 44); LPGTSG (SEQ ID NO: 45); LPSTGS (SEQ ID NO: 46); and LPSTSG (SEQ ID NO: 47).

[0015] In some embodiments of any of the aspects, the protein has the sequence of any of SEQ ID NOs: 1-4 and 6-23.

[0016] In one aspect of any of the embodiments, described herein is a covalently circularized nanodisc comprising: a) a phospholipid bilayer; and b) a looped membrane scaffold protein forming a belt

around the bilayer, wherein the looped membrane scaffold protein comprises: i) an N-terminal circularization domain covalently linked to a C-terminal circularization domain; and ii) a plurality of amphipathic alpha helix domains.

[0017] In some embodiments of any of the aspects, an N-terminal circularization domain covalently linked to a C-terminal circularization domain comprises the sequence LPXT followed by a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; wherein X represents any amino acid. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to 10 amino acids in length and comprises amino acids selected from the group consisting of: glycine and alanine. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 glycine residues. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 alanine residues. In some embodiments of any of the aspects, the sequence LPXT is LPGT (SEQ ID NO: 48); LPST (SEQ ID NO: 49); or LPET (SEQ ID NO: 50).

[0018] In some embodiments of any of the aspects, the N-terminal circularization domain covalently linked to a C-terminal circularization domain is a cysteine residue.

[0019] In some embodiments of any of the aspects, the diameter is greater than about 6 nm. In some embodiments of any of the aspects, the diameter is greater than about 15 nm. In some embodiments of any of the aspects, the diameter is greater than about 50 nm. In some embodiments of any of the aspects, the diameter is about 80 nm.

[0020] In one aspect of any of the embodiments, described herein is a method of making a nanodisc as described herein, the method comprising: contacting the loopable membrane scaffold protein described herein with a sortase enzyme; and contacting the looped protein produced in step (a) with solubilized phospholipids. In some embodiments of any of the aspects, the sortase is selected from the group consisting of: *Staphylococcus aureus* wild type sortase A (Srt A); evolved sortase; eSrtA; eSrtA(2A-9); eSrtA(4S-9); and *Streptococcus pyogenes* sortase A. In some embodiments of any of the aspects, the contacting step is stopped by quenching the reaction with a sortase inhibitor. In some embodiments of any of the aspects, the sortase inhibitor is AAEK2 or high concentrations of EDTA. In some embodiments of any of the aspects, the sortase is bound to a substrate. In some embodiments of any of the aspects, the sortase further comprises a His tag.

[0021] In one aspect of any of the embodiments, described herein is a method of making a nanodisc as described herein, the method comprising: a) contacting the loopable membrane scaffold protein described herein with a solution having a pH low enough to induce cleavage of the dipeptide sequence of Asn-Cys; and a thiol reagent to induce cleavage at the Cys residue, resulting in circularization of the

protein; and b) contacting the looped protein produced in step (a) with solubilized phospholipids. In some embodiments of any of the aspects, the thiol reagent is DTT.

[0022] In one aspect of any of the embodiments, described herein is a method of making a nanodisc as described herein, the method comprising a) contacting the loopable membrane scaffold protein described herein with a thiol reagent, resulting in circularization of the protein; and b) contacting the looped protein produced in step (a) with solubilized phospholipids.

[0023] In some embodiments of any of the aspects, the phospholipids are solubilized in detergent. In some embodiments of any of the aspects, step (b) occurs at a temperature of from about 0 C to about 42 C. In some embodiments of any of the aspects, step (b) occurs at a temperature of from about 0 C to about 37 C. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:2000 to about 1:5000. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:3500 to about 1:4500; and circular nanodiscs are formed. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:4000; and circular nanodiscs are formed. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:2500 to about 1:3700; and non-circular nanodiscs are formed. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:3100 to about 1:3400; and non-circular nanodiscs are formed.

[0024] In some embodiments of any of the aspects, the protein further comprises a cap sequence N-terminal of the N-terminal circularization domain and the method comprises a first step of contacting the protein with an enzyme to cleave the cap sequence. In some embodiments of any of the aspects, the protein further comprises a C-terminal His tag and the method comprises a first step of binding the protein to a substrate. In some embodiments of any of the aspects, the substrate comprises Cu²⁺, Ni²⁺ or Co²⁺. In some embodiments of any of the aspects, the substrate is a chip or bead. In some embodiments of any of the aspects, the substrate is a bead and step (a) is performed in the presence of phospholipids. In some embodiments of any of the aspects, the substrate is a bead and the steps (a) and (b) are performed concurrently. In some embodiments of any of the aspects, the method further comprises, after step (a), a step of eluting the looped protein produced in step (a) from the substrate.

[0025] In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a concentration of less than 100 uM prior to the contacting step. In some embodiments of any of the aspects, the looped protein formed in step (a) is passed through a column, substrate, and/or solution comprising Cu²⁺, Co²⁺ or Ni²⁺.

[0026] In one aspect of any of the embodiments, described herein is a vaccine scaffold polypeptide comprising at least the transmembrane domain of a single α -helix transmembrane protein, or a variant thereof. In some embodiments of any of the aspects, the single α -helix transmembrane protein is selected from the group consisting of phospholamban; influenza M2 channel protein; stannin; Vpu; HIV Vpu; parainfluenza 5 fusion protein; parainfluenza fusion protein; E protein; SARS-CoV E; SH protein; and RSV-SH. In some embodiments of any of the aspects, the transmembrane domain is the sequence of one of SEQ ID NOs: 95-101, 113, or 114. In some embodiments of any of the aspects, the variant of the transmembrane domain is a sequence with at least 90% homology to one of SEQ ID NO: 95-101, 113 or 114.

[0027] In one aspect of any of the embodiments, described herein is an immunogenic conjugate polypeptide comprising: a) at least one antigen polypeptide conjugated to; b) a vaccine scaffold polypeptide as described herein. In some embodiments of any of the aspects, the antigen polypeptide is conjugated to the N-terminus of the vaccine scaffold polypeptide. In some embodiments of any of the aspects, the antigen polypeptide is conjugated to the C-terminus of the vaccine scaffold polypeptide.

[0028] In some embodiments of any of the aspects, the antigen polypeptide is a viral envelope protein, a cell surface protein, or portion thereof. In some embodiments of any of the aspects, the antigen polypeptide is an HIV protein or portion thereof. In some embodiments of any of the aspects, the antigen polypeptide is the membrane proximal external region (MPER) of Env. In some embodiments of any of the aspects, the antigen polypeptide is the 10E8 epitope or the 4E10 epitope.

[0029] In some embodiments of any of the aspects, the immunogenic conjugate polypeptide comprises multiple repeats of the same antigen polypeptide. In some embodiments of any of the aspects, the immunogenic conjugate polypeptide comprises multiple different antigen polypeptides. In some embodiments of any of the aspects, the antigen polypeptides or repeats thereof are separated by a linker.

[0030] In some embodiments of any of the aspects, the immunogenic polypeptide comprises from 1 to 5 copies of the antigen polypeptide. In some embodiments of any of the aspects, the antigen polypeptide and the vaccine scaffold polypeptide are joined by a linker.

[0031] In some embodiments of any of the aspects, the linker is hydrophobic. In some embodiments of any of the aspects, the linker is hydrophilic. In some embodiments of any of the aspects, the linker is from 1 to 5 amino acids in length.

[0032] In one aspect of any of the embodiments, described herein is a nucleic acid encoding a vaccine scaffold polypeptide or an immunogenic conjugate polypeptide as described herein. In one aspect of any of the embodiments, described herein is a polypeptide complex comprising a polymeric complex of a vaccine scaffold polypeptide and/or an immunogenic conjugate polypeptide as described herein.

[0033] In one aspect of any of the embodiments, described herein is a composition comprising a polypeptide complex and/or immunogenic conjugate polypeptide as described herein embedded in the phospholipid bilayer of a liposome or nanodisc. In some embodiments of any of the aspects, the phospholipid bilayer comprises a viral membrane lipid; anionic lipid; or synthetic lipid. In some embodiments of any of the aspects, the phospholipid bilayer comprises cholesterol. In some embodiments of any of the aspects, each liposome or nanodisc comprises 10 or fewer copies of the antigen polypeptide. In some embodiments of any of the aspects, each liposome or nanodisc comprises 5 or fewer copies of the antigen polypeptide.

[0034] In one aspect of any of the embodiments, described herein is a pharmaceutical composition comprising a polypeptide complex and/or immunogenic conjugate polypeptide as described herein embedded in the phospholipid bilayer of a liposome or nanodisc and a pharmaceutically acceptable carrier. In some embodiments of any of the aspects, the pharmaceutical composition further comprises an adjuvant.

[0035] In one aspect of any of the embodiments, described herein is a method of inducing an immune response in a subject, the method comprising administering to the subject a composition comprising a polypeptide complex and/or immunogenic conjugate polypeptide as described herein embedded in the phospholipid bilayer of a liposome or nanodisc. In some embodiments of any of the aspects, the immune response comprises immunization. In some embodiments of any of the aspects, the subject is further administered an adjuvant. In some embodiments of any of the aspects, the compositions is admixed with the adjuvant prior to administration. In some embodiments of any of the aspects, the administration is a prime administration and/or a boost administration. In some embodiments of any of the aspects, the subject is further administered a second vaccine. In some embodiments of any of the aspects, the method further comprises a step of isolating from the subject:

- a. an antibody that specifically binds to the antigen polypeptide;
- b. a nucleic acid encoding an antibody that specifically binds to the antigen polypeptide; or
- c. a cell comprising a nucleic acid sequence encoding an antibody that specifically binds to the antigen polypeptide.

BRIEF DESCRIPTION OF THE DRAWINGS

[0036] Fig. 1 depicts a schematic representation of a nanodisc.

[0037] Figs. 2A-2E depict schematics of the production of covalently circularized MSP and nanodiscs. Fig. 2A depicts a general outline of the constructs that are used for making covalently circularized nanodiscs. Constructs containing a sortase A recognition sequence (LPXTG (SEQ ID NO:

51)) and N-terminal glycine can undergo circularization (intramolecular transpeptidation) after exposing the N-terminal glycine and incubating with evolved sortase. Fig. 2A discloses "HHHHHH" as SEQ ID NO: 54. Fig. 2B depicts an outline of the procedure for creating circularized proteins over a Cu²⁺ chip. Fig. 2B discloses "GSTFSKL" as SEQ ID NO: 80, "KLNTQLPGTAAALEHHHHHH" as SEQ ID NO: 81, "KLNTQLPGTGSTFSKLR" as SEQ ID NO: 82 and "AALEHHHHHH" as SEQ ID NO: 83. Fig. 2C shows SDS-PAGE analysis of NW11 before (lane 1) and after (lane 2) circularization. Fig. 2C discloses "LPXTG" as SEQ ID NO: 51 and "HHHHHH" as SEQ ID NO: 54. Fig. 2D depicts MS/MS spectrum of a tryptic peptide of cNW11 confirming the ligation of the N-terminal residues (GSTFSK (SEQ ID NO: 52)) to the C-terminal LPGTG motif (SEQ ID NO: 53). The b and y ions that were identified in the MS/MS spectrum are highlighted in blue and red. Fig. 2D discloses "KLNTQLPGTGSTFSK" as SEQ ID NO: 84. Fig. 2E depicts the diameter distribution for nanodiscs made using linear MSP (top) and circularized MSP (bottom) and their representative negative stain images. Fig. 2E discloses "NTQLPGTAA" as SEQ ID NO: 85 and "NTQLPGTGST" as SEQ ID NO: 86.

[0038] Figs 3A-3C demonstrate the production of large covalently circularized nanodiscs. Fig. 3A demonstrate SDS-PAGE analysis of NW30 before and after circularization. The circularized NW30 migrates slower than the linear form. Right: negative-stain EM analysis of the nanodiscs made using circularized NW30 (cNW30) shows the formation of ~15 nm nanodiscs. Fig. 3B demonstrate SDS-PAGE analysis of NW50 before and after circularization. Right: negative-stain EM analysis of the nanodiscs made using circularized NW50 (cNW50) shows the formation of ~50 nm nanodiscs. Bottom: cryo-EM images for individual nanodiscs showing top and side views. Fig. 3C depicts negative stain images showing single nanodisc particles with increasing sizes as indicated.

[0039] Fig. 4A depicts cryo-EM images of individual nanodiscs with different shapes. Fig. 4B depicts potential molecular arrangements of scaffold protein molecules in the different shaped nanodiscs. Fig. 4C depicts possible crystal packing of nanodiscs with the corresponding shapes. The triangular, square and hexagonal shapes should, in principle, provide better crystal packing than pentagonal or circular ones. Fig. 4D shows that each triangular disc is made of 3 copies of amphipathic protein that are connected together with a long flexible loop.

[0040] Fig. 5 depicts a schematic of the strategy for making polygonal nanodiscs and the associated scaffold proteins. Fig. 5 discloses "LPXTG" as SEQ ID NO: 51 and "HHHHHH" as SEQ ID NO: 54.

[0041] Figs. 6A-6J depict poliovirus caught in the act. Fig. 6A depicts negative-stain EM of 50-nm circularized nanodiscs plus poliovirus. A control (nanodiscs without CD155) is shown to illustrate the relative dimensions of the 30-nm poliovirus and the 50-nm nanodisc. Fig. 6B depicts an outline of the

procedure used to initiate poliovirus bridging and fusion with nanodiscs decorated with CD155. Fig. 6C depicts negative-stain EM images showing individual viruses tethered to nanodiscs. Fig. 6D depicts cryo-EM image of 50-nm nanodiscs plus poliovirus. Fig. 6E depicts cryo-EM image showing a tilted view of 50-nm nanodisc. Fig. 6F depicts cryo-EM images showing individual viruses tethered to nanodiscs. Fig. 6G-6H depict cryo-EM images showing the creation of a putative pore in the nanodisc by the poliovirus. Fig. 6I depicts cryo-EM image showing three viral particles around a 15-nm nanodisc. Fig. 6J depicts cryo-EM image showing individual viruses ejecting RNA after incubation with CD155-decorated 15-nm nanodiscs.

[0042] Figs 7A-7D demonstrate the analysis of VDAC-1 in different size nanodiscs. Fig. 7A depicts 2D- $[^1\text{H}, ^{15}\text{N}]$ -TROSY spectrum of $^2\text{H}, ^{15}\text{N}$ -labeled VDAC-1 in cNW9 nanodiscs. Fig. 7B depicts representative image of negatively stained cNW9 nanodiscs containing single VDAC-1 channel. The stain-filled channels appear as dark spot inside the nanodisc. Right: SDS-PAGE analysis of cNW9 nanodiscs containing VDAC-1. Fig. 7C depicts 2D- $[^1\text{H}, ^{15}\text{N}]$ -TROSY spectrum of $^2\text{H}, ^{15}\text{N}$ -labeled VDAC-1 in cNW11 nanodiscs. Fig. 7D depicts representative image of negatively stained cNW11 nanodiscs containing VDAC-1 channels. Each nanodisc appear to contain 2 channels. Right: SDS-PAGE analysis of cNW11 nanodiscs containing VDAC-1.

[0043] Fig. 8 depicts a schematic outline of the procedure for creating circularized NW proteins over Cu+2 chip. His6- tagged NWs ("His6" disclosed as SEQ ID NO: 54) is attached to Cu+2 chip. Addition of His6-tagged evolved sortase ("His6" disclosed as SEQ ID NO: 54) creates circularized non- His tagged NW (cNW), which can be eluted while cleaved His6- tagged products ("His6" disclosed as SEQ ID NO: 54) and evolved sortase remain bound to the Cu+2 chip. The eluted cMSP is passed through Ni column to get rid of any traces of His6- tagged products ("His6" disclosed as SEQ ID NO: 54) and evolved sortase. Fig. 8 discloses "GSTFSKL" as SEQ ID NO: 80, "KLNTQLPGTAAALEHHHHHH" as SEQ ID NO: 81, "AAALEHHHHHH" as SEQ ID NO: 83 and "LEEYTKKLNTQLPGTGSTFSKLREQL" as SEQ ID NO: 87.

[0044] Fig. 9A depicts a schematic of circularization of NW proteins over Ni²⁺ beads. His6- tagged linear NW ("His6" disclosed as SEQ ID NO: 54) is attached to Ni-beads followed by adding lipids to NW. Addition of lipids to NW helps to bring the C and N terminal ends close so the circularization pathway will dominate. Adding His6-tagged sortase ("His6" disclosed as SEQ ID NO: 54) creates circularized non- tagged NW (cNW), which can be eluted while cleaved His6- tagged cleavage product ("His6" disclosed as SEQ ID NO: 54) and sortase remain bound to the beads. Fig. 9A discloses "KLNTQLPGTGSTFSK" as SEQ ID NO: 84. Fig. 9B SDS-PAGE gel showing the final products after circularization over Ni beads with or without lipids.

[0045] Fig. 10 depicts a schematic outline of the procedure for creating circularized NW in solution. His6- tagged linear NW ("His6" disclosed as SEQ ID NO: 54) is diluted before adding sortase to favor the circularization. Adding the sortase covalent inhibitor AAEK2 quenches the reaction.

[0046] Fig. 11 depicts a schematic of adding evolved sortase to concentrated NW11 solution in order to multimerize NW11, followed by circularization. The oligomeric, circularized species containing variable numbers of NW11 assemble into nanodiscs of various sizes. Bottom: Negative-stain EM image showing large nanodiscs made using oligomeric, circularized NW11. Scale bar represents 100 nm. Fig. 11 discloses

"LEEKLNTQLPXTGSTFSKLRELEEKLNTQLPXTGSTFSKLRELEEKLNTQLPXTGSTFSKLRE" as SEQ ID NO: 88 and "LEEYTKKLNTQLPGTGSTFSKLRELEEKLNTQLPGTGSTFSKLRE" as SEQ ID NO: 89.

[0047] Figs. 12A-12E demonstrate that covalent circularization stabilizes MSPs against thermal unfolding without and with lipids, stabilizes embedded VDAC1 and allows control of the number of channels embedded. Thermal unfolding of MSP1D1 and cNW11 without (Fig. 12A) and with lipids (Fig. 12B) followed by circular dichroism (CD) spectroscopy at 222 nm, the wavelength most characteristic of helical secondary structure. POPC/POPG lipids at a molar ratio of 3:2 were used to generate nanodiscs. Fig. 12C demonstrates that placement in MSP1D1 nanodiscs increases the melting temperature of VDAC1 by 9.2 degrees over that of VDAC1 in an LDAO micelle environment, and covalent circularization of the scaffold protein (cNW11) raises T_m by additional 9.2 degrees. Thermal unfolding of human VDAC-1 was followed by CD spectroscopy at 218 nm, the wavelength most characteristic of β -sheet secondary structure. Orange: VDAC1 in 0.1% LDAO. Depicted are VDAC1 reconstituted into conventional nanodiscs (assembled using MSP1D1), and VDAC1 reconstituted into circularized nanodiscs (assembled using cNW11). Nanodiscs were made with POPC/POPG 3:2 lipids. Fig. 12D depicts thermal unfolding of VDAC1 reconstituted into circularized nanodiscs (assembled using cNW9) followed by CD spectroscopy at 218 nm. Nanodiscs were made with POPC/POPG lipids at a molar ratio of 3:2. All samples were in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl. Fig. 12E depicts analysis of the VDAC1 nanodisc assembly reaction. Top: size-exclusion chromatography and negative-stain EM of VDAC1 in cNW9 nanodiscs. Negative-stain image shows nanodiscs containing a single channel. The stain-filled channels appear as dark spots inside nanodiscs. Bottom: SDS-PAGE analysis of the nanodisc assembly. Fractions 1-6 were collected and analyzed.

[0048] Figs. 13A-13D demonstrate that covalent circularization stabilizes MSPs against digestion by V8 protease without and with lipids. Figs. 13A-13B depict SDS-PAGE analysis of the proteolysis of lipid-free MSP Δ H5 and cNW9. Samples were treated with V8 protease for 0 min (before addition of V8),

and for 1, 3 and 18 hours. Lanes are labeled with times of protease treatment. Proteolysis was performed in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl at 37°C and using a protein:protease ratio (w/w) of 1000:1. Treatment of MSPΔH5 with V8 resulted in the appearance of a large peptide, which is close in size to MSPΔH5 and is not discernible until about 18 hours after addition of V8; this peptide may be generated earlier but is not visible due to the low amount or overlap with uncleaved MSPΔH5. The intensity of the MSPΔH5 band decreased by 75% after 3 hours and by 93% after 18 hours. On the other hand, the intensity of the cNW9 band decreased by only 20% after 18 hours. A band that corresponds to linearized NW9 was observed after 3 hours and increased in intensity after 18 hours. Figs. 13C-13D depict SDS-PAGE analysis of the proteolysis of nanodiscs assembled with MSPΔH5 and cNW9. Samples were treated with V8 protease for 0 min (before addition of V8), and for 20 min, 1, 3 and 18 hours. Proteolysis was performed at 37°C at pH 7.5 in 20 mM Tris-HCl, 100 mM NaCl using a 100:1 protein:protease (w/w). The band intensity of MSPΔH5 decreased by 81% after 3 hours. There were no decreases in cNW9 band intensity up to 3 hours. ImageJ™ software was for analyzing band intensities.

[0049] Fig. 14 depicts characterization of cNW9 by MS/MS confirms the ligation of the C terminus to the N terminus. MS/MS spectrum of a tryptic fragment of cNW9 (SEQ ID NO: 91) showing the ligation of the C-terminal motif (LNTQLPGTG-His6 (SEQ ID NO: 55)) to the N-terminal residues (GSTFSK (SEQ ID NO: 52)). Expected masses for b and y ions along with the peptide sequence are listed in the table (SEQ ID NO: 91). The b and y ions that were identified in the MS/MS spectrum are highlighted. The full amino acid sequence of linear NW9 is shown at the top (SEQ ID NO: 90).

[0050] Fig. 15 depicts characterization of cNW30 by MS/MS confirms the ligation of the C terminus to the N terminus. MS/MS spectrum of a tryptic fragment of cNW30 (SEQ ID NO: 91) showing the ligation of the C-terminal motif (LNTQLPGTG-His6 (SEQ ID NO: 55)) to the N-terminal residues (GSTFSK (SEQ ID NO: 52)). Expected masses for b and y ions along with the peptide sequence are listed in the table (SEQ ID NO: 91). The b and y ions that were identified in the MS/MS spectrum are highlighted in. The full amino acid sequence of linear NW30 is shown at the top (SEQ ID NO: 92).

[0051] Fig. 16 depicts characterization of cNW50 by MS/MS. Spectrum of a tryptic fragment of cNW50 (SEQ ID NO: 91) showing the ligation of the C-terminal motif (LNTQLPGTG-His6 (SEQ ID NO: 55)) to the N-terminal residues (GSTFSK (SEQ ID NO: 52)). Expected masses for b and y ions along with the peptide sequence are listed in the table (SEQ ID NO: 91). The b and y ions that were positively identified in the MS/MS spectrum are highlighted. The amino acid sequence of linear NW50 (after TEV cleavage) is shown at the top (SEQ ID NO: 93).

[0052] Fig. 17 depict a diagram illustrating the production of circularized proteins or poly peptides using the two-intein system. The intein has a chitin binding domain (CBD) which allows it to bind to

chitin beads. The protein splicing of the inteins is then induced by adding DTT and changing the pH. The target protein then released from the chitin bound intein. The exposed N and C termini of the released protein interact together and result in the circularization of the protein.

[0053] Fig. 18 depicts chemically mediated ligation of scaffold proteins.

[0054] Fig. 19 depicts SDS-PAGE analysis of NW9_3C before (lane 1) and after covalent circularization (lane 2).

[0055] Figs 20A-20B depict phospholamban. Fig. 20A depicts the structure of phospholamban (PDB #1ZLL). Fig. 20B depicts the primary amino acid sequence of phospholamban from different species.

[0056] Figs. 21A-21B depict models of PLN nanodisc structures. Fig. 21A depicts a cartoon of MPER-Phospholamban (PLN)-TM hybrid in bilayer. Fig. 21B depicts MPER-Phospholamban hybrids displaying different number of MPER copies per complex.

[0057] Figs. 22A-22B depict an EM study of MPER-Phospholamban (PLN)-TM hybrid in complex with 10E8 Fab. Fig. 22A demonstrates that negative stain EM reveals multiple arms of 10E8 per complex. Fig. 22B depicts a cartoon for MPER-Phospholamban (PLN)-TM hybrid in complex with 10E8 Fab.

[0058] Figs. 23A-23C demonstrate MPER-Phospholamban (PLN)-TM interactions with antibodies. Fig. 23A depicts 2d class averages of MPER-Phospholamban (PLN)-TM hybrid in complex with 10E8 Fab. Fig. 23B depicts low resolution 3D reconstruction of the complex based on the negative-stain data. Fig. 23C depicts a cartoon of the interaction.

[0059] Fig. 24 depicts a schematic illustration of an example of strategy to produce immunogens incorporated in a nanodisc.

[0060] Figs. 25A-25B depict a schematic illustration of an example of strategy to produce trimer immunogen in nanodisc according to one embodiment of the present invention. gp140 could be covalently linked to MPER-PLB after (Fig. 25A) or before (Fig. 25B) incorporation into nanodisc.

[0061] Figs. 26A-26C depict a schematic illustration of an example of strategy to produce DNA immunogen comprises a DNA Duplex scaffold covalently linked to N terminus of MPER-phospholamban transmembrane fusion protein or 10E8 -phospholamban transmembrane fusion protein. DNA scaffold can be covalently linked to MPER-PLB after (Fig. 26A) or before (Fig. 26B) incorporation into nanodisc. (Fig. 26C) some duplex DNA designs that can be used for modifying 2, 3 or 5 MPER-PLB arms.

[0062] Fig. 27 demonstrates that a strong MPER-specific antibody response is generated to nanodisc delivered MPER-immunogen. MPER-peptide quantification by ELISA. Immulon 1B plates were pre-coated with streptavidin followed by washing, blocking and binding of biotinylated MPER-peptide.

Serially diluted sera collected 19 days following a third and final immunization were tested against naive mouse sera. Immunized mice received membrane presented MPER-phospholamban fusion protein displayed on nanodiscs (100ug of MPER-phospholamban:nanodisc per mouse). Mice were immunized three times at 3-week intervals via subcutaneous tail-base injection in conjunction with 5ug/mouse CDN adjuvant-loaded liposomes. Non-circularized (open circles) and circularized (filled circles) MSP nanodiscs result in identical IgG antibody titers. No MPER-specific IgG antibodies are detected in any naive (filled circle; dashed line) mice. Symbols represent mean with SEM shown. N=4 female BALB/c mice; 8 weeks old at time of primary immunization.

[0063] Fig. 28 demonstrates the immunogenicity of MPER-phospholamban transmembrane/nanodiscs. MPER- and human MSP-specific antibodies titer were measured by ELISA (left panel). 96 well plates were coated with MPER-phospholamban transmembrane in nanodisc (PLMPER/nanodiscs), empty nanodiscs made with human MSP (hMSP) or biotin-MPER, respectively. Right panel shows the frequency of MPER-specific bone marrow (BM) plasma cells at day 10 after the second immunization. Numbers of MPER-binding antibody-forming cells (AFCs) were determined by ELISPOT using liposomes containing MPER-TM, MPER-N (E662-T676) or MPER-C (A667-K683) as the capture antigen. When nanodiscs containing MPER-Phospholamban were administered to mice, high titer immune responses specific to both MPER and human MSP, the membrane scaffold protein that wrap around nanodisc, were generated after two immunizations (left panel)

DETAILED DESCRIPTION

[0064] The inventors have discovered methods of making nanodisc which permit previously unachievable sizes, uniformity, and geometry. This new approach to nanodisc construction involves the use of circularized membrane scaffold proteins, as opposed to the linear membrane scaffold proteins previously described. Provided herein are loopable membrane scaffold proteins, nanodiscs comprising such proteins, and methods of producing these nanodiscs. Further described herein are methods and compositions relating to the presentation of antigens on nanodiscs and/or liposome.

[0065] In one aspect, described herein is a vaccine scaffold polypeptide comprising at least the transmembrane domain of a single α -helix transmembrane protein, or a variant thereof. In some embodiments of any of the aspects, the vaccine scaffold polypeptide consists of the transmembrane domain of a single α -helix transmembrane protein, or a variant thereof. In some embodiments of any of the aspects, the vaccine scaffold polypeptide consists essentially of the transmembrane domain of a single α -helix transmembrane protein or a variant thereof.

[0066] In some embodiments of any of the aspects, a polypeptide described herein is engineered.

[0067] As used herein “single α -helix transmembrane protein” refers to a naturally-occurring protein which comprises one α -helix transmembrane domain. A single α -helix transmembrane protein can exist in situ as a monomer or a multimer. Single α -helix transmembrane proteins are well known in the art and can be readily identified, e.g., by structure and/or annotation searches of protein and sequence databases. Non-limiting examples of suitable single α -helix transmembrane protein can include phospholamban; influenza M2 channel; stannin; Vpu; HIV Vpu; parainfluenza fusion protein; parainfluenza 5 fusion protein; E protein; SARS-CoV E; SH protein; and RSV-SH.

[0068] As used herein “phospholamban” refers to a single α -helix transmembrane protein that forms a homopentameric integral membrane protein complex to regulate the Ca^{2+} pump. Sequences for phospholamban are known for a number of species, including, e.g., human (NCBI Gene ID: 5350) (mRNA, NCBI Seq Ref: NM_002667.4) (polypeptide, NCBI Ref Seq: NP_002658.1). In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the N-terminus of the transmembrane domain of phospholamban. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus of the transmembrane domain of phospholamban. In some embodiments of any of the aspects, the transmembrane domain of phospholamban can be the portion of the sequence of a phospholamban corresponding to SEQ ID NO: 113 or 114. In some embodiments of any of the aspects, the transmembrane domain of phospholamban can be the sequence of SEQ ID NO: 113 or 114.

[0069] As used herein, the influenza “M2 channel protein” or “matrix-2 (M2) protein” refers to a single α -helix transmembrane protein that forms a homotetramer ion channel in the viral envelop of influenza A. Tetramers comprising the transmembrane domain of M2 channel protein display a structure which widens right before the cytoplasmic side of the plasma membrane, permitting antigenic polypeptides to be spread more widely on the membrane surface. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus of the transmembrane domain of M2 channel protein. Sequences for M2 channel protein are known, e.g., (NCBI Gene ID: 956528 and 3655107) (polypeptide, NCBI Ref Seq: NP_040979.2 or YP_308853.1). In some embodiments of any of the aspects, the transmembrane domain of M2 channel protein can be the portion of the sequence of a M2 channel protein corresponding to SEQ ID NO: 95 or 96. In some embodiments of any of the aspects, the transmembrane domain of M2 channel protein can be the sequence of SEQ ID NO: 95 or 96.

SEQ ID NO: 95

LAIAA NIIGILHLIL WILDRL

SEQ ID NO: 96

LVVAA SIIGILHFIL WILDRL

[0070] As used herein, “stannin” refers to a single α -helix transmembrane protein that forms monomers which cross the plasma membrane at approximately a 20° angle. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus of the transmembrane domain of stannin. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the N-terminus of the transmembrane domain of stannin. Sequences for stannin are known, e.g., human stannin (NCBI Gene ID: 51061) (polypeptide, NCBI Ref Seq: NP_001290376.1, NP_001310951.1, NP_001310953.1, NP_001310954.1, and NP_056998.4) (mRNA, NCBI Ref Seq: NM_001303447.1, NM_001324022.1, NM_001324024.1, NM_001324025.1, and NM_015914.6). In some embodiments of any of the aspects, the transmembrane domain of stannin can be the portion of the sequence of a stannin corresponding to SEQ ID NO: 97. In some embodiments of any of the aspects, the transmembrane domain of stannin can be the sequence of SEQ ID NO: 97.

SEQ ID NO: 97

TGVVTVIVILIAIAALGALILGCW

[0071] As used herein, “Vpu” or “viral protein U” refers to a single α -helix transmembrane protein that forms pentameric ion channels which are wider at the C-terminus than the N-terminus. In some embodiments of any of the aspects, a Vpu can be an HIV Vpu. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus of the transmembrane domain of Vpu, e.g., HIV Vpu. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the N-terminus of the transmembrane domain of Vpu, e.g., HIV Vpu. Sequences for Vpu are known, e.g., HIV Vpu (NCBI Gene ID: 155945) (polypeptide, NCBI Ref Seq: NP_057855.1). In some embodiments of any of the aspects, the transmembrane domain of Vpu can be the portion of the sequence of a Vpu corresponding to SEQ ID NO: 98. In some embodiments of any of the aspects, the transmembrane domain of Vpu can be the sequence of SEQ ID NO: 98.

SEQ ID NO: 98

IVAIV ALVVAIIIAI VVWSIVIIIEY

[0072] As used herein, a “parainfluenza fusion protein” refers to a single α -helix transmembrane protein that forms trimeric complexes. In some embodiments of any of the aspects, the parainfluenza fusion protein can be a parainfluenza 5 fusion protein. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus

of the transmembrane domain of a fusion protein, e.g., parainfluenza 5 fusion protein. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the N-terminus of the transmembrane domain of a fusion protein, e.g., parainfluenza 5 fusion protein. Sequences for fusion proteins are known, e.g., parainfluenza 5 fusion protein (NCBI Gene ID: 3160801) (polypeptide, NCBI Ref Seq: YP_138515.1). In some embodiments of any of the aspects, the transmembrane domain of a parainfluenza fusion protein can be the portion of the sequence of a parainfluenza fusion protein corresponding to SEQ ID NO: 99. In some embodiments of any of the aspects, the transmembrane domain of a parainfluenza fusion protein can be the sequence of SEQ ID NO: 99.

SEQ ID NO: 99

VLSIIA ICLGSLGLIL IILLSVVVWK LLTIVVA

[0073] As used herein, a “SH” or “Small Hydrophobic” protein refers to a single α -helix transmembrane protein that forms pentameric complexes. In some embodiments of any of the aspects, the SH protein can be a Respiratory syncytial virus (RSV) SH protein. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus of the transmembrane domain of a SH protein, e.g., RSV SH. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the N-terminus of the transmembrane domain of a SH protein, e.g., RSV SH. Sequences for SH proteins are known, e.g., RSV SH (NCBI Gene ID: 1494473) (polypeptide, NCBI Ref Seq: NP_044594.1). In some embodiments of any of the aspects, the transmembrane domain of a SH protein can be the portion of the sequence of a SH protein corresponding to SEQ ID NO: 100. In some embodiments of any of the aspects, the transmembrane domain of a SH protein can be the sequence of SEQ ID NO: 100.

SEQ ID NO: 100

HMITTIISL LIISIMIAI L

[0074] As used herein, a “E” or “Envelope small membrane” protein refers to a single α -helix transmembrane protein that forms pentameric complexes. In some embodiments of any of the aspects, the E protein can be a SARS-Cov E protein. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the C-terminus of the transmembrane domain of an E protein, e.g., SARS-Cov E. In some embodiments of any of the aspects, an immunogenic conjugate polypeptide can comprise an antigenic polypeptide conjugated to the N-terminus of the transmembrane domain of an E protein, e.g., SARS-Cov E. Sequences for E proteins are known, e.g., SARS-Cov E (NCBI Gene ID: 1489671) (polypeptide, NCBI Ref Seq: NP_828854.1). In some embodiments of any of the aspects, the transmembrane domain of an E protein can be the portion of

the sequence of an E protein corresponding to SEQ ID NO: 101. In some embodiments of any of the aspects, the transmembrane domain of an E protein can be the sequence of SEQ ID NO: 101.

SEQ ID NO: 101

NSVLLF LAFVVFLVLT LAILTALRL

[0075] In some embodiments of any of the aspects, the vaccine scaffold polypeptide is less than 60 amino acids in length. In some embodiments of any of the aspects, the vaccine scaffold polypeptide is less than 50 amino acids in length. In some embodiments of any of the aspects, the vaccine scaffold polypeptide is less than 40 amino acids in length.

[0076] As used herein, the term “transmembrane domain” refers to a portion, part, or polypeptide sequence of a protein which forms an α -helix with a least one hydrophobic face. The transmembrane domain can target and/or anchor a translated polypeptide comprising the transmembrane domain to the plasma membrane of a host cell. The transmembrane domains of a given polypeptide sequence can be predicted and/or identified with any of various computer programs designed to analyze polypeptide sequences, or by consulting the annotation found in various publically-available databases, e.g., NCBI or PDB.

[0077] In some embodiments of any of the aspects, the transmembrane domain of phospholamban is the sequence of SEQ ID NO: 113 or 114. In some embodiments of any of the aspects, a variant of the transmembrane domain of phospholamban is a sequence with at least 90% homology to SEQ ID NO: 113 or 114. A variant of a given transmembrane domain can include sequences in which cysteine residues are substituted with another residue, e.g., serine. Such substitutions can improve the ease with which the vaccine scaffold polypeptide can be manipulated.

[0078] In one aspect of any of the embodiments, described herein is an immunogenic conjugate polypeptide comprising: a) at least one antigen polypeptide conjugated to; b) a vaccine scaffold polypeptide as described herein. In one aspect of any of the embodiments, described herein is an immunogenic conjugate polypeptide comprising: a) at least one antigen polypeptide conjugated to; b) a transmembrane domain of a single α -helix transmembrane protein. In one aspect of any of the embodiments, described herein is an immunogenic conjugate polypeptide consisting essentially of: a) at least one antigen polypeptide conjugated to; b) a vaccine scaffold polypeptide as described herein. In one aspect of any of the embodiments, described herein is an immunogenic conjugate polypeptide consisting essentially of: a) at least one antigen polypeptide conjugated to; b) a transmembrane domain of a single α -helix transmembrane protein. In one aspect of any of the embodiments, described herein is an immunogenic conjugate polypeptide consisting of: a) at least one antigen polypeptide conjugated to; b) a vaccine scaffold polypeptide as described herein. In one aspect of any of the embodiments, described

herein is an immunogenic conjugate polypeptide consisting of: a) at least one antigen polypeptide conjugated to; b) a transmembrane domain of a single α -helix transmembrane protein.

[0079] As used herein, an "antigen" is a molecule that is specifically bound by a B cell receptor (BCR), T cell receptor (TCR), and/or antibody, thereby activating an immune response. An antigen can be pathogen-derived, or originate from a pathogen. An antigen can be a polypeptide, protein, nucleic acid or other molecule or portion thereof. The term "antigenic determinant" refers to an epitope on the antigen recognized by an antigen-binding molecule, and more particularly, by the antigen-binding site of said molecule. An antigen polypeptide can be any antigenic polypeptide sequence.

[0080] In some embodiments of any of the aspects, the immunogenic conjugate polypeptide is less than 60 amino acids in length. In some embodiments of any of the aspects, the immunogenic conjugate polypeptide is less than 50 amino acids in length. In some embodiments of any of the aspects, the immunogenic conjugate polypeptide is less than 40 amino acids in length.

[0081] In some embodiments of any of the aspects, the antigen polypeptide is a viral protein or a bacterial protein or antigenic portion thereof. In some embodiments of any of the aspects, the antigen polypeptide is a protein or antigenic portion thereof, obtained from a pathogenic organism. In some embodiments of any of the aspects, the antigen polypeptide is a protein or antigenic portion thereof, specific to a pathogenic organism, e.g., as compared to related nonpathogenic and/or commensal organisms.

[0082] In some embodiments of any of the aspects, the antigen polypeptide is a viral envelope protein, a cell surface protein, or antigenic portion thereof. In some embodiments of any of the aspects, the antigen polypeptide is encoded by a viral env gene, e.g., the HIV env gene (NCBI Gene ID: 155971). Env genes and corresponding polypeptides for various virii are known in the art, e.g., in the NCBI database.

[0083] In some embodiments of any of the aspects, the antigen polypeptide is an HIV protein, or antigenic portion thereof. In some embodiments of any of the aspects, the antigen polypeptide is the membrane proximal external region (MPER) of Env, e.g., HIV Env. Exemplary MPER sequences for several HIV isolates are provided in Table 2.

[0084] Table 2: Sequence of MPER from different HIV isolates

HIV Strain	MPER sequence	SEQ ID NO:
HxB2	ELDKWASLWNWFNITNWLWYIK	102
SF162.LS	ELDKWASLWNWFDISKWLWYIK	103
6535.3	ELDKWGSLWNWFSISNWLWYIR	104
YU2.DG	ALDKWASLWNWFDITKWLWYIK	105
PVO.4	ALDKWESLWNWFSITKWLWYIK	106
RHPA.4259.7	ALDKWASLWSWFSITHWLWYIK	107
TRO.11	ELDS WASLWNWFDISKWLWYIK	108

WITO.4160.33	ELDKWASLWNWFNITNWLWYIK	109
REJO.4541.67	ELDKWASLWNWFDITNWLWYIK	110
AC10.0.29	ALDKWANLWNWFNITEWLWYIK	111

[0085] In some embodiments of any of the aspects, the antigen polypeptide is the 10E8 epitope (SLWNWFNITNWLWYIK; SEQ ID NO: 112) or the 4E10 epitope (NWFNITNWLW; SEQ ID NO: 94) of HIV Env.

[0086] In some embodiments, the antigen polypeptide can be the ectodomain of HIV gp160 (e.g., NCBI Gene ID: 155971), influenza HA (e.g., NCBI Gene ID: 956529, 3654620, 3655151, 23308115, 23308115, 1460996, 23104227, or 3655103), hepatitis C E1, or hepatitis C E2 protein (see, e.g., NCBI Gene ID: 951475) or an antigenic portion thereof.

[0087] The term “conjugated” refers to the attachment of at least two entities to form one entity. The joining of the two entities can be direct (e.g., via covalent or non-covalent bonds) or indirect (e.g., via linkers etc.). Conjugation can be by means of linkers, chemical modification, peptide linkers, chemical linkers, covalent or non-covalent bonds, or protein fusion or by any means known to one skilled in the art. The joining can be permanent or reversible. In some embodiments of any of the aspects, the antigen polypeptide is conjugated to the N-terminus of the vaccine scaffold polypeptide. In some embodiments of any of the aspects, the antigen polypeptide is conjugated to the C-terminus of the vaccine scaffold polypeptide.

[0088] An immunogenic conjugate polypeptide can comprise more than one occurrence of an antigen polypeptide, e.g., multiple repeats of the same antigen polypeptide and/or multiple different antigen polypeptides. In some embodiments of any of the aspects, the immunogenic conjugate polypeptide comprises multiple repeats of the same antigen polypeptide. In some embodiments of any of the aspects, the immunogenic conjugate polypeptide comprises multiple different antigen polypeptides.

[0089] In some embodiments of any of the aspects, multiple antigen polypeptides or repeats thereof are separated by a linker. In some embodiments of any of the aspects, the antigen polypeptide and the vaccine scaffold polypeptide are joined by a linker. The term “linker” refers to any means, entity or moiety used to join two or more entities. Linker moieties include, but are not limited to, chemical linker moieties, or for example a peptide linker moiety. In some embodiments of any of the aspects, the linker is hydrophobic. In some embodiments of any of the aspects, the linker is hydrophilic. In some embodiments of any of the aspects, the linker is from 1 to 5 amino acids in length, e.g., is 1, 2, 3, 4, or 5 amino acids in length.

[0090] In some embodiments of any of the aspects, the immunogenic conjugate polypeptide comprises from 1 to 5 antigen polypeptides, e.g., 1 to 5 repeats of the same antigen polypeptide and/or 1 to 5 different antigen polypeptides.

[0091] In one aspect of any of the embodiments, described herein is a nucleic acid encoding the vaccine scaffold polypeptide as described herein or the immunogenic conjugate polypeptide as described herein. In some embodiments of any of the aspects, described herein is a vector comprising a nucleic acid encoding the vaccine scaffold polypeptide as described herein or the immunogenic conjugate polypeptide as described herein.

[0092] The vaccine scaffold polypeptides and/or immunogenic conjugate polypeptides described herein can exist as multimers, e.g., as polypeptide complexes. The potential for multimerization can be controlled by the vaccine scaffold polypeptide selection, e.g., phospholamban-derived vaccine scaffold polypeptides will provide vaccine scaffold polypeptides and/or immunogenic conjugate polypeptides which form pentamers. In a multimer of immunogenic conjugate polypeptides, the antigen polypeptide(s) of each immunogenic conjugate polypeptide can be the same or differ. In some embodiments of any of the aspects, the individual vaccine scaffold polypeptides and/or immunogenic conjugate polypeptides of a polypeptide complex are not covalently bound to each other.

[0093] In one aspect of any of the embodiments, described herein is a composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc. In some embodiments of any of the aspects, the phospholipid bilayer comprises a viral membrane lipid; anionic lipid; or synthetic lipid. In some embodiments of any of the aspects, the phospholipid bilayer comprises cholesterol. Suitable lipids discussed in further detail below herein.

[0094] In some embodiments of any of the aspects, each liposome or nanodisc comprises 10 or fewer copies of the antigen polypeptide. In some embodiments of any of the aspects, each liposome or nanodisc comprises 5 or fewer copies of the antigen polypeptide.

[0095] In one aspect of any of the embodiments, described herein is a pharmaceutical composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc and a pharmaceutically acceptable carrier. In some embodiments of any of the aspects, the pharmaceutical composition can further comprise an adjuvant.

[0096] As used herein in the context of immunization, immune response and vaccination, the term “adjuvant” refers to any substance than when used in combination with a specific antigen that produces a more robust immune response than the antigen alone. When incorporated into a vaccine formulation, an

adjuvant acts generally to accelerate, prolong, or enhance the quality of specific immune responses to the vaccine antigen(s). Non-limiting examples of adjuvants include alum, paraffin oil, killed bacteria, squalene, detergents, plant saponins, cytokines, Freund's complete adjuvant, Freund's incomplete adjuvant, adjuvant 65, aluminum hydroxide, aluminum phosphate, and/or calcium phosphate hydroxide.

[0097] Described herein are covalently circularized nanodiscs and the loopable membrane scaffold proteins used to construct these nanodiscs. As used herein, "nanodisc" refers to a discoidal, nanoscale phospholipid bilayer which is "belted" or "ringed" by a membrane scaffold protein. The membrane scaffold protein comprises amphipathic alpha helices which provide an hydrophobic surface next to the hydrocarbon tails of the phospholipid bilayer and an external hydrophilic surface. As used herein, "covalently circularized nanodisc" refers to a nanodisc in which the membrane scaffold protein has been covalently bonded to itself to form continuous polypeptide loop (i.e. it has been circularized).

[0098] In one aspect of any of the embodiments, described herein is a loopable membrane scaffold protein comprising, from N-terminus to C-terminus: i) an N-terminal circularization domain; ii) at least one amphipathic alpha helix domain; and iii) a C-terminal circularization domain. In one aspect of any of the embodiments, described herein is a loopable membrane scaffold protein comprising, from N-terminus to C-terminus: i) an N-terminal circularization domain; ii) a plurality of amphipathic alpha helix domains; and iii) a C-terminal circularization domain. As used herein, "loopable membrane scaffold protein" refers to a protein which is capable of being circularized by a covalent bond in such a way that the interior circumference of the protein provides a hydrophobic environment and the exterior circumference forms a hydrophilic environment. Once circularized, the protein is referred to herein as a "looped membrane scaffold protein." This circularized protein can accommodate a phospholipid bilayer within the interior circumference.

[0099] As used herein, "circularized" or "looped" when used in reference to a polypeptide refers to the fact that the peptide sequence is not linear in nature, e.g., it does not have an N-terminus or C-terminus. A polypeptide which is circularized or looped can form any shape, e.g., a circle, an oval, or a polygon. In contrast, a polypeptide or nanodisc referred to herein which is "circular in shape," has a shape or cross-section which forms the shape of a circle.

[0100] As used herein, "a circularization domain" refers to a section of a linear polypeptide which upon interaction with a corresponding second circularization domain in the same polypeptide, can cause circularization of the polypeptide. The circularization domain can comprise one or more amino acid residues. A number of circularization technologies are known in the art and any such technology can be applied to the proteins described herein. Exemplary circularization domains are described below.

[00101] Proteins can be circularized by the action of a sortase enzyme. Sortase recognizes a LPXTG/A (e.g., LPXTG (SEQ ID NO: 51) or LPXTA (SEQ ID NO: 56)) sequence, cleaving the sequence after the threonine. The threonine residue can then form a new peptide bond with a glycine or alanine residue found at a second, more N-terminal location in the polypeptide, causing it to be circularized. The use of sortase enzymes to circularize polypeptide is described in more detail in, e.g., Cowper et al. *ChemBioChem* 2013 14:809-812; Antos et al., *Journal of Biological Chemistry* 2009 284:16028-36; and Tsukiji et al. *ChemBioChem* 2009 10:787-798; each of which is incorporated by reference herein in its entirety. In some embodiments of any of the aspects described herein, one circularization domain comprises: a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and the other circularization domain comprises: a sequence LPXTG/A; wherein X represents any amino acid. In some embodiments of any of the aspects described herein, the N-terminal circularization domain comprises: a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and the C-terminal circularization domain comprises: a sequence LPXTG/A; wherein X represents any amino acid.

[00102] As used herein, “non-polar amino acid sequence” refers to a sequence comprising non-polar amino acids, the sequence being at least one amino acid in length. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to 20 amino acids in length. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to about 10 amino acids in length. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to 10 amino acids in length. In some embodiments of any of the aspects, the non-polar amino acid sequence comprises glycine and alanine residues. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of glycine and alanine residues. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to 10 amino acids in length and comprises amino acids selected from the group consisting of: glycine and alanine. In some embodiments of any of the aspects, the non-polar amino acid sequence is from 1 to 10 amino acids in length and consists of amino acids selected from the group consisting of: glycine and alanine. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 residues selected from glycine and alanine. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 glycine residues. In some embodiments of any of the aspects, the non-polar amino acid sequence consists of 1-5 alanine residues.

[00103] In some embodiments of any of the aspects, the sequence LPXTG/A can be covalently bound to the at least one glycine or alanine residue of (i).

[00104] In some embodiments of any of the aspects, the LPXTG/A is LPGTG/A (SEQ ID NO: 24); LPSTG/A (SEQ ID NO: 25); or LPETG/A (SEQ ID NO: 26). In some embodiments of any of the aspects,

the LPXTG/A is LPGTG (SEQ ID NO: 53); LPGTA (SEQ ID NO: 57); LPSTG (SEQ ID NO: 58); LPSTA (SEQ ID NO: 59); LPETG (SEQ ID NO: 60); or LPETA (SEQ ID NO: 61).

[00105] Proteins can also be circularized by use of two intein domains (e.g., using the pTWIN vectors commercially available from New England Biolabs; Ipswich, MA (e.g., Cat. No. N6951S)). Inteins can excise themselves from a polypeptide, causing splicing of the remaining sequence (e.g., exteins). When inteins are used to flank a central sequence, the central sequence can be circularized when the inteins excise themselves. Various inteins and their use are described in more detail in, e.g., Elleuche et al. *Applied Microbiology and Biotechnology*. 2010 87:479-489; Cowper et al. *ChemBioChem* 2013 14:809-812; Ahlmann-Eltze et al. 2015 hdl.handle.net/1721.1/96071; Evans et al. *Journal of Biological Chemistry* 1999 274:18359-18363; and Evans et al. *Biopoly* 1999 51: 333-342; each of which is incorporated by reference herein in its entirety.

[00106] In some embodiments of any of the aspects, one circularization domain comprises an intein separated from the domain of (ii) by a Cys; and the other circularization domain comprises an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein. In some embodiments of any of the aspects, the N-terminal circularization domain comprises an intein separated from the domain of (ii) by a Cys; and the C-terminal circularization domain comprises an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein. In some embodiments of any of the aspects, at least one intein is flanked by a chitin binding domain (CBD). Intein-mediated cyclization is described further in, e.g., Xu et al. *Methods*. 2001 Jul;24(3):257-77; which is incorporated by reference herein in its entirety.

[00107] Proteins can also be circularized by utilizing 2 cysteine residues, one at each end of the sequence to be circularized. The C-terminal cysteine is capable of initiating an intramolecular N→S acyl shift. This results in an S-acyl intermediate that can be intercepted by an added thiol, yielding a C-terminal thioester and liberating cysteine. The cysteine at the N terminus interacts with the C-terminal thioester and leads to spontaneous backbone circularization, through intramolecular transthioesterification. Next, an S→N acyl shift rearrangement occurs that lead to formation of a peptide bond. The final circularized product will have a cysteine next to the newly formed peptide bond. Such methods of circularization are described in more detail at, e.g., Cowper, B et. al. *ChemBioChem* 2013, 14, 809 – 812; which is incorporated by reference herein in its entirety. In some embodiments of any of the aspects, each circularization domain comprises a Cys. In some embodiments of any of the aspects, one of the Cys residues is preceded by a Gly residue. In some embodiments of any of the aspects, one of the Cys residues is the C-terminal residue of the linear polypeptide.

[00108] As used herein, “amphipathic alpha helix domain” refers to a domain comprising a continuous amphipathic alpha helix, e.g., an alpha helix with a face which is hydrophobic/non-polar and a second face which is hydrophilic/polar. In some embodiments of any of the aspects, a loopable membrane scaffold protein can comprise a plurality of amphipathic alpha helix domains, e.g., multiple domains separated by non-helical sequences. Several exemplary amphipathic alpha helix domains are provided herein. By way of non-limiting example, an amphipathic alpha helix can comprise an ApoA-I polypeptide or repeats thereof; at least one alpha helical polypeptide with secondary amphipathicity or a repeat thereof; at least one amphipathic proline-rich polypeptide or a repeat thereof; or at least one repeat of the sequence XZXPPP; wherein X is any hydrophobic residue (e.g., any of A, I, L, F, V, P, W, M, Y and G) and Z is any hydrophilic residue (e.g., any of R, K, D, E, Q, N, H, S, T, or C). A given amphipathic alpha helix domain and/or loopable membrane scaffold protein can comprise a combination of any of the specific sequences and domains described herein.

[00109] ApoA-I (e.g., NCBI Gene ID: 335), also referred to as ApoA1, refers to the major protein component of high density lipoprotein (HDL) in plasma, which also serves a cofactor for lecithin cholesterolacyltransferase (LCAT). The sequences for ApoA-I are known for a number of species, e.g., human ApoA-I mRNA (e.g., NCBI Ref Seq: NM_000039.2) and protein (e.g., NCBI Ref Seq: NP_000030.1; SEQ ID NO: 5) sequences. The native ApoA-I protein comprises a 43 residue N-terminal globular domain and a 200 residue C-terminal lipid binding domain (e.g., residues 68-267 of SEQ ID NO: 5) which forms an amphipathic alpha helix. As used herein, “ApoA-I polypeptide” refers to a polypeptide comprising at least one repeat of the ApoA-I C-terminal lipid binding domain (e.g., residues 68-267 of SEQ ID NO: 5). In some embodiments of any of the aspects, an ApoA-I polypeptide comprises from 1 to 10 repeats of the ApoA-I C-terminal lipid binding domain (e.g., residues 68-267 of SEQ ID NO: 5).

[00110] As used herein, “secondary amphipathicity” refers to amphipathicity existing in the secondary structure of a polypeptide, e.g., the polar residues are found on one face of the secondary structure and the nonpolar residues are found on the opposite face of the secondary structure.

[00111] In some embodiments of any of the aspects described herein, an alpha helical polypeptide with secondary amphipathicity can have a sequence selected from the group consisting of:

DWFKAFYDKLA EKFK EAF (SEQ ID NO: 27);

DFLKAFYDKVA EKFK EAAPDWFKAFYDKVA EKFK EAF (SEQ ID NO: 28);

KLALRLALKAFKAALKLA (SEQ ID NO: 29); RLALDLALRAFKAAWKLA (SEQ ID NO: 30);

ELAWDLAFEALDAELKLD (SEQ ID NO: 31); FLKLLKKFLKLFKKLLKLF (SEQ ID NO: 32);

WLKLLKKWLKLWKKLLKL (SEQ ID NO: 33); LLKFLKRLKLLKDLWKLL (SEQ ID NO: 34);

WEAAFAEALAEAWAEHLAEALAEAVEALAA (SEQ ID NO: 35);

WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 36);
 RAFARALARALKALARALKALAR (SEQ ID NO: 37); GLFEALLELLESLWELLLEA (JTS1
 peptide) (SEQ ID NO: 38); KLLKLLLKLWKKLLKLLK (SEQ ID NO: 39);
 DWLKAFYDKVAEKLKEAF (18A peptide) (SEQ ID NO: 40); DWFKAFYDKVAEKFKKEAF (4F
 peptide) (SEQ ID NO: 41); or variants thereof. Alpha helical polypeptide with secondary amphipathicity
 are readily identified by one of skill in the art, e.g., the Helical Wheel Prediction algorithm available at
 rzlab.ucr.edu/scripts/wheel/ can be used to characterize the secondary amphipathicity of a query
 polypeptide sequence. Secondary amphipathic peptides are described further in the art at, e.g., S.
 Gottschalk et al., Gene Ther. 1996; 3: 448-457; T. Niidome et al. Bioconjugate Chem. 1999; 10(5): 773-
 780); B. Chung et al., J. Bio. Chem. 1985; 260:10256-10262; and H. Kariyazono et al. J. Pept. Sci. 2016;
 22: 116-122; each of which is incorporated by reference herein in its entirety.

[00112] The loopable membrane scaffold proteins described herein can further comprise sequences
 useful for preparation of the covalently circularized nanodiscs, but which will not be present in the
 covalently circularized nanodisc itself. For example, sequences flanking the N- or C-terminal
 circularization domains can be used to purify the loopable membrane scaffold proteins and/or to tether
 them to a substrate or surface for the circularization process. Circularization itself will necessarily
 remove such flanking sequences from the membrane scaffold protein as it is looped at the circularization
 domains.

[00113] In some embodiments of any of the aspects, the loopable membrane scaffold protein can
 comprise a cap sequence N-terminal of the N-terminal circularization domain that can be cleaved to
 expose the N-terminal circularization domain as the N-terminus of the protein. In some embodiments of
 any of the aspects, the loopable membrane scaffold protein can comprise a cap sequence N-terminal of
 the N-terminal circularization domain comprising a non-polar amino acid sequence that can be cleaved to
 expose the N-terminal circularization domain a non-polar amino acid sequence as the N-terminus of the
 protein. Such cap sequences can comprise and be cleaved by means of any sequence recognized by a
 protease. By way of non-limiting example, the cap sequence can comprise the recognition sequence Leu-
 Val-Pro-Arg-Gly-Ser (SEQ ID NO: 42). This recognition sequence can be cleaved by thrombin between
 Arg and Gly. By way of further non-limiting example, the cap sequence can comprise the recognition
 sequence Glu-X-Leu-Tyr-Φ-Gln-φ where X is any residue, Φ is any large or medium hydrophobic residue
 (e.g., W, F, M, Y, L, I, or V) and φ is glycine or alanine (SEQ ID NO: 43). Glu-X-Leu-Tyr-Φ-Gln-φ
 (SEQ ID NO: 43) is recognized by TEV protease and cleaved after the Gln residue.

[00114] In some embodiments of any of the aspects, the cap sequence, or any sequence flanking the
 circularization domains can further comprise an affinity purification tag. In some embodiments of any of

the aspects, the cap sequence can further comprise an affinity purification tag N-terminal of the cleavage site. In some embodiments of any of the aspects, the loopable membrane scaffold protein can comprise an affinity tag flanking one of the circularization domains. In some embodiments of any of the aspects, the loopable membrane scaffold protein can comprise an affinity tag C-terminal of the C-terminal circularization domain. In some embodiments of any of the aspects, the loopable membrane scaffold protein can comprise an affinity tag N-terminal of the N-terminal circularization domain.

[00115] Affinity tags can be used for, e.g., affinity purification or to bind a protein to a desired substrate or surface. Affinity tags can be used to bind the loopable membrane scaffold tag to a surface or substrate prior to circularization, or to remove flanking sequences (e.g. cleaved N-terminal cap sequences) after they have been cleaved. A number of affinity tags and their binding partner ligands are well known in the art and are described, e.g., in Lichty et al. *Protein Expr Purif* 2005 41:98-105; Zhao et al. *J Analytical Methods in Chemistry* 2013; Kimple et al. *Current Protocols in Protein Science* 2004 36:939:9.1-9.9.19; and Giannone et al. *Methods and Protocols "Protein Affinity Tags"* Humana Press 2014; each of which is incorporated by reference herein in its entirety. Non-limiting examples of affinity tags can include His tags, Flag tags, or Strep tags.

[00116] Affinity tags can be separated from a circularization domain by a linker. Accordingly, a linker can be located C-terminal of the C-terminal circularization domain and N-terminal of an affinity tag, or N-terminal of the N-terminal circularization domain and C-terminal of an affinity tag.

[00117] Illustrative examples of linkers include glycine polymers (G)_n; glycine-serine polymers (G₁₋₅S₁₋₅)_n (SEQ ID NO: 63), where n is an integer of at least one, two, three, four, or five; glycine-alanine polymers; alanine-serine polymers; and other flexible linkers known in the art. Glycine and glycine-serine polymers are relatively unstructured, and therefore may be able to serve as a neutral tether between domains of the proteins described herein. Glycine accesses significantly more phi-psi space than even alanine, and is much less restricted than residues with longer side chains (*see* Scheraga, *Rev. Computational Chem.* 11173-142 (1992)). The ordinarily skilled artisan will recognize that design of a protein in particular embodiments can include linkers that are all or partially flexible, such that the linker can include a flexible linker as well as one or more portions that confer less flexible structure to provide for a desired nanodisc structure.

[00118] Other exemplary linkers include, but are not limited to the following amino acid sequences: GGG; DGGGS (SEQ ID NO: 64); TGEKP (SEQ ID NO: 65) (*see, e.g.,* Liu et al., *PNAS* 5525-5530 (1997)); GGRR (SEQ ID NO: 66) (Pomerantz et al. 1995, *supra*); (GGGS)_n wherein n = 1, 2, 3, 4 or 5 (SEQ ID NO: 67) (Kim et al., *PNAS* 93, 1156-1160 (1996.)); EGKSSGSGSESKVD (SEQ ID NO: 68) (Chaudhary et al., 1990, *Proc. Natl. Acad. Sci. U.S.A.* 87:1066-1070); KESGSVSSEQLAQFRSLD

(SEQ ID NO: 69) (Bird *et al.*, 1988, Science 242:423-426), GGRRGGGS (SEQ ID NO: 70); LRQRDGERP (SEQ ID NO: 71); LRQKDGGGSERP (SEQ ID NO: 72); LRQKD(GGGS)₂ ERP (SEQ ID NO: 73). Alternatively, flexible linkers can be rationally designed using a computer program capable of modeling both DNA-binding sites and the peptides themselves (Desjarlais & Berg, *PNAS* 90:2256-2260 (1993), *PNAS* 91:11099-11103 (1994) or by phage display methods. In one embodiment, the linker comprises the following amino acid sequence: GSTSGSGKPGSGEGSTKG (SEQ ID NO: 74) (Cooper *et al.*, *Blood*, 101(4): 1637-1644 (2003)).

[00119] In some embodiments of any of the aspects, a covalently circularized nanodisc can, when viewed from an angle which is perpendicular to a hydrophilic surface of the phospholipid layer, be circular in shape. In some embodiments of any of the aspects, a covalently circularized nanodisc can, when viewed from an angle which is perpendicular to a hydrophilic surface of the phospholipid layer, be polygonal (e.g., triangular, rectangular, pentagonal, etc) in shape.

[00120] In some embodiments of any of the aspects, the shape of the nanodisc can be directed by the ratio of phospholipids to looped scaffold membrane protein. In some embodiments of any of the aspects, the shape of the nanodisc can be directed by the structure of the looped scaffold membrane protein. A looped scaffold membrane protein comprising a single amphipathic alpha helix domain or multiple amphipathic alpha helix domains which are joined by sequences with constrained geometry will tend to form nanodiscs which, when viewed from an angle which is perpendicular to a hydrophilic surface of the phospholipid layer, are circular in shape.

[00121] In some embodiments of any of the aspects, the loopable scaffold membrane protein can comprise a plurality of amphipathic alpha helix domains separated by flexible linker sequences. Each linker sequence will tend to allow a sharper bend to be assumed by a looped scaffold membrane protein at that location. Accordingly, such variants of the looped scaffold membrane protein can, when viewed from an angle which is perpendicular to a hydrophilic surface of the phospholipid layer, be polygonal in shape. The particular polygon assumed by the assembled nanodisc can be directed by the number of separate flexible linker sequences present in the looped scaffold membrane protein, with each forming a corner of a polygon. Linker sequences are known to those of skill in the art and can be any sequence of less than about 30 amino acids in length which do not assume a secondary structure. In some embodiments of any of the aspects, the linker sequence can be less than 30 amino acids in length. In some embodiments of any of the aspects, the linker sequence can be less than about 25 amino acids in length. In some embodiments of any of the aspects, the linker sequence can be less than about 20 amino acids in

length. In some embodiments of any of the aspects, the linker sequence can be less than 20 amino acids in length.

[00122] In some embodiments of any of the aspects, the flexible linker sequence is selected from the group consisting of: LPGTGS (SEQ ID NO: 44); LPGTSG (SEQ ID NO: 45); LPSTGS (SEQ ID NO: 46); and LPSTSG (SEQ ID NO: 47).

[00123] In some embodiments of any of the aspects, the loopable scaffold membrane protein has the sequence of any of SEQ ID NOs: 1-4 and 6-23.

[00124] Mutation of certain residues to cysteine can, e.g., permit immobilizing nanodiscs to surfaces and/or solid substrates, e.g., by surface plasmon resonance or with use of resins or beads configured to bind to cysteine residues. Provision of cysteine residues can further permit attachment of tags, e.g., small molecule or protein tags for NMR or crystallization applications). Accordingly, In some embodiments of any of the aspects, the loopable scaffold membrane protein can comprise substitution with cysteine at one or more residues.

[00125] In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at one or more residues corresponding to positions 31, 36, 137, and 153 of SEQ ID NO: 1 or 6, or positions 31, 36, 93, 110, 159, and 175 of SEQ ID NO: 2 or 7. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at one residue corresponding to positions 31, 36, 137, and 153 of SEQ ID NO: 1 or 6. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at two residues corresponding to positions 31, 36, 137, and/or 153 of SEQ ID NO: 1 or 6. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at three residues corresponding to positions 31, 36, 137, and/or 153 of SEQ ID NO: 1 or 6. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at the four residues corresponding to positions 31, 36, 137, and 153 of SEQ ID NO: 1 or 6.

[00126] In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at the residue corresponding to positions 31, 36, 93, 110, 159, or 175 of SEQ ID NO: 2 or 7. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at the two residues corresponding to positions 31, 36, 93, 110, 159, and/or 175 of SEQ ID NO: 2 or 7. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at the three residues corresponding to positions 31, 36, 93, 110, 159, and/or 175 of SEQ ID NO: 2 or 7. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a

substitution with cysteine at the four residues corresponding to positions 31, 36, 93, 110, 159, and/or 175 of SEQ ID NO: 2 or 7. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at the five residues corresponding to positions 31, 36, 93, 110, 159, and/or 175 of SEQ ID NO: 2 or 7. In some embodiments of any of the aspects, the loopable scaffold membrane protein can further comprise a substitution with cysteine at the six residues corresponding to positions 31, 36, 93, 110, 159, and 175 of SEQ ID NO: 2 or 7.

[00127] In some embodiments of any of the aspects, the loopable scaffold membrane protein can comprise a sequence of SEQ ID NO: 1 or 6 with a substitution with cysteine at one or more residues corresponding to positions 31, 36, 137, and 153 of SEQ ID NO: 1 or 6. In some embodiments of any of the aspects, the loopable scaffold membrane protein can comprise a sequence of SEQ ID NO: 2 or 7 with a substitution with cysteine at one or more residues corresponding to positions 31, 36, 93, 110, 159, and 175 of SEQ ID NO: 2 or 7.

[00128] In some embodiments of any of the aspects, the loopable scaffold membrane protein can have or comprise the sequence of SEQ ID NO: 62.

SEQ ID NO: 62 NW11_3C

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWCNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGECMRDRARAHVDALRTH
LAPYSDELQRRLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALCDLRQGLLPVLESFK
VSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00129] In one aspect of any of the embodiments, described herein is a nucleic acid encoding a loopable membrane scaffold protein as described herein. In one aspect of any of the embodiments, described herein is a vector comprising a nucleic acid encoding a loopable membrane scaffold protein as described herein.

[00130] The loopable scaffold membrane proteins described herein can be circularized and combined with phospholipids to form a covalently circularized nanodisc. In one aspect of any of the embodiments, described herein is a covalently circularized nanodisc comprising: a) a phospholipid bilayer; and b) a looped membrane scaffold protein forming a belt around the bilayer, wherein the looped membrane scaffold protein comprises: i) an N-terminal circularization domain covalently linked to a C-terminal circularization domain; and ii) at least one amphipathic alpha helix domains. In one aspect of any of the embodiments, described herein is a covalently circularized nanodisc comprising: a) a phospholipid bilayer; and b) a looped membrane scaffold protein forming a belt around the bilayer, wherein the looped membrane scaffold protein comprises: i) an N-terminal circularization domain covalently linked to a C-terminal circularization domain; and ii) a plurality of amphipathic alpha helix domains.

[00131] Suitable exemplary circularization domains and methods of circularizing proteins comprising them are described elsewhere herein. The looped proteins that result from use of these exemplary circularization domains are described herein.

[00132] In some embodiments of any of the aspects, a loopable membrane scaffold protein can comprise one circularization domain comprising: a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and a second circularization domain comprising: a sequence LPXTG/A; wherein X represents any amino acid. In the resulting looped membrane scaffold protein a first circularization domain covalently linked to a second circularization domain comprises the sequence LPXT followed by a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; wherein X represents any amino acid. In some embodiments of any of the aspects, the sequence LPXT is LPGT (SEQ ID NO: 48); LPST (SEQ ID NO: 49); or LPET (SEQ ID NO: 50).

[00133] In some embodiments of any of the aspects, a loopable membrane scaffold protein can comprise one circularization domain comprising an intein separated from the domain of (ii) by a Cys; and a second circularization domain comprising an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein. In the resulting looped membrane scaffold protein a first circularization domain covalently linked to a second circularization domain comprises a cysteine residue.

[00134] In some embodiments of any of the aspects, each circularization domain of a loopable membrane scaffold protein can comprise a Cys residue. In some embodiments of any of the aspects, one of the Cys residues is preceded by a Gly residue. In some embodiments of any of the aspects, one of the Cys residues is the C-terminal residue of the linear polypeptide. In the resulting looped membrane scaffold protein a first circularization domain covalently linked to a second circularization domain comprises a cysteine residue.

[00135] In some embodiments of any of the aspects, the diameter of the nanodisc is greater than about 6 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is greater than 6 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is greater than about 15 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is greater than 15 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is greater than about 50 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is greater than 50 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is about 80 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is 80 nm.

[00136] In some embodiments of any of the aspects, the diameter of the nanodisc is about 6 nm to about 80 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is 6 nm to 80 nm.

In some embodiments of any of the aspects, the diameter of the nanodisc is about 15 nm to about 80 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is 15 nm to 80 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is about 50 nm to about 80 nm. In some embodiments of any of the aspects, the diameter of the nanodisc is 50 nm to 80 nm.

[00137] The nanodiscs described herein comprise phospholipids. As used herein, the term "phospholipid" refers to phosphatidic acids, phosphoglycerides, and phosphosphingolipids. Phosphatidic acids include a phosphate group coupled to a glycerol group, which may be mono- or diacylated. Phosphoglycerides (or glycerophospholipids) include a phosphate group intermediate an organic group (e.g., choline, ethanolamine, serine, inositol) and a glycerol group, which may be mono- or diacylated. Phosphosphingolipids (or sphingomyelins) include a phosphate group intermediate an organic group (e.g., choline, ethanolamine) and a sphingosine (non-acylated) or ceramide (acylated) group. It will be appreciated that in certain embodiments, the phospholipids useful in the compositions and methods of the invention include their salts (e.g., sodium, ammonium). For phospholipids that include carbon-carbon double bonds, individual geometrical isomers (cis, trans) and mixtures of isomers are included.

[00138] Representative phospholipids include phosphatidylcholines, phosphatidylethanolamines, phosphatidylglycerols, phosphatidylserines, phosphatidylinositols, and phosphatidic acids, and their lysophosphatidyl (e.g., lysophosphatidylcholines and lysophosphatidylethanolamine) and diacyl phospholipid (e.g., diacylphosphatidylcholines, diacylphosphatidylethanolamines, diacylphosphatidylglycerols, diacylphosphatidylserines, diacylphosphatidylinositols, and diacylphosphatidic acids) counterparts.

[00139] The acyl groups of the phospholipids may be the same or different. In certain embodiments, the acyl groups are derived from fatty acids having C10-C24 carbon chains (e.g., acyl groups such as lauroyl, myristoyl, palmitoyl, stearoyl or oleoyl groups). Representative diacylphosphatidylcholines (i.e., 1,2-diacyl-sn-glycero-3-phosphocholines) include distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dilinoleoylphosphatidylcholine (DLPC), palmitoyl-oleoylphosphatidylcholine (POPC), palmitoyl-linoleoylphosphatidylcholine, stearoyl-linoleoylphosphatidylcholine, stearoyl-oleoylphosphatidylcholine, stearoyl-arachidoylphosphatidylcholine, didecanoylphosphatidylcholine (DDPC), dierucoylphosphatidylcholine (DEPC), dilinoleoylphosphatidylcholine (DLOPC), dimyristoylphosphatidylcholine (DMPC), myristoyl-palmitoylphosphatidylcholine (MPPC), myristoyl-stearoylphosphatidylcholine (MSPC), stearoyl-myristoylphosphatidylcholine (SMPC), palmitoyl-myristoylphosphatidylcholine (PMPC),

palmitoylstearylphosphatidylcholine (PSPC), stearylpalmitoylphosphatidylcholine (SPPC), and stearylleoylphosphatidylcholine (SOPC).

[00140] Representative diacylphosphatidylethanolamines (i.e., 1,2-diacyl-sn-glycero-3-phosphoethanolamines) include dioleoylphosphatidylethanolamine (DOPE), dipalmitoylphosphatidylethanolamine (DPPE), distearoylphosphatidylethanolamine (DSPE), dilauroylphosphatidylethanolamine (DLPE), dimyristoylphosphatidylethanolamine (DMPE), dierucoylphosphatidylethanolamine (DEPE), and palmitoylloeylphosphatidylethanolamine (POPE).

[00141] Representative diacylphosphatidylglycerols (i.e., 1,2-diacyl-sn-glycero-3-phosphoglycerols) include dioleoylphosphatidylglycerol (DOPG), dipalmitoylphosphatidylglycerol (DPPG), dierucoylphosphatidylglycerol (DEPG), dilauroylphosphatidylglycerol (DLPG), dimyristoylphosphatidylglycerol (DMPG), distearoylphosphatidylglycerol (DSPG), and palmitoylloeylphosphatidylglycerol (POPG).

[00142] Representative diacylphosphatidylserines (i.e., 1,2-diacyl-sn-glycero-3-phosphoserines) include dilauroylphosphatidylserine (DLPS), dioleoylphosphatidylserine (DOPS), dipalmitoylphosphatidylserine (DPPS), and distearoylphosphatidylserine (DSPS).

[00143] Representative diacylphosphatidic acids (i.e., 1,2-diacyl-sn-glycero-3-phosphates) include dierucoylphosphatidic acid (DEPA), dilauroylphosphatidic acid (DLPA), dimyristoylphosphatidic acid (DMPA), dioleoylphosphatidic acid (DOPA), dipalmitoylphosphatidic acid (DPPA), and distearoylphosphatidic acid (DSPA).

[00144] Representative phospholipids include phosphosphingolipids such as ceramide phosphoryllipid, ceramide phosphorylcholine, and ceramide phosphorylethanolamine.

[00145] In some embodiments of any of the aspects, the nanodics of the invention can comprise one type of phospholipid. In some embodiments of any of the aspects, the nanodics of the invention can comprise two or more different phospholipids. In some embodiments of any of the aspects, the nanodics of the invention can comprise two different phospholipids. In some embodiments of any of the aspects, the nanodics of the invention can comprise three different phospholipids. In some embodiments of any of the aspects, the nanodics of the invention can comprise four different phospholipids.

[00146] In some embodiments of any of the aspects, the phospholipid can be a native lipid extract. In some embodiments of any of the aspects, the phospholipid can be a headgroup modified lipid, e.g., alkyl phosphates (e.g. 16:0 Monomethyl PE; 18:1 Monomethyl PE; 16:0 Dimethyl PE; 18:1 Dimethyl PE; 16:0 Phosphatidylmethanol; 18:1 Phosphatidylmethanol; 16:0 Phosphatidylethanol; 18:1 Phosphatidylethanol; 16:0-18:1 Phosphatidylethanol; 16:0 Phosphatidylpropanol; 18:1 Phosphatidylpropanol; 16:0 Phosphatidylbutanol; and 18:1 Phosphatidylbutanol); MRI imaging reagents (e.g. 14:0 PE-DTPA (Gd);

DTPA-BSA (Gd); 16:0 PE-DTPA (Gd); 18:0 PE-DTPA (Gd); bis(14:0 PE)-DTPA (Gd); bis(16:0 PE)-DTPA (Gd); and bis(18:0 PE)-DTPA (Gd)); chelators (e.g., 14:0 PE-DTPA (Gd); DTPA-BSA (Gd); 16:0 PE-DTPA (Gd); 18:0 PE-DTPA (Gd); bis(14:0 PE)-DTPA (Gd); bis(16:0 PE)-DTPA (Gd); and bis(18:0 PE)-DTPA (Gd)); glycosylated lipids (e.g., 18:1 Lactosyl PE); pH sensitive lipids (e.g., N-palmitoyl homocysteine; 18:1 DGS; 16:0 DGS; and DOBAQ); antigenic lipids (e.g., DNP PE and DNP Cap PE); adhesive lipids (e.g., 16:0 DG Galloyl); and functionalized lipids (e.g., 16:0 PA-PEG3-mannose; 16:0 Caproylamine PE; 18:1 Caproylamine PE; 16:0 MPB PE; 18:1 MPB PE; 16:0 Ptd Ethylene Glycol; 18:1 Ptd Ethylene Glycol; 16:0 Folate Cap PE; 16:0 Cyanur PE; 16:0 Biotinyl PE; 18:1 Biotinyl PE; 16:0 Biotinyl Cap PE; 16:0 Cyanur Cap PE; 18:1 Biotinyl Cap PE; 16:0 Dodecanoyl PE; 18:1 Dodecanoyl PE; 16:0 Glutaryl PE; 18:1 Glutaryl PE; 16:0 Succinyl PE; 18:1 Succinyl PE; 16:0 PDP PE; 18:1 PDP PE; 16:0 Ptd Thioethanol; 18:1 Dodecanylamine PE; 16:0 Dodecanylamine PE; 16:0 PE MCC; 18:1 PE MCC; 16:0 hexynoyl PE; and 16:0 azidocaproyl PE).

[00147] In some embodiments of any of the aspects described herein, the phospholipids can further comprise a molecule of interest, e.g., a membrane protein, a receptor, a transmembrane protein or channel, hydrophobic small molecules, hydrophobic drugs, RNA, peptides, and the like. In some embodiments of any of the aspects described herein, the covalently circularized nanodiscs described herein can be or be utilized as a drug delivery vehicle.

[00148] Provided herein are methods of making the covalently circularized nanodiscs described herein. In one aspect of any of the embodiments, described herein is a method comprising: a) circularizing the loopable membrane scaffold protein described herein; b) contacting the looped protein produced in step (a) with solubilized phospholipids.

[00149] In one aspect of any of the embodiments, described herein is a method comprising: a) contacting the loopable membrane scaffold protein described herein with a sortase enzyme; b) contacting the looped protein produced in step (a) with solubilized phospholipids. A number of sortase enzymes are known in the art, e.g., ones that recognize the various exemplary circularization domain sequences provided herein. By way of non-limiting example, the sortase can be *Staphylococcus aureus* wild type sortase A (Srt A); evolved sortase; eSrtA; eSrtA(2A-9); eSrtA(4S-9); or *Streptococcus pyogenes* sortase A. Wild type SrtA and eSrtA recognize LPXTG (SEQ ID NO: 51) while eSrtA(2A-9) recognizes LAXTG (SEQ ID NO: 75), LSETG (SEQ ID NO: 76) and accepts glycine nucleophiles and eSrt A(4S-9) recognizes LPXSG (SEQ ID NO: 77), LPXCG (SEQ ID NO: 78), LPXAG (SEQ ID NO: 79) and accepts glycine nucleophiles. Evolved sortases are described further in, e.g., Dorr et al. PNAS 2014 111(37):13343-8 and Chen et al. PNAS 2011 108(28):11399-404; each of which is incorporated by reference herein in its entirety. *Streptococcus* sortase A recognizes the LPXTA motif (SEQ ID NO: 56)

and accepts alanine nucleophiles (see, e.g., P. Race, *J. Biol. Chem.* **284**: 6924–6933; which is incorporated by reference herein in its entirety).

[00150] In some embodiments of any of the aspects, the contacting step can be stopped by quenching the reaction with a sortase inhibitor. Sortase inhibition is known in the art and can include the use of AAEK2 or a high concentration of EDTA. As used herein, a high concentration of EDTA can be from about 100 μ M to about 400 mM, e.g., 100 μ M to 400 mM, about 1mM to about 300 mM, 1mM to 300 mM, about 1mM to about 200 mM, or 1mM to 200 mM.

[00151] In one aspect of any of the embodiments, described herein is a method comprising a) contacting the loopable membrane scaffold protein comprising intein circularization domains with: i) a solution having a pH low enough to induce cleavage of the dipeptide sequence of Asn-Cys; ii) and a thiol reagent to induce cleavage at the Cys residue; resulting in circularization of the protein; and b) contacting the looped protein produced in step (a) with solubilized phospholipids. In some embodiments of any of the aspects, the thiol reagent is DTT.

[00152] In some embodiments of any of the aspects, the loopable membrane scaffold protein comprises a chitin binding domain flanking each circularization domain and is first purified by chitin binding prior to circularization (e.g., the loopable membrane scaffold protein is expressed from a pTWIN vector in *E. coli* and purified by binding to chitin beads).

[00153] In one aspect of any of the embodiments, described herein is a method comprising a) contacting a loopable membrane scaffold protein comprising circularization domains comprising cysteine residues with a thiol reagent or thiol donor, resulting in circularization of the protein; and b) contacting the looped protein produced in step (a) with solubilized phospholipids. In some embodiments of any of the aspects, the thiol can be DTT, sodium 2-mercaptoethane sulfonate (MESNa) or guanidinium hydrochloride. Circularization via two cysteine residues is described in the art, e.g., Cowper, B et. al. *ChemBioChem* 2013, 14, 809 – 812; which is incorporated by reference herein in its entirety. Briefly, the C-terminal cysteine initiates an intramolecular N $\rightarrow$ S acyl shift. This results in an S-acyl intermediate that can be intercepted by an added thiol, yielding a C-terminal thioester and liberating cysteine. The cysteine at the N terminus interacts with the C-terminal thioester and leads to spontaneous backbone circularization, through intramolecular transthioesterification. Next, an S $\rightarrow$ N acyl shift rearrangement occurs that lead to formation of a peptide bond. The final circularized product has a cysteine next to the newly formed peptide bond.

[00154] In some embodiments of any of the aspects, a reagent and/or enzyme used to catalyze or permit circularization (e.g., a sortase enzyme) can be bound to a substrate or surface.

[00155] In some embodiments of any of the aspects, a reagent and/or enzyme used to catalyze or permit circularization (e.g., a sortase enzyme) can further comprise an affinity tag to, e.g., permit binding a substrate or to assist in removal of the reagent and/or enzyme when the reaction is to be stopped. In some embodiments of any of the aspects, the affinity tag is a His tag. In some embodiments of any of the aspects, a sortase can further comprise a His tag.

[00156] In some embodiments of any of the aspects, the loopable membrane scaffold protein can further comprise an affinity tag that permits the protein to be bound to a substrate or surface. In some embodiments of any of the aspects, the loopable membrane scaffold protein can further comprise a His tag that permits the protein to be bound to a substrate or surface. In some embodiments of any of the aspects, the loopable membrane scaffold protein can further comprise a C-terminal affinity tag that permits the protein to be bound to a substrate or surface. In some embodiments of any of the aspects, the method described herein can further comprise a first step of binding a loopable membrane scaffold protein comprising an affinity tag to a substrate or surface.

[00157] In some embodiments of any of the aspects, the substrate comprises Cu^{2+} , Ni^{2+} or Co^{2+} . In some embodiments of any of the aspects, the substrate is a column, chip, or bead.

[00158] In some embodiments of any of the aspects, the loopable membrane scaffold protein is bound to a substrate, e.g., a bead, and the first contacting step is performed in the presence of phospholipids. In some embodiments of any of the aspects, the loopable membrane scaffold protein is bound to a substrate, e.g., a bead, and the first and second contacting steps are performed concurrently.

[00159] In some embodiments of any of the aspects, the loopable membrane scaffold protein is bound to a substrate, e.g., a bead, and after the first contacting step, the looped membrane scaffold protein is eluted from the substrate.

[00160] In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution prior to the first contacting step. In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a concentration of less than about 100 μM prior to the first contacting step. In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a concentration of less than 100 μM prior to the first contacting step. In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a concentration of from about 1 μM to about 1 mM prior to the first contacting step. In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a concentration of from 1 μM to 1 mM prior to the first contacting step. In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a concentration of from about 1 μM to about 100 μM prior to the first contacting step. In some embodiments of any of the aspects, the loopable membrane scaffold protein is in solution at a

concentration of from 1 μ M to 100 μ M prior to the first contacting step. In some embodiments of any of the aspects, the looped membrane scaffold protein formed after the first contacting step is passed through a column, substrate, and/or solution which can be bound by an affinity tag found on the loopable membrane scaffold protein. In some embodiments of any of the aspects, the looped membrane scaffold protein formed after the first contacting step is passed through a column, substrate, and/or solution Cu^{2+} , Co^{2+} or Ni^{2+} , which can be bound by an affinity tag (e.g. a His tag) found on the loopable membrane scaffold protein.

[00161] In some embodiments of any of the aspects, the loopable membrane protein further comprises a cap sequence N-terminal of the N-terminal circularization domain and the method comprises a first step of contacting the protein with an enzyme to cleave the cap sequence.

[00162] In some embodiments of any of the aspects, the phospholipids of the second contacting step are solubilized in detergent. Exemplary detergents can include, but are not limited to Decyl β -D-maltopyranoside, Deoxycholic acid, Digitonin, n-Dodecyl β -D-glucopyranoside, n-Dodecyl β -D-maltoside, N-Lauroylsarcosine sodium salt, Sodium cholate, Sodium deoxycholate, Undecyl β -D-maltoside, Triton X-100, CHAPS, 5-Cyclohexylpentyl β -D-maltoside, n-dodecyl phosphatidylcholine, n-octyl- β -d-glucoside, and Brij 97.

[00163] In some embodiments of any of the aspects, the second contacting step occurs at a temperature of from about 0 C to about 42 C. In some embodiments of any of the aspects, the second contacting step occurs at a temperature of from 0 C to 42 C. In some embodiments of any of the aspects, the second contacting step occurs at a temperature of from about 0 C to about 37 C. In some embodiments of any of the aspects, the second contacting step occurs at a temperature of from 0 C to 37 C.

[00164] In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:1000 to about 1:10,000. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from 1:1000 to 1:10,000. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from about 1:2000 to about 1:5000. In some embodiments of any of the aspects, the ratio of loopable membrane scaffold protein to phospholipids is from 1:2000 to 1:5000.

[00165] In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is from about 1:3500 to about 1:4500; and nanodiscs which are circular in shape formed. In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is about 1:4000; and nanodiscs which are circular in shape are formed. In some embodiments of any of the aspects, the loopable membrane

scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is from 1:3500 to 1:4500; and nanodiscs which are circular in shape are formed. In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is 1:4000; and nanodiscs which are circular in shape are formed. In some embodiments of any of the aspects, the phospholipids are POPC/POPG lipids.

[00166] In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is from about 1:2500 to about 1:3700; and nanodiscs which are polygonal in shape are formed. In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is from about 1:3100 to about 1:3400; and nanodiscs which are polygonal in shape are formed. In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is from 1:2500 to 1:3700; and nanodiscs which are polygonal in shape are formed. In some embodiments of any of the aspects, the loopable membrane scaffold protein has the sequence of SEQ ID NO: 4; the ratio of loopable membrane scaffold protein to phospholipids is from 1:3100 to 1:3400; and nanodiscs which are polygonal in shape are formed. In some embodiments of any of the aspects, the phospholipids are POPC/POPG lipids.

[00167] The nanodiscs described herein can be used, e.g., in structural and/or functional studies of membrane proteins or as vaccine platforms. It is specifically contemplated that the nanodiscs described herein can be used for solution NMR, cryo-EM studies, crystallization, and 2D lattice of, e.g., membrane proteins in or on the bilayer. Membrane proteins can be studied to examine, e.g., protein-protein interactions with other membrane proteins, or cytosolic or extracellular proteins; behavior and structure of membrane pores; virus entry and inhibition thereof; and cell-free expression of membrane proteins. The nanodiscs described herein can also permit study of, e.g., lipid rafts themselves, vesicle fusion, and Residual Dipolar Coupling (RDC) measurements (detergent free, homogenous effect, charge tunable).

[00168] In one aspect of any of the embodiments, described herein is a method of inducing an immune response in a subject, the method comprising administering to the subject a composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc. As used herein, an “immune response” refers to a response by a cell of the immune system, such as a B cell, T cell (CD4 or CD8), regulatory T cell, antigen-presenting cell, dendritic cell, monocyte, macrophage, NKT cell, NK cell, basophil, eosinophil, or neutrophil, to a stimulus (e.g., to an adjuvant). In some embodiments of the

aspects described herein, the response is specific for a particular antigen (an "antigen-specific response"), and refers to a response by a CD4 T cell, CD8 T cell, or B cell via their antigen-specific receptor. In some embodiments of the aspects described herein, an immune response is a T cell response, such as a CD4+ response or a CD8+ response. Such responses by these cells can include, for example, cytotoxicity, proliferation, cytokine or chemokine production, trafficking, or phagocytosis, and can be dependent on the nature of the immune cell undergoing the response. Stimulation of an immune response refers to an induction or increase of the immune response.

[00169] An immune response to an antigen can be the development in a subject of a humoral and/or a cell-mediated immune response to molecules present in the antigen or vaccine composition of interest. For purposes of the present invention, a "humoral immune response" is an antibody-mediated immune response and involves the induction and generation of antibodies that recognize and bind with some affinity for the antigen in the immunogenic composition of the invention, while a "cell-mediated immune response" is one mediated by T-cells and/or other white blood cells. A "cell-mediated immune response" is elicited by the presentation of antigenic epitopes in association with Class I or Class II molecules of the major histocompatibility complex (MHC), CD1 or other non-classical MHC-like molecules. This activates antigen-specific CD4+ T helper cells or CD8+ cytotoxic lymphocyte cells ("CTLs"). CTLs have specificity for peptide antigens that are presented in association with proteins encoded by classical or non-classical MHCs and expressed on the surfaces of cells. CTLs help induce and promote the intracellular destruction of intracellular microbes, or the lysis of cells infected with such microbes. Another aspect of cellular immunity involves an antigen-specific response by helper T-cells. Helper T-cells act to help stimulate the function, and focus the activity of, nonspecific effector cells against cells displaying peptide or other antigens in association with classical or non-classical MHC molecules on their surface. A "cell-mediated immune response" also refers to the production of cytokines, chemokines and other such molecules produced by activated T-cells and/or other white blood cells, including those derived from CD4+ and CD8+ T-cells. The ability of a particular antigen or composition to stimulate a cell-mediated immunological response may be determined by a number of assays, such as by lymphoproliferation (lymphocyte activation) assays, CTL cytotoxic cell assays, by assaying for T-lymphocytes specific for the antigen in a sensitized subject, or by measurement of cytokine production by T cells in response to re-stimulation with antigen. Such assays are well known in the art. See, e.g., Erickson et al. (1993) *J. Immunol.* 151:4189-4199; and Doe et al. (1994) *Eur. J. Immunol.* 24:2369-2376.

[00170] In some embodiments of any of the aspects, an immune response can comprise immunization. Immunization refers to the process of inducing a high level of antibody and/or cellular immune response in a subject, that is directed against a pathogen or antigen to which the organism has been exposed.

Immunization can provide protection against the relevant pathogen, e.g., prevent disease arising from a pathogen comprising the antigen if the subject is exposed to said pathogen after immunization.

Immunization may be life-long or require booster administrations.

[00171] In some embodiments of any of the aspects, a subject administered a composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc is further administered an adjuvant, e.g. in the same composition or in a different composition, at the same time, or at a separate time. In some embodiments of any of the aspects, the composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc is admixed with the adjuvant prior to administration.

[00172] In some embodiments of any of the aspects, the administration is a prime administration and/or a boost administration. In some embodiments of any of the aspects, the subject is further administered a second vaccine, e.g., against the same or a different pathogen(s).

[00173] The compositions and methods described herein can be administered to a subject in need of vaccination, immunization, and/or stimulation of an immune response. In some embodiments of any of the aspects, the methods described herein comprise administering an effective amount of compositions described herein, e.g. to a subject in order to stimulate an immune response or provide protection against the relevant pathogen the antigen was derived from. Providing protection against the relevant pathogen is stimulating the immune system such that later exposure to the antigen (e.g., on or in a live pathogen) triggers a more effective immune response than if the subject was naïve to the antigen. Protection can include faster clearance of the pathogen, reduced severity and/or time of symptoms, and/or lack of development of disease or symptoms. As compared with an equivalent untreated control, such reduction is by at least 5%, 10%, 20%, 40%, 50%, 60%, 80%, 90%, 95%, 99% or more as measured by any standard technique. A variety of means for administering the compositions described herein to subjects are known to those of skill in the art. Such methods can include, but are not limited to oral, parenteral, intravenous, intramuscular, subcutaneous, transdermal, airway (aerosol), pulmonary, cutaneous, injection, or topical, administration. Administration can be local or systemic. In some embodiments of any of the aspects, the administration can be intramuscular or subcutaneous.

[00174] The term "effective amount" as used herein refers to the amount of adjuvant needed to stimulate the immune system, or in combination with an antigen, to provide a protective effect against subsequent infections, and relates to a sufficient amount of pharmacological composition to provide the desired effect. The term "therapeutically effective amount" therefore refers to an amount of the adjuvant (and optionally, the antigen) that is sufficient to provide a particular immune stimulatory effect when

administered to a typical subject. An effective amount as used herein, in various contexts, would also include an amount sufficient to delay the development of a symptom of the disease, alter the course of a symptom of the disease (for example but not limited to, slowing the progression of a symptom of the disease), or prevent a symptom of the disease. Thus, it is not generally practicable to specify an exact "effective amount". However, for any given case, an appropriate "effective amount" can be determined by one of ordinary skill in the art using only routine experimentation.

[00175] Effective amounts, toxicity, and therapeutic efficacy can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dosage can vary depending upon the dosage form employed and the route of administration utilized. The dose ratio between toxic and therapeutic effects is the therapeutic index and can be expressed as the ratio LD50/ED50. Compositions and methods that exhibit large therapeutic indices are preferred. A therapeutically effective dose can be estimated initially from cell culture assays. Also, a dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of a composition which achieves a half-maximal inhibition of symptoms or induction of desired responses) as determined in cell culture, or in an appropriate animal model. Levels in plasma can be measured, for example, by high performance liquid chromatography. The effects of any particular dosage can be monitored by a suitable bioassay, *e.g.*, assay for antibody titers, among others. The dosage can be determined by a physician and adjusted, as necessary, to suit observed effects of the treatment.

[00176] In some embodiments of any of the aspects, the technology described herein relates to a pharmaceutical composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc as described herein, and optionally a pharmaceutically acceptable carrier. In some embodiments of any of the aspects, the active ingredients of the pharmaceutical composition comprises a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc. In some embodiments of any of the aspects, the active ingredients of the pharmaceutical composition consist essentially of a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc. In some embodiments of any of the aspects, the active ingredients of the pharmaceutical composition consist of a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc. Pharmaceutically acceptable carriers and diluents include saline, aqueous buffer solutions, solvents and/or

dispersion media. The use of such carriers and diluents is well known in the art. Some non-limiting examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, methylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids (23) serum component, such as serum albumin, HDL and LDL; (22) C₂-C₁₂ alcohols, such as ethanol; and (23) other non-toxic compatible substances employed in pharmaceutical formulations.

Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservative and antioxidants can also be present in the formulation. The terms such as "excipient", "carrier", "pharmaceutically acceptable carrier" or the like are used interchangeably herein.

In some embodiments of any of the aspects, the carrier inhibits the degradation of the active agent.

[00177] In some embodiments of any of the aspects, the pharmaceutical composition as described herein can be a parenteral dose form. Since administration of parenteral dosage forms typically bypasses the patient's natural defenses against contaminants, parenteral dosage forms are preferably sterile or capable of being sterilized prior to administration to a patient. Examples of parenteral dosage forms include, but are not limited to, solutions ready for injection, dry products ready to be dissolved or suspended in a pharmaceutically acceptable vehicle for injection, suspensions ready for injection, and emulsions. In addition, controlled-release parenteral dosage forms can be prepared for administration of a patient, including, but not limited to, DUROS[®]-type dosage forms and dose-dumping.

[00178] Suitable vehicles that can be used to provide parenteral dosage forms of an adjuvant as disclosed within are well known to those skilled in the art. Examples include, without limitation: sterile water; water for injection USP; saline solution; glucose solution; aqueous vehicles such as but not limited to, sodium chloride injection, Ringer's injection, dextrose Injection, dextrose and sodium chloride injection, and lactated Ringer's injection; water-miscible vehicles such as, but not limited to, ethyl alcohol, polyethylene glycol, and propylene glycol; and non-aqueous vehicles such as, but not limited to, corn oil, cottonseed oil, peanut oil, sesame oil, ethyl oleate, isopropyl myristate, and benzyl benzoate. Compounds

that alter or modify the solubility of a pharmaceutically acceptable salt of a composition as disclosed herein can also be incorporated into the parenteral dosage forms of the disclosure, including conventional and controlled-release parenteral dosage forms.

[00179] Conventional dosage forms generally provide rapid or immediate drug release from the formulation. Depending on the pharmacology and pharmacokinetics of the drug, use of conventional dosage forms can lead to wide fluctuations in the concentrations of the drug in a patient's blood and other tissues. These fluctuations can impact a number of parameters, such as dose frequency, onset of action, duration of efficacy, maintenance of therapeutic blood levels, toxicity, side effects, and the like. Advantageously, controlled-release formulations can be used to control a drug's onset of action, duration of action, plasma levels within the therapeutic window, and peak blood levels. In particular, controlled- or extended-release dosage forms or formulations can be used to ensure that the maximum effectiveness of a drug is achieved while minimizing potential adverse effects and safety concerns, which can occur both from under-dosing a drug (i.e., going below the minimum therapeutic levels) as well as exceeding the toxicity level for the drug. In some embodiments of any of the aspects, the adjuvant can be administered in a sustained release formulation.

[00180] Controlled-release pharmaceutical products have a common goal of improving drug therapy over that achieved by their non-controlled release counterparts. Ideally, the use of an optimally designed controlled-release preparation in medical treatment is characterized by a minimum of drug substance being employed to cure or control the condition in a minimum amount of time. Advantages of controlled-release formulations include: 1) extended activity of the drug; 2) reduced dosage frequency; 3) increased patient compliance; 4) usage of less total drug; 5) reduction in local or systemic side effects; 6) minimization of drug accumulation; 7) reduction in blood level fluctuations; 8) improvement in efficacy of treatment; 9) reduction of potentiation or loss of drug activity; and 10) improvement in speed of control of diseases or conditions. Kim, Cherng-ju, *Controlled Release Dosage Form Design*, 2 (Technomic Publishing, Lancaster, Pa.: 2000).

[00181] Most controlled-release formulations are designed to initially release an amount of drug (active ingredient) that promptly produces the desired therapeutic effect, and gradually and continually release other amounts of drug to maintain this level of therapeutic or prophylactic effect over an extended period of time. In order to maintain this constant level of drug in the body, the drug must be released from the dosage form at a rate that will replace the amount of drug being metabolized and excreted from the body. Controlled-release of an active ingredient can be stimulated by various conditions including, but not limited to, pH, ionic strength, osmotic pressure, temperature, enzymes, water, and other physiological conditions or compounds.

[00182] A variety of known controlled- or extended-release dosage forms, formulations, and devices can be adapted for use with the salts and compositions of the disclosure. Examples include, but are not limited to, those described in U.S. Pat. Nos.: 3,845,770; 3,916,899; 3,536,809; 3,598,123; 4,008,719; 5,674,533; 5,059,595; 5,591,767; 5,120,548; 5,073,543; 5,639,476; 5,354,556; 5,733,566; and 6,365,185 B1; each of which is incorporated herein by reference. These dosage forms can be used to provide slow or controlled-release of one or more active ingredients using, for example, hydroxypropylmethyl cellulose, other polymer matrices, gels, permeable membranes, osmotic systems (such as OROS[®] (Alza Corporation, Mountain View, Calif. USA)), or a combination thereof to provide the desired release profile in varying proportions.

[00183] In some embodiments of any of the aspects, the methods described herein can further comprise administering a second agent and/or treatment to the subject, e.g. as part of a combinatorial therapy.

[00184] In certain embodiments, an effective dose of a composition as described herein can be administered to a patient once. In certain embodiments, an effective dose of a composition can be administered to a patient repeatedly. For systemic administration, subjects can be administered a therapeutic amount of a composition as described herein, such as, e.g. 0.1 mg/kg, 0.5 mg/kg, 1.0 mg/kg, 2.0 mg/kg, 2.5 mg/kg, 5 mg/kg, 10 mg/kg, 15 mg/kg, 20 mg/kg, 25 mg/kg, 30 mg/kg, 40 mg/kg, 50 mg/kg, or more.

[00185] The dosage of a composition as described herein can be determined by a physician and adjusted, as necessary, to suit observed effects of the treatment. With respect to duration and frequency of treatment, it is typical for skilled clinicians to monitor subjects in order to determine when the treatment is providing therapeutic benefit, and to determine whether to increase or decrease dosage, increase or decrease administration frequency, discontinue treatment, resume treatment, or make other alterations to the treatment regimen. The dosing schedule can vary from once a week to daily depending on a number of clinical factors, such as the subject's sensitivity to the adjuvant and/or the antigen. The desired dose or amount of activation can be administered at one time or divided into subdoses, e.g., 2-4 subdoses and administered over a period of time, e.g., at appropriate intervals through the day or other appropriate schedule. In some embodiments of any of the aspects, administration can be chronic, e.g., one or more doses over a period of weeks or months.

[00186] The dosage ranges for the administration of a composition according to the methods described herein depend upon, for example, the form of the adjuvant, its potency, and the extent to which symptoms, markers, or indicators of a response described herein are desired to be induced, for example the percentage induction desired for an immune response. The dosage should not be so large as to cause

adverse side effects, such as inflammatory responses. Generally, the dosage will vary with the age, condition, and sex of the patient and can be determined by one of skill in the art. The dosage can also be adjusted by the individual physician in the event of any complication.

[00187] The efficacy of a composition as described herein in, e.g. inducing a response as described herein (e.g. an immune response or immunization) can be determined by the skilled clinician. However, a treatment is considered "effective treatment," as the term is used herein, if one or more of the signs or symptoms of a condition described herein are altered in a beneficial manner, other clinically accepted signs or symptoms are improved, or a desired response is induced e.g., by at least 10% following treatment according to the methods described herein. Efficacy can be assessed, for example, by measuring a marker, indicator, symptom, and/or the incidence of a condition treated according to the methods described herein or any other measurable parameter appropriate. Immune responses can be detected by a variety of methods known to those skilled in the art, including but not limited to, antibody production, cytotoxicity assay, proliferation assay and cytokine release assays. For example, samples of blood can be drawn from the immunized mammal and analyzed for the presence of antibodies against the antigen administered in the respective vaccine and the titer of these antibodies can be determined by methods known in the art.

[00188] Efficacy of an agent can be determined by assessing physical indicators of a desired response, (e.g. immune response, cytokine production, antibody titers, etc). It is well within the ability of one skilled in the art to monitor efficacy of administration and/or treatment by measuring any one of such parameters, or any combination of parameters. Efficacy can be assessed in animal models of a condition described herein, for example immunization of monkeys. When using an experimental animal model, efficacy of treatment is evidenced when a statistically significant change in a marker is observed.

[00189] *In vitro* and animal model assays are provided herein which allow the assessment of a given dose of an adjuvant and/or antigen. By way of non-limiting example, the effects of a dose of adjuvant can be assessed by measuring the antibody titers or cytokine production.

[00190] The efficacy of a given dosage combination can also be assessed in an animal model, e.g. immunization of mice as described in the Examples herein.

[00191] In one aspect of any of the embodiments, described herein is a kit comprising a composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc. A kit is any manufacture (e.g., a package or container) comprising at least one reagent, e.g., a composition as described herein, the manufacture being promoted, distributed, or sold as a unit for performing the methods described herein. The kits described herein can optionally comprise additional components useful for performing the

methods described herein. By way of example, the kit can comprise fluids and compositions (e.g., buffers, needles, syringes etc.) suitable for performing one or more of the administrations according to the methods described herein, an instructional material which describes performance of a method as described herein, and the like. Additionally, the kit may comprise an instruction leaflet.

[00192] The term "treatment" (including variations thereof, e.g., "treat" or "treated") as used herein means any one or more of the following: (i) the prevention of infection or re-infection, as in a traditional vaccine, (ii) the reduction in the severity of, or, in the elimination of symptoms, and (iii) the substantial or complete elimination of the pathogen or disorder in question. Hence, treatment may be effected prophylactically (prior to infection) or therapeutically (following infection). In the present invention, prophylactic treatment is the preferred mode. According to a particular embodiment of the present invention, compositions and methods are provided that treat, including prophylactically and/or therapeutically immunize, a host animal against a microbial infection (e.g., a bacterium or virus). The methods of the present invention are useful for conferring prophylactic and/or therapeutic immunity to a subject. The methods of the present invention can also be practiced on subjects for biomedical research applications.

[00193] In some embodiments of any of the aspects, an immunogenic amount or immunologically effective amount of a composition described herein is administered. The term an "immunogenic amount," and "immunologically effective amount," both of which are used interchangeably herein, refers to the amount of the antigen or immunogenic composition sufficient to elicit an immune response, either a cellular (T-cell) or humoral (B-cell or antibody) response, or both, as measured by standard assays known to one skilled in the art.

[00194] In one aspect of any of the embodiments, after the subject is administered the composition comprising a vaccine scaffold polypeptide, immunogenic conjugate polypeptide, and/or a multimer thereof embedded in the phospholipid bilayer of a liposome or nanodisc, the method further comprises isolating from the subject: a) an antibody that specifically binds to the antigen polypeptide; b) a nucleic acid encoding an antibody that specifically binds to the antigen polypeptide; or c) a cell comprising a nucleic acid sequence encoding an antibody that specifically binds to the antigen polypeptide. Such compositions can be utilized to provide antibodies (polyclonal or monoclonal) which can be used for research, diagnostic, and/or therapeutic purposes according to methods known in the art.

[00195] As used herein, the term "pharmaceutical composition" refers to the active agent in combination with a pharmaceutically acceptable carrier e.g. a carrier commonly used in the pharmaceutical industry. The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical

judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio. In some embodiments of any of the aspects, a pharmaceutically acceptable carrier can be a carrier other than water. In some embodiments of any of the aspects, a pharmaceutically acceptable carrier can be a cream, emulsion, gel, liposome, nanoparticle, and/or ointment. In some embodiments of any of the aspects, a pharmaceutically acceptable carrier can be an artificial or engineered carrier, e.g., a carrier that the active ingredient would not be found to occur in nature.

[00196] As used herein, the term "administering," refers to the placement of a compound as disclosed herein into a subject by a method or route which results in at least partial delivery of the agent at a desired site. Pharmaceutical compositions comprising the compounds disclosed herein can be administered by any appropriate route which results in an effective treatment in the subject.

[00197] For convenience, the meaning of some terms and phrases used in the specification, examples, and appended claims, are provided below. Unless stated otherwise, or implicit from context, the following terms and phrases include the meanings provided below. The definitions are provided to aid in describing particular embodiments, and are not intended to limit the claimed invention, because the scope of the invention is limited only by the claims. Unless otherwise defined, 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 invention belongs. If there is an apparent discrepancy between the usage of a term in the art and its definition provided herein, the definition provided within the specification shall prevail.

[00198] For convenience, certain terms employed herein, in the specification, examples and appended claims are collected here.

[00199] As used herein, "liposome" refers to a vesicular structure having lipid-containing membranes enclosing an aqueous interior. In cell biology, a vesicular structure is a hollow, lamellar, spherical structure, and provides a small and enclosed compartment, separated from the cytosol by at least one lipid bilayer. Liposomes can have one or more lipid membranes. Liposome compositions can be prepared by a variety of methods that are known in the art.

[00200] As used herein, a "subject" means a human or animal. Usually the animal is a vertebrate such as a primate, rodent, domestic animal or game animal. Primates include chimpanzees, cynomolgous monkeys, spider monkeys, and macaques, e.g., Rhesus. Rodents include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals include cows, horses, pigs, deer, bison, buffalo, feline species, e.g., domestic cat, canine species, e.g., dog, fox, wolf, avian species, e.g., chicken, emu, ostrich, and fish, e.g., trout, catfish and salmon. In some embodiments of any of the aspects, the subject is a

mammal, e.g., a primate, e.g., a human. The terms, “individual,” “patient” and “subject” are used interchangeably herein.

[00201] Preferably, the subject is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but is not limited to these examples. Mammals other than humans can be advantageously used as subjects that represent animal models of immunization and immune response. A subject can be male or female.

[00202] A subject can be one who has been previously diagnosed with or identified as suffering from or having a condition in need of treatment (e.g. susceptibility to infection) or one or more complications related to such a condition, and optionally, have already undergone treatment for the condition or the one or more complications related to the condition. Alternatively, a subject can also be one who has not been previously diagnosed as having the condition or one or more complications related to the condition. For example, a subject can be one who exhibits one or more risk factors for the condition or one or more complications related to the condition or a subject who does not exhibit risk factors.

[00203] A “subject in need” of treatment for a particular condition can be a subject having that condition, diagnosed as having that condition, or at risk of developing that condition.

[00204] As used herein, the terms “protein” and “polypeptide” are used interchangeably herein to designate a series of amino acid residues, connected to each other by peptide bonds between the alpha-amino and carboxy groups of adjacent residues. The terms “protein”, and “polypeptide” refer to a polymer of amino acids, including modified amino acids (e.g., phosphorylated, glycosylated, etc.) and amino acid analogs, regardless of its size or function. “Protein” and “polypeptide” are often used in reference to relatively large polypeptides, whereas the term “peptide” is often used in reference to small polypeptides, but usage of these terms in the art overlaps. The terms “protein” and “polypeptide” are used interchangeably herein when referring to a gene product and fragments thereof. Thus, exemplary polypeptides or proteins include gene products, naturally occurring proteins, homologs, orthologs, paralogs, fragments and other equivalents, variants, fragments, and analogs of the foregoing.

[00205] As used herein, the term “nucleic acid” or “nucleic acid sequence” refers to any molecule, preferably a polymeric molecule, incorporating units of ribonucleic acid, deoxyribonucleic acid or an analog thereof. The nucleic acid can be either single-stranded or double-stranded. A single-stranded nucleic acid can be one nucleic acid strand of a denatured double-stranded DNA. Alternatively, it can be a single-stranded nucleic acid not derived from any double-stranded DNA. In one aspect, the nucleic acid can be DNA. In another aspect, the nucleic acid can be RNA. Suitable nucleic acid molecules are DNA, including genomic DNA or cDNA. Other suitable nucleic acid molecules are RNA, including mRNA.

[00206] In some embodiments of any of the aspects, a nucleic acid encoding a polypeptide as described herein (e.g. a loopable membrane scaffold protein) is comprised by a vector. In some of the aspects described herein, a nucleic acid sequence encoding a given polypeptide as described herein, or any module thereof, is operably linked to a vector. The term "vector", as used herein, refers to a nucleic acid construct designed for delivery to a host cell or for transfer between different host cells. As used herein, a vector can be viral or non-viral. The term "vector" encompasses any genetic element that is capable of replication when associated with the proper control elements and that can transfer gene sequences to cells. A vector can include, but is not limited to, a cloning vector, an expression vector, a plasmid, phage, transposon, cosmid, chromosome, virus, virion, etc.

[00207] As used herein, the term "expression vector" refers to a vector that directs expression of an RNA or polypeptide from sequences linked to transcriptional regulatory sequences on the vector. The sequences expressed will often, but not necessarily, be heterologous to the cell. An expression vector may comprise additional elements, for example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in human cells for expression and in a prokaryotic host for cloning and amplification. The term "expression" refers to the cellular processes involved in producing RNA and proteins and as appropriate, secreting proteins, including where applicable, but not limited to, for example, transcription, transcript processing, translation and protein folding, modification and processing. "Expression products" include RNA transcribed from a gene, and polypeptides obtained by translation of mRNA transcribed from a gene. The term "gene" means the nucleic acid sequence, which is transcribed (DNA) to RNA *in vitro* or *in vivo* when operably linked to appropriate regulatory sequences. The gene may or may not include regions preceding and following the coding region, e.g. 5' untranslated (5' UTR) or "leader" sequences and 3' UTR or "trailer" sequences, as well as intervening sequences (introns) between individual coding segments (exons).

[00208] As used herein, the term "viral vector" refers to a nucleic acid vector construct that includes at least one element of viral origin and has the capacity to be packaged into a viral vector particle. The viral vector can contain the nucleic acid encoding encoding a polypeptide as described herein in place of non-essential viral genes. The vector and/or particle may be utilized for the purpose of transferring any nucleic acids into cells either *in vitro* or *in vivo*. Numerous forms of viral vectors are known in the art.

[00209] By "recombinant vector" is meant a vector that includes a heterologous nucleic acid sequence, or "transgene" that is capable of expression *in vivo*. It should be understood that the vectors described herein can, In some embodiments of any of the aspects, be combined with other suitable compositions and therapies. In some embodiments of any of the aspects, the vector is episomal. The use of a suitable episomal vector provides a means of maintaining the nucleotide of interest in the subject in

high copy number extra chromosomal DNA thereby eliminating potential effects of chromosomal integration.

[00210] As used herein, the term “pharmaceutical composition” refers to the active agent in combination with a pharmaceutically acceptable carrier e.g. a carrier commonly used in the pharmaceutical industry. The phrase “pharmaceutically acceptable” is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

[00211] The term “statistically significant” or “significantly” refers to statistical significance and generally means a two standard deviation (2SD) or greater difference.

[00212] Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction conditions used herein should be understood as modified in all instances by the term “about.” The term “about” when used in connection with percentages can mean $\pm 1\%$.

[00213] As used herein the term “comprising” or “comprises” is used in reference to compositions, methods, and respective component(s) thereof, that are essential to the method or composition, yet open to the inclusion of unspecified elements, whether essential or not.

[00214] The term “consisting of” refers to compositions, methods, and respective components thereof as described herein, which are exclusive of any element not recited in that description of the embodiment.

[00215] As used herein the term “consisting essentially of” refers to those elements required for a given embodiment. The term permits the presence of elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment.

[00216] As used herein, the term “corresponding to” refers to an amino acid or nucleotide at the enumerated position in a first polypeptide or nucleic acid, or an amino acid or nucleotide that is equivalent to an enumerated amino acid or nucleotide in a second polypeptide or nucleic acid. Equivalent enumerated amino acids or nucleotides can be determined by alignment of candidate sequences using degree of homology programs known in the art, e.g., BLAST.

[00217] The singular terms “a,” “an,” and “the” include plural referents unless context clearly indicates otherwise. Similarly, the word “or” is intended to include “and” unless the context clearly indicates otherwise. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of this disclosure, suitable methods and materials are described below. The

abbreviation, "e.g." is derived from the Latin *exempli gratia*, and is used herein to indicate a non-limiting example. Thus, the abbreviation "e.g." is synonymous with the term "for example."

[00218] Unless otherwise defined herein, scientific and technical terms used in connection with the present application shall have the meanings that are commonly understood by those of ordinary skill in the art to which this disclosure belongs. It should be understood that this invention is not limited to the particular methodology, protocols, and reagents, etc., described herein and as such can vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims. Definitions of common terms in immunology and molecular biology can be found in *The Merck Manual of Diagnosis and Therapy*, 19th Edition, published by Merck Sharp & Dohme Corp., 2011 (ISBN 978-0-911910-19-3); Robert S. Porter *et al.* (eds.), *The Encyclopedia of Molecular Cell Biology and Molecular Medicine*, published by Blackwell Science Ltd., 1999-2012 (ISBN 9783527600908); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8); *Immunology* by Werner Luttmann, published by Elsevier, 2006; *Janeway's Immunobiology*, Kenneth Murphy, Allan Mowat, Casey Weaver (eds.), Taylor & Francis Limited, 2014 (ISBN 0815345305, 9780815345305); *Lewin's Genes XI*, published by Jones & Bartlett Publishers, 2014 (ISBN-1449659055); Michael Richard Green and Joseph Sambrook, *Molecular Cloning: A Laboratory Manual*, 4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (2012) (ISBN 1936113414); Davis *et al.*, *Basic Methods in Molecular Biology*, Elsevier Science Publishing, Inc., New York, USA (2012) (ISBN 044460149X); *Laboratory Methods in Enzymology: DNA*, Jon Lorsch (ed.) Elsevier, 2013 (ISBN 0124199542); *Current Protocols in Molecular Biology (CPMB)*, Frederick M. Ausubel (ed.), John Wiley and Sons, 2014 (ISBN 047150338X, 9780471503385), *Current Protocols in Protein Science (CPPS)*, John E. Coligan (ed.), John Wiley and Sons, Inc., 2005; and *Current Protocols in Immunology (CPI)* (John E. Coligan, ADA M Kruisbeek, David H Margulies, Ethan M Shevach, Warren Strobe, (eds.) John Wiley and Sons, Inc., 2003 (ISBN 0471142735, 9780471142737), the contents of which are all incorporated by reference herein in their entireties.

[00219] Other terms are defined herein within the description of the various aspects of the invention.

[00220] All patents and other publications; including literature references, issued patents, published patent applications, and co-pending patent applications; cited throughout this application are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the technology described herein. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not

entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

[00221] The description of embodiments of the disclosure is not intended to be exhaustive or to limit the disclosure to the precise form disclosed. While specific embodiments of, and examples for, the disclosure are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the disclosure, as those skilled in the relevant art will recognize. For example, while method steps or functions are presented in a given order, alternative embodiments may perform functions in a different order, or functions may be performed substantially concurrently. The teachings of the disclosure provided herein can be applied to other procedures or methods as appropriate. The various embodiments described herein can be combined to provide further embodiments. Aspects of the disclosure can be modified, if necessary, to employ the compositions, functions and concepts of the above references and application to provide yet further embodiments of the disclosure. Moreover, due to biological functional equivalency considerations, some changes can be made in protein structure without affecting the biological or chemical action in kind or amount. These and other changes can be made to the disclosure in light of the detailed description. All such modifications are intended to be included within the scope of the appended claims.

[00222] Specific elements of any of the foregoing embodiments can be combined or substituted for elements in other embodiments. Furthermore, while advantages associated with certain embodiments of the disclosure have been described in the context of these embodiments, other embodiments may also exhibit such advantages, and not all embodiments need necessarily exhibit such advantages to fall within the scope of the disclosure.

[00223] The technology described herein is further illustrated by the following examples which in no way should be construed as being further limiting.

[00224] Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:

1. A loopable membrane scaffold protein comprising, from N-terminus to C-terminus:
 - i. an N-terminal circularization domain;
 - ii. a plurality of amphipathic alpha helix domains; and
 - iii. a C-terminal circularization domain.
2. The protein of paragraph 1, wherein one circularization domain comprises:

- a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and
 - the other circularization domain comprises:
 - a sequence LPXTG/A; wherein X represents any amino acid.
- 3. The protein of paragraph 1, wherein the N-terminal circularization domain comprises:
 - a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue; and
 - the C-terminal circularization domain comprises:
 - a sequence LPXTG/A; wherein X represents any amino acid.
- 4. The protein of any of paragraphs 1-3, wherein the non-polar amino acid sequence is from 1 to 10 amino acids in length and comprises amino acids selected from the group consisting of: glycine and alanine.
- 5. The protein of any of paragraphs 1-4, wherein the non-polar amino acid sequence consists of 1-5 glycine residues.
- 6. The protein of any of paragraphs 1-4, wherein the non-polar amino acid sequence consists of 1-5 alanine residues.
- 7. The protein of paragraph 1-6, wherein the sequence LPXTG/A is covalently bound to the at least one glycine or alanine residue of (i).
- 8. The protein of any of paragraphs 1-7, wherein the sequence LPXTG/A is LPGTG/A (SEQ ID NO: 24); LPSTG/A (SEQ ID NO: 25); or LPETG/A (SEQ ID NO: 26).
- 9. The protein of paragraph 1, wherein one circularization domain comprises an intein separated from the domain of (ii) by a Cys; and
 - the other circularization domain comprises an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein.
- 10. The protein of paragraph 1, wherein the N-terminal circularization domain comprises an intein separated from the domain of (ii) by a Cys; and
 - the C-terminal circularization domain comprises an intein separated from the domain of (ii) by a dipeptide sequence of Asn-Cys, with the Asn being closest to the intein.
- 11. The protein of any of paragraphs 9-10, wherein at least one intein is flanked by a chitin binding domain (CBD).
- 12. A protein of paragraph 1, wherein each circularization domain comprises a Cys.

13. The protein of any of paragraphs 1-12, wherein the amphipathic alpha helix domain comprises:
- (a) an ApoA-I polypeptide; or
 - (b) at least one alpha helical polypeptide with secondary amphipathicity or a repeat thereof;
 - (c) at least one amphipathic proline-rich polypeptide or a repeat thereof; or
 - (d) at least one repeat of the sequence XZXPPP; wherein X is any hydrophobic residue and Z is any hydrophilic residue.
14. The protein of paragraph 13, wherein the alpha helical polypeptide has a sequence selected from the group consisting of:
- DWFKAFYDKLAEKFKEAF (SEQ ID NO: 27);
 DFLKAFYDKVAEKFEAAPDWFKAFYDKVAEKFEAF (SEQ ID NO: 28);
 KLALRLALKAFKAALKLA (SEQ ID NO: 29);
 RLALDLALRAFKAAWKLA (SEQ ID NO: 30);
 ELAWDLAFEALDAELKLD (SEQ ID NO: 31);
 FLKLLKKFLKLFKKLLKLF (SEQ ID NO: 32);
 WLKLLKKWLKLWKKLLKL (SEQ ID NO: 33);
 LLKFLKRLLKLLKDLWKLL (SEQ ID NO: 34);
 WEAAFAEALAEAWAEHLAEALAEAVEALAA (SEQ ID NO: 35);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 36);
 RAFARALARALKALARALKALAR (SEQ ID NO: 37);
 GLFEALLELLESLWELLLEA (SEQ ID NO: 38);
 KLLKLLKLWKLLKLLK (SEQ ID NO: 39);
 DWLKAFYDKVAEKLKEAF (SEQ ID NO: 40); and
 DWFKAFYDKVAEKFEAF (SEQ ID NO: 41).
15. The protein of any of paragraphs 1-14, further comprising a cap sequence N-terminal of the N-terminal circularization domain that can be cleaved to expose the N-terminal circularization domain as the N-terminus of the protein.
16. The protein of paragraph 15, wherein the cap sequence comprises a sequence selected from the group consisting of:
- a. Leu-Val-Pro-Arg-Gly-Ser (SEQ ID NO: 42); and
 - b. Glu-X-Leu-Tyr- Φ -Gln- ϕ where X is any residue, Φ is any large or medium hydrophobic residue and ϕ is glycine or alanine (SEQ ID NO: 43).
17. The protein of any of paragraphs 1-16, wherein the cap sequence further comprises an affinity purification tag N-terminal of the cleavage site.
18. The protein of any of paragraphs 1-17, further comprising a His tag flanking one of the circularization domains.

19. The protein of paragraph 18, wherein the His tag is located C-terminal of the C-terminal circularization domain.
20. The protein of any of paragraphs 1-19, further comprising a linker flanking one of the circularization domains.
21. The protein of any of paragraphs 1-20, wherein the linker is located C-terminal of the C-terminal circularization domain.
22. The protein of any of paragraphs 1-21, wherein the plurality of amphipathic alpha helix domains are separated by flexible linker sequences;
23. The protein of paragraph 22, wherein the flexible linker sequence is selected from the group consisting of:
LPGTGS (SEQ ID NO: 44); LPGTSG (SEQ ID NO: 45); LPSTGS (SEQ ID NO: 46);
and LPSTSG (SEQ ID NO: 47).
24. The protein of any of paragraphs 1-23, having the sequence of any of SEQ ID NOs: 1-4 and 6-23.
25. A covalently circularized nanodisc comprising:
 - a. a phospholipid bilayer; and
 - b. a looped membrane scaffold protein forming a belt around the bilayer, wherein the looped membrane scaffold protein comprises:
 - iv. an N-terminal circularization domain covalently linked to a C-terminal circularization domain; and
 - v. a plurality of amphipathic alpha helix domains.
26. The nanodisc of paragraph 25, wherein an N-terminal circularization domain covalently linked to a C-terminal circularization domain comprises the sequence LPXT followed by a non-polar amino acid sequence comprising at least one N-terminal glycine or alanine residue;
wherein X represents any amino acid.
27. The nanodisc of any of paragraphs 25-26, wherein the non-polar amino acid sequence is from 1 to 10 amino acids in length and comprises amino acids selected from the group consisting of: glycine and alanine.
28. The nanodisc of any of paragraphs 25-27, wherein the non-polar amino acid sequence consists of 1-5 glycine residues.

29. The nanodisc of any of paragraphs 25-27, wherein the non-polar amino acid sequence consists of 1-5 alanine residues.
30. The nanodisc of any of paragraphs 25-29, wherein the sequence LPXT is LPGT (SEQ ID NO: 48); LPST (SEQ ID NO: 49); or LPET (SEQ ID NO: 50).
31. The nanodisc of any paragraph 25, wherein the N-terminal circularization domain covalently linked to a C-terminal circularization domain is a cysteine residue.
32. The nanodisc of any of paragraphs 25-31, wherein amphipathic alpha helix domain comprises:
- (a) an ApoA-I polypeptide; or
 - (b) at least one alpha helical polypeptide with secondary amphipathicity or a repeat thereof;
 - (c) at least one amphipathic proline-rich polypeptide or a repeat thereof; or
 - (d) at least one repeat of the sequence XZXPPP; wherein X is any hydrophobic residue and Z is any hydrophilic residue.
33. The nanodisc of paragraph 32, wherein the alpha helical polypeptide has a sequence selected from the group consisting of:
- DWFKAFYDKLAEKFKEAF (SEQ ID NO: 27);
 DFLKAFYDKVAEKFKEAAPDWFKAFYDKVAEKFKEAF (SEQ ID NO: 28);
 KLALRLALKAFKAALKLA (SEQ ID NO: 29);
 RLALDLALRAFKA AWKLA (SEQ ID NO: 30);
 ELAWDLAFEALDAELKLD (SEQ ID NO: 31);
 FLKLLKKFLKLFKKLLKLF (SEQ ID NO: 32);
 WLKLLKKWLKLWKKLLKL (SEQ ID NO: 33);
 LLKFLKRLLKLLKDLWKLL (SEQ ID NO: 34);
 WEAFAEALAEAWAEHLAEALAEAVEALAA (SEQ ID NO: 35);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 36);
 RAFARALARALKALARALKALAR (SEQ ID NO: 37);
 GLFEALLELLESLWELLLEA (SEQ ID NO: 38);
 KLLKLLKLWKKLLKLLK (SEQ ID NO: 39);
 DWLKA FYDKVAEKLKEAF (SEQ ID NO: 40); and
 DWFKAFYDKVAEKFKEAF (SEQ ID NO: 41).
34. The nanodisc of any of paragraphs 25-33, wherein the plurality of amphipathic alpha helix domains are separated by flexible linker sequences;
35. The nanodisc of paragraph 25-34, wherein the flexible linker sequence is selected from the group consisting of:
- LPGTGS (SEQ ID NO: 44); LPGTSG (SEQ ID NO: 45); LPSTGS (SEQ ID NO: 46);
 and LPSTSG (SEQ ID NO: 47).

36. The nanodisc of any of paragraphs 25-35, wherein the protein has the sequence of any of SEQ ID NOs: 1-4 and 6-23.
37. The nanodisc of any of paragraphs 25-36, wherein the diameter is greater than about 6 nm.
38. The nanodisc of any of paragraphs 25-37, wherein the diameter is greater than about 15 nm.
39. The nanodisc of any of paragraphs 25-38, wherein the diameter is greater than about 50 nm.
40. The nanodisc of any of paragraphs 25-39, wherein the diameter is about 80 nm.
41. A method of making a nanodisc of any of paragraphs 25-30 or 32-40, the method comprising:
- contacting the loopable membrane scaffold protein of any of paragraphs 1-8 or 13-24 with a sortase enzyme;
 - contacting the looped protein produced in step (a) with solubilized phospholipids.
42. The method of paragraph 41, wherein the sortase is selected from the group consisting of: *Staphylococcus aureus* wild type sortase A (Srt A); evolved sortase; eSrtA; eSrtA(2A-9); eSrtA(4S-9); and *Streptococcus pyogenes* sortase A.
43. The method of any of paragraphs 41-42, wherein the contacting step is stopped by quenching the reaction with a sortase inhibitor.
44. The method of paragraph 43, wherein the sortase inhibitor is AA EK2 or high concentrations of EDTA.
45. The method of any of paragraphs 41-44, wherein the sortase is bound to a substrate.
46. The method of any of paragraphs 41-45, wherein the sortase further comprises a His tag.
47. A method of making a nanodisc of any of paragraphs 31-40, the method comprising:
- contacting the loopable membrane scaffold protein of any of paragraphs 9-11 or 13-24 with:
 - a solution having a pH low enough to induce cleavage of the dipeptide sequence of Asn-Cys;
 - and a thiol reagent to induce cleavage at the Cys residue;resulting in circularization of the protein; and

- b. contacting the looped protein produced in step (a) with solubilized phospholipids.
- 48. The method of paragraph 47, wherein the thiol reagent is DTT.
- 49. A method of making a nanodisc of any of paragraphs 31-40, the method comprising:
 - a. contacting the loopable membrane scaffold protein of any of paragraphs 12-24 with a thiol reagent, resulting in circularization of the protein; and
 - b. contacting the looped protein produced in step (a) with solubilized phospholipids.
- 50. The method of any of paragraphs 41-49, wherein the phospholipids are solubilized in detergent.
- 51. The method of any of paragraphs 41-50, wherein step (b) occurs at a temperature of from about 0 C to about 42 C.
- 52. The method of any of paragraphs 41-51, wherein step (b) occurs at a temperature of from about 0 C to about 37 C.
- 53. The method of any of paragraphs 41-52, wherein the ratio of loopable membrane scaffold protein to phospholipids is from about 1:2000 to about 1:5000.
- 54. The method of any of paragraphs 41-53, wherein the ratio of loopable membrane scaffold protein of any of paragraphs 1-21 and 24 to phospholipids is from about 1:3500 to about 1:4500; and circular nanodiscs are formed.
- 55. The method of any of paragraphs 41-54, wherein the ratio of loopable membrane scaffold protein of any of paragraphs 1-21 and 24 to phospholipids is from about 1:4000; and circular nanodiscs are formed.
- 56. The method of any of paragraphs 41-55, wherein the ratio of loopable membrane scaffold protein of any of paragraphs 22-23, to phospholipids is from about 1:2500 to about 1:3700; and non-circular nanodiscs are formed.
- 57. The method of any of paragraphs 41-56, wherein the ratio of loopable membrane scaffold protein of any of paragraphs 22-23 to phospholipids is from about 1:3100 to about 1:3400; and non-circular nanodiscs are formed.

58. The method of any of paragraphs 41-57, wherein the protein further comprises a cap sequence N-terminal of the N-terminal circularization domain and the method comprises a first step of contacting the protein with an enzyme to cleave the cap sequence.
59. The method of any of paragraphs 41-58, wherein the protein further comprises a C-terminal His tag and the method comprises a first step of binding the protein to a substrate.
60. The method of any of paragraphs 41-59, wherein the substrate comprises Cu²⁺, Ni²⁺ or Co²⁺.
61. The method of any of paragraphs 41-60, wherein the substrate is a chip or bead.
62. The method of any of paragraphs 41-61, wherein the substrate is a bead and step (a) is performed in the presence of phospholipids.
63. The method of any of paragraphs 41-62, wherein the substrate is a bead and the steps (a) and (b) are performed concurrently.
64. The method of any of paragraphs 41-63, further comprising, after step (a), a step of eluting the looped protein produced in step (a) from the substrate.
65. The method of any of paragraphs 41-64, wherein the loopable membrane scaffold protein is in solution at a concentration of less than 100 uM prior to the contacting step.
66. The method of paragraph 65, wherein the looped protein formed in step (a) is passed through a column, substrate, and/or solution comprising Cu²⁺, Co²⁺ or Ni²⁺.

[00225] Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:

1. A vaccine scaffold polypeptide comprising at least the transmembrane domain of a single α -helix transmembrane protein, or a variant thereof.
2. The polypeptide of paragraph 1, wherein the single α -helix transmembrane protein is selected from the group consisting of:
phospholamban; influenza M2 channel protein; stannin; Vpu; HIV Vpu; parainfluenza 5 fusion protein; parainfluenza fusion protein; E protein; SARS-CoV E; SH protein; and RSV-SH.
3. The polypeptide of any of paragraphs 1-2, wherein the transmembrane domain of phospholamban is the sequence of one of SEQ ID NOs: 95-101, 113, or 114.
4. The polypeptide of any of paragraphs 1-2, wherein the variant of the transmembrane domain of phospholamban is a sequence with at least 90% homology to one of SEQ ID NO: 95-101, 113 or 114.

5. An immunogenic conjugate polypeptide comprising:
 - a. at least one antigen polypeptide conjugated to;
 - b. a vaccine scaffold polypeptide of paragraphs 1-4.
6. The polypeptide of paragraph 5, wherein the antigen polypeptide is conjugated to the N-terminus of the vaccine scaffold polypeptide.
7. The polypeptide of paragraph 5 wherein the antigen polypeptide is conjugated to the C-terminus of the vaccine scaffold polypeptide.
8. The polypeptide of any of paragraphs 5-7, wherein the antigen polypeptide is a viral envelope protein, a cell surface protein, or portion thereof.
9. The polypeptide of any of paragraphs 5-8, wherein the antigen polypeptide is an HIV protein or portion thereof.
10. The polypeptide of any of paragraphs 5-9, wherein the antigen polypeptide is the membrane proximal external region (MPER) of Env.
11. The polypeptide of any of paragraphs 5-9, wherein the antigen polypeptide is the 10E8 epitope or the 4E10 epitope.
12. The polypeptide of any of paragraphs 5-11, comprising multiple repeats of the same antigen polypeptide.
13. The polypeptide of any of paragraphs 5-12, comprising multiple different antigen polypeptides.
14. The polypeptide of any of paragraphs 12-13, wherein the antigen polypeptides or repeats thereof are separated by a linker.
15. The polypeptide of paragraph 14, wherein the linker is hydrophobic.
16. The polypeptide of paragraph 14, wherein the linker is hydrophilic.
17. The polypeptide of any of paragraphs 14-16, wherein the linker is from 1 to 5 amino acids in length.
18. The polypeptide of any of paragraphs 5-17, wherein the immunogenic polypeptide comprises from 1 to 5 copies of the antigen polypeptide.
19. The polypeptide of any of paragraphs 5-18, wherein the antigen polypeptide and the vaccine scaffold polypeptide are joined by a linker.
20. The polypeptide of paragraph 19, wherein the linker is hydrophobic.
21. The polypeptide of paragraph 19, wherein the linker is hydrophilic.
22. The polypeptide of any of paragraphs 19-21, wherein the linker is from 1 to 5 amino acids in length.

23. A nucleic acid encoding the vaccine scaffold polypeptide of any of paragraphs 1-5 or the immunogenic conjugate polypeptide of any of paragraphs 5-22.
24. A polypeptide complex comprising a polymeric complex of the polypeptides of any of paragraphs 1-23.
25. A composition comprising a polypeptide complex of paragraph 24 and/or an immunogenic conjugate polypeptide of any of claims 5-22 embedded in the phospholipid bilayer of a liposome or nanodisc.
26. The composition of paragraph 25, wherein the phospholipid bilayer comprises a viral membrane lipid; anionic lipid; or synthetic lipid.
27. The composition of paragraph 25, wherein the phospholipid bilayer comprises cholesterol.
28. The composition of any of paragraphs 25-27, wherein each liposome or nanodisc comprises 10 or fewer copies of the antigen polypeptide.
29. The composition of any of paragraphs 25-28, wherein each liposome or nanodisc comprises 5 or fewer copies of the antigen polypeptide.
30. A pharmaceutical composition comprising the composition of any of paragraphs 25-29 and a pharmaceutically acceptable carrier.
31. The composition of paragraph 30, further comprising an adjuvant.
32. A method of inducing an immune response in a subject, the method comprising administering to the subject a composition of any of paragraphs 25-31.
33. The method of paragraph 32, wherein the immune response comprises immunization.
34. The method of paragraphs 32-33, wherein the subject is further administered an adjuvant.
35. The method of paragraph 34, wherein the compositions of any of paragraphs 24-30 is admixed with the adjuvant prior to administration.
36. The method of any of paragraphs 32-35, wherein the administration is a prime administration and/or a boost administration.
37. The method of any of paragraphs 32-36, wherein the subject is further administered a second vaccine.
38. The method of any of paragraphs 32-37, further comprising a step of isolating from the subject:
 - a. an antibody that specifically binds to the antigen polypeptide;
 - b. a nucleic acid encoding an antibody that specifically binds to the antigen polypeptide; or

- c. a cell comprising a nucleic acid sequence encoding an antibody that specifically binds to the antigen polypeptide.

EXAMPLES

EXAMPLE 1

[00226] Described herein are three different nanodiscs embodiments and/or structures: (1) covalently circularized nanodiscs (cNDs); and (2) polygonal nanodiscs). Specifically contemplated herein are the use of these nanodiscs for, e.g., facilitating solution NMR and Cryo-EM studies, permitting structural and functional studies of large membrane proteins, and the use of the nanodiscs as vaccine platforms.

[00227] Nanodiscs have been designed to conduct in vitro biophysical studies of membrane proteins. A nanodisc is an assembly of a phospholipid bilayer surrounded or “belted” by two copies of an amphipathic helical membrane scaffold protein (MSP). A schematic representation of a nanodisc is depicted in Fig. 1.

[00228] The size of the MSP (e.g., circumference) defines the size of the nanodiscs and various lipids that can be incorporated to form the discoidal bilayer construct. Nanodiscs have been shown to provide a stable and more biologically relevant, “native-like” lipid bilayer environment compared to agents such as detergents or liposomes. Numerous proteins have been assembled into nanodiscs and these constructs have proven to be useful in studying membrane protein structure and function by solution NMR and Cryo-electron microscopy (Cryo-EM).

[00229] As used herein, the term “conventional nanodiscs” refers to discoidal, nanoscale phospholipid bilayers encompassed by a membrane scaffold protein (MSP) e.g., two molecules of amphipathic alpha-helical protein wrapped around the perimeter of the disc in an anti-parallel fashion. The hydrophobic face of the scaffold protein serves to sequester the hydrocarbon tails of the phospholipids away from the solvent and limits the size of the disc. In some embodiments of any of the aspects, the membrane scaffold protein can be apolipoprotein. These conventional nanodiscs are lipid-protein nanoparticles roughly 10 nm in diameter.

[00230] Described herein are three “non-conventional” nanodiscs: covalently circularized nanodiscs (cNDs) and polygonal nanodiscs.

[00231] Covalently circularized nanodiscs (cNDs): Described herein is the covalently linkage of the N and C termini of ApoA-I protein variants and circularization of the proteins on a Cu²⁺ chip to produce uniformly sized nanodiscs. To accommodate the analysis of larger membrane proteins, it is also

demonstrated herein that high molecular weight circularized variants can be used to prepare circularized nanodiscs of up to 80 nm in diameter.

[00232] Polygonal nanodiscs: The present technology also relates to the production of polygonal nanodiscs. The inventors determined that at suboptimal lipid concentrations, the circularized high molecular weight variants could form nanodiscs having non-circular shapes and that these can increase the chance for crystallization.

[00233] Described herein are several applications for the nanodisc structures including their use in solution NMR and Cryo-EM studies, permitting structural and functional studies of larger membrane proteins, permitting the study of membrane protein interactions with other proteins, permitting the study of viral entry into cells, and their use as vaccine platforms.

EXAMPLE 2

[00234] Described herein are covalently circularized phospholipid bilayer nanodiscs (cNDs) of various sizes up to 80 nm in diameter with a very narrow size distribution. Producing nanodiscs of increased size is highly desirable for incorporating larger integral membrane proteins (IMP), molecular complexes, oligomers of IMP, and pore-forming proteins into the discoidal bilayer for detailed biophysical investigation. This alone makes the cNDs described herein of interest to chemists engaged in pharmacological, membrane/lipid, and general biophysics research. Also, the cNDs described herein can permit new opportunities in crystallization, drug and nucleic acid delivery, vaccine platform, cell free expression of larger membrane proteins (IMP) and molecular complexes.

[00235] In Addition to cNDs, described herein are polygonal nanodiscs that offer numerous exciting applications.

[00236] Using an in vitro system that best mimics the in vivo physicochemical environment surrounding an integral membrane protein (IMP) is critical for fully understanding its molecular behavior in a biologically relevant context. Although micelles and bicelles have been used extensively in the study of IMPs, the presence of detergent often causes destabilization and rapid loss of function of the incorporated protein [1]. Furthermore, detergents can hinder interactions between an IMP and its cytosolic partner(s). Until recently, liposomes were the only means to harbor membrane proteins in a detergent-free bilayer environment. However, liposomes have a tendency to aggregate and are heterogeneous in size. Moreover, liposomes are not compatible with crystallization and solution-state NMR.

[00237] Nanodiscs on the other hand provide a detergent-free lipid bilayer model that is compatible with solution NMR, are not prone to aggregation, and do not suffer from the geometric distortion often

inherent in other membrane substitutes. A conventional nanodisc is composed of a nanometer-sized phospholipid bilayer encircled by two copies of α helical, amphipathic membrane scaffold proteins (MSPs) [2, 3]. MSPs are truncated forms of apolipoprotein A1, which is the major protein component of high-density lipoprotein. There are several different variants of MSP that can be used to assemble nanodiscs of various sizes ranging from 7-17 nm average diameter [4,5]. These different variants of MSP are commercially available (cube-biotech.com) and have been used for over a decade to study a variety of IMPs.

[00238] The use of nanodiscs for multidimensional NMR experiments, as required for complete three-dimensional structural assignment, remains challenging and is still in its early stages. Furthermore, to the best of our knowledge, there has been no membrane proteins successfully crystallized in nanodiscs. A challenging hurdle is that the diameter of the nanodisc particle can vary broadly independent of the MSP chain length in a single preparation [5, 6]. This size heterogeneity can result in a variable number of membrane proteins being incorporated into each disc [6].

[00239] Covalently circularized nanodiscs (cNDs). To improve homogeneity, stability and expand the sizes of nanodiscs, sortase A [7] was used to covalently link the N and C termini of several apoA1 variants. As shown in Figure 2A, the protein constructs contain the consensus sequence recognized by sortase A, which are a LPXTG sequence (SEQ ID NO: 51) near the C terminus and a single glycine residue at the N terminus. These two sites are sufficient to ensure covalent linkage between the N and C termini of a protein [8] while still conserving the function to form nanodiscs.

[00240] Initially, circularization was performed over a Cu⁺² chip (Figure 2B). Immobilizing proteins on the Cu⁺² chip for circularization reduces the chances of intermolecular head-to-tail linkage between two different molecules and decreases reaction time. Negative stain EM was used to confirm that circularized protein retains its ability to assemble into nanodiscs and to assess the size-homogeneity of the discs. The acquired EM images revealed uniformly sized nanodiscs and approximately 90% of the nanodiscs were found to have a diameter between 11-12 nm, as opposed to 32% of nanodiscs assembled with the un-circularized construct (Fig. 1E). Using these circularization methods, several high molecular weight (MW) circularized variants were produced and used to prepare circularized nanodiscs (cND) of up to 80 nm in diameter (Fig. 3C). To scale up the circularization reaction, we a method for circularization in solution and also over nickel beads was developed, permitting production of miligram quantities of circularized proteins.

[00241] Polygonal nanodiscs. It was observed that while performing screens to optimize the lipid-to-scaffold-protein ratio, at suboptimal lipid concentrations the circularized high MW variants could also form large nanodiscs of well-defined geometric non-circular shapes (Figs. 4A-4D). The 3X-, 4X-, 5X-,

and 6X-circularized scaffold proteins spontaneously formed triangular, square, pentagonal, and hexagonal shaped nanodiscs. The flexible linkers (LPGTGS (SEQ ID NO: 44)) between each copy of scaffold protein that result from sortase ligation enable these high MW circularized species to assemble nanodiscs with these atypical but well-defined shapes. Each side of these shapes appears to be formed by one copy of scaffold protein. These polygonal nanodiscs can increase the chance for crystallization by encouraging more efficient crystal packing, reducing the degrees of freedom in the scaffold protein by covalent closure, and increasing the stability and homogeneity of the resulting nanodisc.

[00242] Applications for the nanodiscs described herein include:

- 1- facilitating solution NMR and cryo EM studies of membrane proteins in physiologically relevant conditions (more stable and more homogenous than conventional nanodiscs).
- 2- Permit structural and functional studies of large membrane proteins (eg. mammalian respiratory complex) in bilayer.
- 3- Permit studying membrane protein interactions with another membrane protein or cytosolic partner.
- 4- Permit constructions and studying many membrane pores (eg. proapoptotic proteins BAX and BAK pores, anthrax pore etc)
- 5- Permit studying virus entry to cells (Fig. 6) and screening for potential inhibitors against virus entry.
- 6- Cell-free expression for large membrane proteins and complexes.
- 7- Permit Residual Dipolar Coupling (RDC) measurements (detergent free, homogenous effect, charge tunable).
- 8- Crystallization and 2D lattice.
- 9- Vaccination (many copies of membrane proteins per disc).
- 10- Enable studying lipid rafts.
- 11- Studying the fusion of synaptic vesicle membranes with planar bilayer membranes.

[00243] Described herein is the validation of circularized nanodiscs as a novel vaccine platform:

Because cNDs particles permit the functional reconstitution of envelope/membrane proteins, the gp41 transmembrane segment plus the highly conserved membrane proximal region (MPER-TM) of HIV, which is an important target for vaccine development, can be incorporated into cNDs and their potential to elicit an immune response and confer immunity to HIV virus challenge investigated. The improved stability, homogeneity and bigger sizes of cNDs relative to conventional nanodiscs make them an ideal platform for vaccine development.

[00244] Described herein is the design of polygonal nanodiscs and testing for crystallization. To improve the yield and prevent these shapes from becoming circular, a series of constructs can be designed with variable numbers of helices (2-6 helices) between each flexible linker. 2 or 3 helices can be fused together to increase the rigidity (Figure 5). In addition, a series of constructs can be based on the amphipathic peptide 18A [9]. The assembled nanodiscs can be examined by negative stain. After scaling up the production of polygonal nanodiscs, 2D or 3D crystallization can be performed with or without incorporated membrane protein. Stable and well-characterized membrane protein can be used for crystallization trials.

[00245] References cited

- [1] Bowie JU. Stabilizing membrane proteins. *Current opinion in structural biology*. 2001;11:397-402.
- [2] Denisov IG, Grinkova YV, Lazarides AA, Sligar SG. Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J Am Chem Soc*. 2004;126:3477-87.
- [3] Bayburt TH, Grinkova YV, Sligar SG. Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett*. 2002;2:853-6.
- [4] Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, et al. Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in enzymology*. 2009;464:211-31.
- [5] Hagn F, Etzkorn M, Raschle T, Wagner G. Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *Journal of the American Chemical Society*. 2013;135:1919-25.
- [6] Raschle T, Hiller S, Yu TY, Rice AJ, Walz T, Wagner G. Structural and functional characterization of the integral membrane protein VDAC-1 in lipid bilayer nanodiscs. *Journal of the American Chemical Society*. 2009;131:17777-9.
- [7] Chen I, Dorr BM, Liu DR. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108:11399-404.
- [8] Antos JM, Popp MW, Ernst R, Chew GL, Spooner E, Ploegh HL. A straight path to circular proteins. *The Journal of biological chemistry*. 2009;284:16028-36.
- [9] Anantharamaiah GM, Jones JL, Brouillette CG, Schmidt CF, Chung BH, Hughes TA, et al. Studies of synthetic peptide analogs of the amphipathic helix. Structure of complexes with dimyristoyl phosphatidylcholine. *The Journal of biological chemistry*. 1985;260:10248-55.

EXAMPLE 3

[00246] Design and applications of circularized nanodiscs for studying membrane proteins and viral entry.

[00247] Described herein is a process for producing covalently circularized nanodiscs covering a wide range of defined sizes up to 80 nm in diameter and in a variety of geometric shapes. Further described herein is the use of large nanodiscs to address the poorly understood question of how simple non-enveloped viruses translocate their genomes across membranes to initiate infection, and demonstrate the feasibility of using 50 nm nanodiscs to determine the structure of the RNA translocation pore of poliovirus by cryo-electron microscopy.

[00248] Phospholipid bilayer nanodiscs provide a detergent-free lipid bilayer model, enabling biochemical and biophysical characterization of membrane proteins in a physiologically relevant environment¹. A conventional nanodisc is composed of a nanometer-sized phospholipid bilayer patch encircled by two α -helical, amphipathic membrane scaffold proteins (MSPs)^{2,3}. MSPs are truncated forms of apolipoprotein A1 (apoA1), which is the major protein component of high-density lipoprotein. To date, however, the utility of this system for structural studies has been severely limited by the heterogeneity of size and the number of membrane proteins enclosed, and only small nanodiscs could be constructed with the currently available protein scaffolds⁴⁻⁷. To resolve these problems described herein are three different methods to covalently link the N and C termini of newly engineered variants based on apoA1, and the production of nanodiscs with a large range of discrete sizes and defined geometric shapes. The protein constructs described herein contain the consensus sequence recognized by sortase A (LPGTG (SEQ ID NO: 53)) near the C terminus and a single glycine residue at the N terminus (Fig. 2A). These two sites are sufficient to ensure covalent linkage between the N and C termini of a protein⁸ while still conserving the function to form nanodiscs.

[00249] Initially, the NW11 construct, which assembles an 11 nm nanodisc, was used to optimize the circularization over a Cu²⁺ chip (Fig. 2B). In this scheme, the Cu²⁺ is saturated with un-circularized NW11 protein prior to evolved sortase⁹ addition. Upon successful completion, the circularized NW11 (cNW11) is liberated to the solution and can be further purified via nickel affinity chromatography. Immobilizing NW11 on the Cu²⁺ chip for circularization reduces the chances for head-to-tail linkage of two neighboring NW11 molecules and also offers a quick reaction time. Reaction completion was confirmed by SDS-PAGE gel shift (Fig. 2C), and reaction fidelity was confirmed via tandem mass spectrometry (MS/MS) (Fig. 2D). Next, it was tested whether the final circularized product was still capable of assembling nanodiscs. Indeed, cNW11 assembled nanodiscs, and the acquired electron microscopy (EM) images revealed uniformly sized nanodiscs. Nearly 89% of the nanodiscs were found to have a diameter between 11-12 nm as opposed to 32% for nanodiscs assembled with the linear

counterpart (Fig. 2E). Even though circularization over a Cu^{+2} chip usually results in a very clean final product, the approach is limited to small-scale production of circularized protein.

[00250] In order to scale up the production of cNW11, the circularization reaction was performed over nickel beads (Fig. 9A). SDS-PAGE analysis showed that cNW11 was produced as a mixture with higher molecular weight species that likely arose from head-to-tail ligation of neighboring NW11 monomers followed by circularization (Fig. 9B, lane 1). Adding lipids to the already immobilized NW11 made the intramolecular circularization the dominant pathway, and higher molecular species were only observed in trace quantities (Fig. 9b, lane 2).

[00251] To increase the yield of cNW11, the circularization reaction was performed in solution (Fig. 10). Evolved sortase was added to a dilute NW11 solution ($[\text{NW11}] < 15 \mu\text{M}$) to suppress linking two or more copies of NW11. The reaction was quenched with a covalent sortase inhibitor AA EK2¹⁰, and cNW11 was purified via reverse nickel affinity chromatography. Reaction completion was readily confirmed by SDS-PAGE analysis. With this method, mg quantities of cNW11 that is >95% monomeric were produced. Also, an array of higher molecular weight circularized species were produced by adding sortase to a concentrated NW11 solution ($>100 \mu\text{M}$). Indeed larger nanodiscs up to 80 nm in diameter were produced (Fig. 11).

[00252] Circular dichroism (CD) spectroscopy was used to assess the effect of circularization on the thermal stability of lipid-free and lipid-bound cNW11. cNW11 has increased thermostability with an apparent midpoint melting temperature (T_m) of 65.4°C compared to 53.8°C for MSP1D1. Similarly, Lipid-bound cNW11 demonstrated better stability compared to lipid-bound MSP1D1 (T_m of 90°C vs 84.5°C) (Fig. 12A, 12B). The stability of Voltage-Dependent Anion Channel 1 (VDAC1) in LDAO micelles is increased from 63.8 to 73°C by insertion into MSP1D1 nanodiscs, and to 82.2°C when inserted into cNW11 nanodiscs (Fig. 12C). Insertion of VDAC1 into the smaller cNW9 nanodiscs maintains the high T_m of 82°C (Fig. 12D) and limits the number of embedded channels to a single one (Fig. 12E) avoiding the previously observed problem of undefined numbers of embedded channels⁵. Furthermore, covalent circularization enhances the proteolytic stability of nanodiscs (Figs. 13A-13D).

[00253] During efforts to optimize the lipids/ circularized scaffold proteins ratios, it was observed that at suboptimal lipid proportions the circularized high molecular weight variants could also form nanodiscs of well-defined, non-circular shapes (Fig. 4). The 3X-, 4X-, 5X-, and 6X-cNW11 spontaneously formed triangular, square, pentagonal, and hexagonal shaped nanodiscs. The flexible linkers (LPGTGS (SEQ ID NO: 44)) between each copy of NW11 that result from sortase ligation enable these high MW circularized species to assemble nanodiscs with these unusual but well-defined shapes. Each side of these

shapes appears to be formed by one copy of NW11. These polygonal nanodiscs are useful for crystallization efforts by encouraging more efficient crystal packing relative to circular nanodiscs.

[00254] Inspired by the large nanodiscs resulting from ligation of two or more NW11 molecules, DNA constructs NW30 and NW50 were designed to produce sortagable variants that assemble 30- and 50-nm nanodiscs, respectively. SDS-PAGE analysis (Figs. 3A, 3B) illustrates the purification and circularization of NW30 and NW50. Unlike cNW11, the circularized NW30 and NW50 migrate slower than the linear forms. Surprisingly, the cNW30 forms homogenous ~15 nm instead of 30 nm nanodiscs as confirmed by negative-stain (Fig 3A). It appears that the protein crosses and folds over itself to form a double belt that can support the bilayer with a single polypeptide. On the other hand, as predicted, circularized NW50 assembled homogenous 50 nm nanodiscs.

[00255] With the characterization of the circularized nanodiscs (cNDs) in place, the cNW50 nanodiscs were used as a model to study the question of how simple non-enveloped viruses transfer their genomes across membranes to initiate infection. Unlike enveloped viruses, non-enveloped viruses lack an external membrane, and the delivery of their genome into cells requires translocation across a membrane to gain access to the inside of the host cell^{11, 12}. Although there are now several model systems being used to study this process, the mechanism of genome translocation remains poorly understood¹³, and a more detailed structural analysis of the membrane-associated forms of the cell-entry intermediates is required. So far, mechanistic insights have been limited, due in part to technical difficulties involved in direct visualization of viral gene delivery and sample heterogeneity due to size heterogeneity of liposomes. The availability of large nanodiscs with defined size encouraged structural studies provides resolutions sufficient to gain insights into the mechanism of RNA translocation. As a proof of principal the cNW50 nanodiscs were used to visualize the RNA-translocation pore of poliovirus.

[00256] Poliovirus (30 nm diameter) is the prototype member of the enterovirus genus of the picornavirus family, which are positive-sense, single-stranded RNA viruses with ~7500b genomes enclosed by an icosahedral capsid, and lacking an envelope¹⁴. Viral infection is mediated by a specific receptor, CD155 (also known as the poliovirus receptor, PVR)¹⁵. Upon raising the temperature from 4°C to 37°C the receptor catalyzes a conformational rearrangement and expansion of the virus particle. The expanded virus is then endocytosed by a non-canonical, actin-independent pathway¹⁶, and the RNA is released across the endosomal membrane, leaving an intact empty particle that is then transported to the perinuclear region. A 50-nm nanodisc is sufficiently large to accommodate multiple CD155 copies and has enough surface area to act as a surrogate membrane for the RNA-translocation complex during viral uncoating (Fig. 6A). Similar to studies that used liposomes¹⁷⁻¹⁹, here 50-nm nanodiscs containing lipids derivatized with a NTA nickel-chelating head group were generated and functionalized with the His-

tagged CD155 ectodomain. The receptor-decorated nanodiscs were incubated with poliovirus for 5 minutes at 4°C. The complex was then heated to 37°C for 15 minutes to initiate receptor-mediated viral uncoating (Fig. 6B). Negative-stain EM confirmed virus binding to the CD155-decorated nanodiscs and subsequent insertion of viral components into and across the membrane (Fig. 6C). Additionally, the negative-stain EM images indicated that the virus started to form a pore in the nanodisc. To obtain a view of the molecular interactions involved in the RNA-translocation complex as well as to elucidate in more detail the formation of a pore, cryo-EM studies were conducted using an FEI Polara electron microscope. Figure 6D and 13 show the dark RNA-filled virus next to the slightly larger 50 nm nanodisc. The nanodisc is tilted in Figure 8E and tethered to the virus in Figure 6F. The formation of a putative pore inside the nanodiscs (Fig. 6G, 6H), through which the virus ejects its RNA was visualized. The images indicate that the pore is formed by several viral proteins whose number and identities are currently being determined. The availability of the circularized 50-nm nanodiscs greatly facilitates imaging as compared to using liposomes because the nanodiscs are more homogenous in size and shape, and allow for the use of a thinner ice layer. Also, RNA can be visualized more easily in the absence of large liposomal membranes. In order to reduce complexity even further, cNW30 nanodiscs (15 nm) decorated with CD155 were used. Surprisingly, the virus also tethers to this smaller nanodisc (Fig. 6I), forms a pore and ejects RNA, leaving an empty viral capsid behind (Fig. 6J). The quality of the data collected on the virus-nanodisc complexes represents a vast improvement over data previously used to obtain low-resolution structural models of the translocation complex using the receptor-decorated liposome model²⁰, and indicates that the nanodisc model permits determination of structures with greatly improved resolution and quality.

[00257] In conclusion, it is demonstrated herein that it is possible to construct covalently circularized nanodiscs with a wide range of geometric shapes and sizes. The ability to make stable cNDs at variable sizes up to 80 nm diameter provides a tool to tightly embed much larger membrane proteins or their intra- and extra-membrane complexes than previous nanodiscs systems have allowed. Moreover, the newly engineered covalently circularized nanodiscs produce nanodiscs with high homogeneity in size and shape and have significantly improved stability compared to non-circularized forms, both of which greatly facilitate their use for cryo-EM and other biophysical approaches. The utility of this model system to probe an outstanding question in the field of virology is demonstrated and the data indicate that the system is suitable for the structural and functional study of other large protein/membrane complexes.

[00258] *Expression of NW9, NW11, NW30 and NW50.* NW9, NW11, NW30 and NW50 (all in pET-28a) containing a TEV-cleavable N-terminal His6 tag (SEQ ID NO: 54) and a C-terminal sortase-cleavable His6 tag (SEQ ID NO: 54) were transformed into BL21-Gold (DE3) competent E. coli cells

(Agilent). 3L cell cultures were grown at 37°C, 200 rpm in Luria broth (LB) medium supplemented with 50 µg/ml Kanamycin, and expression was induced with 1 mM IPTG at an OD₆₀₀ of 0.8 for 3 hours at 37°C (NW9 and NW11) or 16 hours at 18°C (NW30 and NW50). Cells were harvested by centrifugation (7000xg, 15 minutes, 4°C) and cell pellets were stored at -80°C.

[00259] *Purification of NW9 and NW11.* Cell pellets expressing NW9 or NW11 were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1% Triton X-100) and lysed by sonication on ice. Lysate was centrifuged (35,000xg, 50 minutes, 4°C) and the supernatant was filtered and loaded onto Ni²⁺-NTA column. The column was washed with lysis buffer then with buffer A (50 mM Tris, pH 8.0, 500 mM NaCl). To recover additional protein from the insoluble fractions, the pellets recovered from lysate centrifugation were dissolved in denaturing buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 6M guanidine hydrochloride), centrifuged (35,000xg, 50 minutes, 4°C), and the supernatant was applied to the same Ni²⁺-NTA column containing bound protein from the soluble fraction. The column was washed with denaturing buffer, and NW9 and NW11 were refolded on-column with 10 column volumes (CV) buffer A. Resin was then washed with 10 CV of the following: buffer A + 1% Triton X-100, buffer A + 50 mM sodium cholate, buffer A, and buffer A + 20 mM imidazole. Proteins were eluted with buffer A + 500 mM imidazole, TEV (His6-tagged (SEQ ID NO: 54); produced in-house) was added to cleave the N-terminal His6 tag (SEQ ID NO: 54), and the samples were dialyzed against 50 mM Tris-HCl, pH 8.0, 20 mM NaCl, 1 mM EDTA, 2 mM DTT at 4°C for 16 hours. NW9 and NW11 (still containing a C-terminal His6 tag (SEQ ID NO: 54)) were exchanged into nanodisc-assembly buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 0.02% NaN₃) using centricon concentrators (10 kDa MW cutoff, Millipore).

[00260] *Purification of NW30 and NW50.* NW30 and NW50 were purified under denaturing conditions and refolded as follows. Pellets of cells expressing NW30 or NW50 were resuspended in denaturing lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 6M guanidine hydrochloride) and lysed by sonication on ice. Lysate was centrifuged (35000xg, 50 minutes, 4°C) and the supernatant was filtered and loaded onto Ni²⁺-NTA column. Resin was washed with 10 column volumes (CV) of denaturing lysis buffer to remove unbound proteins, and NW30 and NW50 were refolded on-column with 10 CV buffer A (50 mM Tris HCl, pH 8.0, 500 mM NaCl). Resin was washed with 10 CV of the following: buffer A + 1% Triton X-100, buffer A + 50 mM sodium cholate, buffer A, and buffer A + 20 mM imidazole. Proteins were eluted with buffer A + 500 mM imidazole, TEV was added to cleave the N-terminal His6 tag (SEQ ID NO: 54), and samples were dialyzed against 50 mM Tris, pH 8.0, 20 mM NaCl, 1 mM EDTA, 2 mM DTT at 4°C for 16 hours. NW30 and NW50 (still containing a Cterminal His6 tag (SEQ ID NO: 54)) were further purified by size exclusion chromatography (SEC; Superdex 200 16/60 [GE Healthcare] equilibrated in 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 50 mM sodium cholate, 0.5 mM

EDTA). SEC fractions containing NW30 and NW50 were further purified over Ni²⁺-NTA resin to remove truncation products (which lack a C-terminal His6 tag (SEQ ID NO: 54)). Purified proteins were exchanged into nanodisc assembly buffer (50 mM Tris HCl, pH 8.0, 500 mM NaCl, 0.02% NaN₃) using centricon concentrators (30 kDa MW cutoff, Millipore).

[00261] *MSP circularization.* A 50 mL reaction was prepared with 10 μ M NWs and 5 μ M (final concentrations) freshly made evolved sortase in 300 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 10 mM CaCl₂. The reaction was incubated at 4°C for 16 hours or at 25°C for 4 hours with gentle shaking on a rotating platform. A covalent sortase inhibitor AAEK2 was added to a concentration of 500 μ M, and the solution was incubated further for 30 minutes at room temperature with gentle shaking. Proteins that did not undergo circularization were removed by binding to Ni²⁺-NTA column. Circularized NWs (cNWs) were further purified by size-exclusion chromatography (Superdex 75 16/60) equilibrated in 20 mM Tris HCl, pH 7.5, 500 mM NaCl, 50 mM sodium cholate, 0.5 mM EDTA.

[00262] *Reconstitution of cNW11, cNW30 and cNW50 nanodiscs.* cNWs:lipid ratios of 1:60, 1:75, 1:1000 and 1:4000 were used to assemble cNW9, cNW11, cNW30, cNW50 nanodiscs respectively. Lipids (POPC:POPG 3:2, solubilized in sodium cholate) and cNWs were incubated on ice for 1 hour. After incubation, sodium cholate was removed by the addition of Bio-beads SM-2 (Bio-Rad) and incubation on ice for 1 hour followed by overnight incubation at 4°C. The nanodisc preparations were filtered through 0.45- μ m nitrocellulose-filter tubes to remove the Bio-beads. The nanodisc preparations were further purified by size-exclusion chromatography while monitoring the absorbance at 280 nm on a Superdex 200 10 $\times$ 300 column (for cNW9 and cNW11 nanodiscs) or Superose 6 10/300 column (for cNW30 and cNW50 nanodiscs) equilibrated in 20 mM Tris HCl, pH 7.5, 100 mM NaCl, 0.5 mM EDTA. Fractions corresponding to the size of each nanodisc were collected and concentrated. The purity of nanodisc preparations was checked using SDS-PAGE.

[00263] *In vitro reconstitution of VDAC-1 into POPC/POPG nanodiscs.* In order to assemble VDAC-1 into POPC/POPG nanodiscs, 15 μ M of VDAC-1 (solubilized in LDAO), 150 μ M of cNW11 and 8.5 mM lipids (POPC:POPG 3:2, solubilized in sodium cholate) were incubated over ice for 1 hour. After incubation, LDAO and sodium cholate were removed by the addition of Bio-beads SM-2™ (Bio-Rad) and incubation on ice for 1 hour followed by overnight incubation at 4°C. The disc preparation was filtered through 0.45- μ m nitrocellulosefilter tubes to remove the Bio-beads™. To remove the VDAC-free nanodiscs, the sample was mixed with Ni²⁺-NTA resin for 1 hour at 4°C. The resin bed volume was equal to the assembly mixture. The resin was washed with buffer E (20 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 15 mM imidazole). Nanodiscs containing VDAC were eluted with buffer E containing 0.3 M imidazole.

[00264] The nanodisc preparation was further purified by size-exclusion chromatography while monitoring the absorbance at 280 nm on a Superdex 200 10×300 column (GE Healthcare) with 20 mM Tris-HCl, pH 7.5, 0.1 mM NaCl at 0.5 mL/minute. Fractions corresponding to the size of the VDAC-nanodisc complex were collected and concentrated. The purity of VDAC containing nanodiscs was checked using SDS-PAGE.

[00265] *Negative-stain electron microscopy.* Samples were prepared by conventional negative staining as described previously²¹. Briefly, 3.5 µl of nanodisc samples were adsorbed to carbon-coated copper grids and stained with 0.75% (w/v) uranyl formate. EM images were collected with a Philips CM10™ electron microscope (FEI) equipped with a tungsten filament and operated at an acceleration voltage of 100 kV. Images were recorded with a Gatan 1 K × 1 K™ CCD camera (Gatan, Inc., Pleasanton, CA, USA). Images were analyzed with the ImageJ™ software²².

[00266] *Cryo-electron microscopy.* A 3.5 µl droplet of the poliovirus-nanodisc complex was loaded onto a glow-discharged holey grid (Protochips, Morrisville, NC). The excess liquid was removed from the grid surface before it was rapidly plunged into liquid ethane. Grids were transferred to an FEI Polara™ electron microscope operating at an acceleration voltage of 300 keV. Micrographs were acquired on a K2 Summit™ camera (Gatan, Pleasanton, CA) in super-resolution mode using SerialEM23™, whereby 25 frames were collected to a total dose of 30 electrons per square angstrom. These frames were aligned and averaged using motioncorr²⁴.

[00267] Table 1 - Characterization of intact NW proteins by mass spectrometry.

	Mass, Da			
	Linear (calculated)	Linear (observed)	Circularized (Calculated)	Circularized (Observed)
NW9	21191.8	21192.8	19838.4	19838.5
NW11	23747.7	23752.4	22394.3	22398.1
NW30	65506.9	65520.6	64153.5	64162.4
NW50	107266.1	107304.7	105912.7	105940.7

[00268] References

1. Bayburt, T.H., Carlson, J.W. & Sligar, S.G. J Struct Biol 123, 37-44 (1998).
2. Bayburt, T.H., Grinkova, Y.V. & Sligar, S.G. Nano Letters 2, 853-856 (2002).
3. Ritchie, T.K. et al. Methods Enzymol 464, 211-31 (2009).
4. Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. J Am Chem Soc 135, 1919-25 (2013).

5. Raschle, T. et al. *J Am Chem Soc* 131, 17777-9 (2009).
6. Grinkova, Y.V., Denisov, I.G. & Sligar, S.G. *Protein Eng Des Sel* 23, 843-8 (2010).
7. Raschle, T., Hiller, S., Etzkorn, M. & Wagner, G. *Current opinion in structural biology* 20, 471-9 (2010).
8. Antos, J.M. et al. *J Biol Chem* 284, 16028-36 (2009).
9. Chen, I., Dorr, B.M. & Liu, D.R. *Proc Natl Acad Sci U S A* 108, 11399-404 (2011).
10. Maresso, A.W. et al. *J Biol Chem* 282, 23129-39 (2007).
11. Hogle, J.M. *Annual Review of Microbiology* 56, 677-702 (2002).
12. Tuthill, T.J., Gropelli, E., Hogle, J.M. & Rowlands, D.J. *Cell Entry by Non-Enveloped Viruses* 343, 43-89 (2010).
13. Marsh, M. & Helenius, A. *Cell* 124, 729-40 (2006).
14. Tuthill, T.J., Gropelli, E., Hogle, J.M. & Rowlands, D.J. *Curr Top Microbiol Immunol* 343, 43-89 (2010).
15. Mendelsohn, C.L., Wimmer, E. & Racaniello, V.R. *Cell* 56, 855-65 (1989).
16. Brandenburg, B. et al. *PLoS Biol* 5, e183 (2007).
17. Tuthill, T.J., Bubeck, D., Rowlands, D.J. & Hogle, J.M. *J Virol* 80, 172-80 (2006).
18. Bubeck, D., Filman, D.J. & Hogle, J.M. *Nat Struct Mol Biol* 12, 615-8 (2005).
19. Strauss, M., Levy, H.C., Bostina, M., Filman, D.J. & Hogle, J.M. *J Virol* 87, 3903-14 (2013).
20. Strauss, M. et al. *J Virol* 89, 4143-57 (2015).
21. Ohi, M., Li, Y., Cheng, Y. & Walz, T. *Biol Proced Online* 6, 23-34 (2004).
22. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. *Nat Methods* 9, 671-5 (2012).
23. Mastronarde, D.N. *J Struct Biol* 152, 36-51 (2005).
24. Li, X., Grigorieff, N. & Cheng, Y. *J Struct Biol* 172, 407-12 (2010).

[00269] Protein sequences

[00270] SEQ ID No: 1 NW9

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQLPGTGA
AALEHHHHHH

[00271] SEQ ID NO: 2 NW11

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLEEMRDRARAHVDALRTH

LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00272] SEQ ID NO: 3 NW30

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
QLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYT
KKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQK
VEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTG
AAALEHHHHHH

[00273] SEQ ID NO: 4 NW50

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
QLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYT
KKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQK
VEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVT
QEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLRAELQEGA
RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYH
AKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKE
TEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEK
LSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTL
SEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

EXAMPLE 4: Covalently Circularized Nanodiscs for Structural and Functional Studies of Membrane Proteins

[00274] The physicochemical environment enveloping a membrane protein in vitro is critical to fully understand the molecular behavior in a biologically relevant context. Although micelles and bicelles have produced significant results, detergents can alter dynamic protein behavior and hinder interactions between a membrane protein and its cytosolic partner(s). Conventional nanodiscs are composed of a nanometer-sized phospholipid bilayer encircled by two α helical, amphipathic membrane scaffold proteins (MSPs). These particles provide a unique detergent free lipid bilayer model enabling biochemical and biophysical characterization of membrane proteins in a physiologically relevant medium. Recently, nanodiscs have been used to investigate the structure and function of different membrane proteins using solution NMR.

[00275] Although the MSPs used in a particular nanodisc preparation are the same length, the distribution in diameter of an assembled disc can be broad. This can result in a variable number of membrane proteins incorporating into each disc, and obtaining a sample with a mixed population of membrane proteins per nanodisc that lead to confounding results. To overcome these issues described herein are nanodiscs assembled with circularized MSPs (cMSP). Described herein are different methods to covalently link the N and C terminus of an MSP polypeptide using evolved sortase. Sortase reaction fidelity has been confirmed by mass spectrometry of the cMSP polypeptide. The homogeneity in particle diameter was a narrow distribution using negative-stain EM. Using the circularization method, nanodiscs of different sizes up to 100 nm in diameter can be constructed.

[00276] Methods

[00277] Described herein are three methods to create circularized MSP.

[00278] A- creating circularized MSP over Cu²⁺ (Fig. 8). This method results in the cleanest product. It is suitable for, e.g., small-scale production of circularized MSP.

[00279] B- Circularization of MSP over Ni²⁺ bead. His-tagged MSP is attached to Ni beads then some lipids are added to MSP and finally sortase is added. Adding lipid to MSP first can minimize the multimerization by-products and maximize the intramolecular circularization. This method is suitable for, e.g., large-scale production of cMSP.

[00280] C- Circularization of MSP in solution. His-tagged MSP is diluted in sortase buffer (MSP final conc < 15 μ M) then sortase is added. The reaction is quenched by adding the sortase covalent inhibitor AAEC2 (final conc 0.5 mM). The reaction mixture is run through Ni column and the flow through collected. This method is suitable for, e.g., large-scale production of cMSP.

[00281] Results

[00282] Figs. 2E demonstrates, via negative stain EM, that the circularized nanodiscs described herein are more homogenous than conventional nanodiscs. Figs. 2C-2E depict characterization of the

circularized nanodiscs. Large nanodisc sizes can be achieved with the methods described herein (Fig. 3A-3C).

[00283] It is specifically contemplated herein that the presently described circularized nanodiscs can be used for the following non-limiting applications:

Small circularized nanodiscs (<12 nm)

- 1- facilitating solution NMR studies of membrane proteins in physiologically relevant conditions (more stable and more homogenous than conventional nanodiscs)
- 2- Allow for detergent titrations and the transfer of NMR assignments (detergent to nanodisc) because of its high stability.

Large circularized nanodiscs (10-100 nm)

- 1- Permit structural and functional studies of large membrane proteins (eg. FepA and mammalian respiratory complex) in bilayer
- 2- Permit studying membrane protein interactions with another membrane protein or cytosolic partner
- 3- Permit constructions and studying many membrane pores (eg. proapoptotic proteins BAX and BAK pores, anthrax pore etc)
- 4- Permit studying virus entry to cells (Figs. 6A-6J) and screening for potential inhibitors against virus entry
- 5- Cell-free expression for large membrane proteins and complexes.
- 6- Permit Residual Dipolar Coupling (RDC) measurements (detergent free, homogenous effect, charge tunable)
- 7- Crystallization and 2D lattice
- 8- Limited proteolysis of membrane proteins (more stable than conventional nanodisc)
- 9- Vaccination (many copies of membrane proteins per disc)
- 10- Studying how Apolipoprotein A-1 interact with lipids since the size of the large nanodisc is suitable for cryo-EM size range.
- 11- Permit studying lipid rafts.
- 12- Studying the fusion of synaptic vesicle membranes with planar bilayer membranes.

[00284] Protein sequences

[00285] SEQ ID NO: 6 NW9 (assemble ~ 9 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKEN

GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQLPGTGAA
AALEHHHHHH

[00286] SEQ ID NO: 7 NW10 (assemble ~10 nm nanodisc)

MGSSHHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGRLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRRLAARLEALKENG GARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFK
VSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00287] SEQ ID NO: 8 NW18 (assemble ~ 18 nm nanodisc)

MGSSHHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGRLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGRLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRRLAARLEALKENG GARLAEYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00288] SEQ ID NO: 9 NW20 (assemble ~ 20 nm nanodisc)

MGSSHHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGRLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRRLAARLEALKENG GARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
RLAARLEALKENG GARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEY
TKKLNTQLPGTGAAALEHHHHHH

[00289] SEQ ID NO: 10 NW27 (assemble ~ 27 nm nanodisc)

MGSSHHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGRLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGRLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRRLAARLEALKENG GARLAEYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQLPGTG
AAALEHHHHHH

[00290] SEQ ID NO: 11 NW30 (assemble ~ 30 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
RLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKV
EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTG
AAALEHHHHHHH

[00291] SEQ ID NO: 12 NW36 (assemble ~ 36 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G LLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHHHHH

[00292] SEQ ID NO: 13 NW40 (assemble ~ 40 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
RLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKV
EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVT
QEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGA

RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYH
AKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHH
HH

[00293] SEQ ID NO: 14 NW45 (assemble ~ 45 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTGA
AAALEHHHHHH

[00294] SEQ ID NO: 15 NW50 (assemble ~ 50 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
QLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQK
VEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVT
QEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLRAELQEGA
RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYH
AKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKE
TEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEK

LSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTL
SEKAKPALEDLRQG L L P V L E S F K V S F L S A L E E Y T K K L N T Q L P G T G A A A L E H H H H H H

[00295] SEQ ID NO: 16 NW54 (assemble ~ 54 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T Q
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T Q E F W D N L E K E T E G L R Q E M S K D L E E V K A K V Q P Y
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T Q
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T Q E F W D N L E K E T E G L R Q E M S K D L E E V K A K V Q P Y
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T Q
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G L L P V L E S F K V S F L S A L E E Y T K K L N T Q L P G T G A A A L E H H H H H H

[00296] SEQ ID NO: 17 NW60 (assemble ~ 60 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG L L P V L E S F K
V S F L S A L E E Y T K K L N T Q G T P V T Q E F W D N L E K E T E G L R Q E M S K D L E E V K A K V Q P Y L D D F Q K K W Q
EEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
QLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG L L P V L E S F K V S F L S A L E E Y T
K K L N T Q G T P V T Q E F W D N L E K E T E G L R Q E M S K D L E E V K A K V Q P Y L D D F Q K K W Q E E M E L Y R Q K V
EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
N G G A R L A E Y H A K A T E H L S T L S E K A K P A L E D L R Q G L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T
Q E F W D N L E K E T E G L R Q E M S K D L E E V K A K V Q P Y L D D F Q K K W Q E E M E L Y R Q K V E P L R A E L Q E G A
R Q K L H E L Q E K L S P L G E E M R D R A R A H V D A L R T H L A P Y S D E L Q R Q L A A R L E A L K E N G G A R L A E Y H
A K A T E H L S T L S E K A K P A L E D L R Q G L L P V L E S F K V S F L S A L E E Y T K K L N T Q G T P V T Q E F W D N L E K E
T E G L R Q E M S K D L E E V K A K V Q P Y L D D F Q K K W Q E E M E L Y R Q K V E P L R A E L Q E G A R Q K L H E L Q E K

LSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTL
SEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMS
KDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMR
DRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALE
DLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00297] SEQ ID NO: 18 NW63 (assemble ~ 63 nm)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
G LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTG
AAALEHHHHHH

[00298] SEQ ID NO: 19 NW70 (assemble ~ 70 nm)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMEL YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
QLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL YRQKV

EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVT
QEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGA
RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGGARLAEYH
AKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKE
TEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEK
LSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGGARLAEYHAKATEHLSTL
SEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLQEMS
KDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMR
DRARAHVDALRTHLAPYSDELQRQLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALE
DLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLQEMSKDLEEVKAK
VQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDA
LRTHLAPYSDELQRQLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVL
ESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00299] SEQ ID NO: 20 NW72 (assemble ~ 72 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ
EFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKEN
GGARLAEYHAKATEHLSTLSEKAKPALEDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQ

EFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLGEEMRDRAR
AHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQ
GLLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALEHHHHHH

[00300] SEQ ID NO: 21 NW80 (assemble ~ 80 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQ
EEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQR
RLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVE
EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKE
NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVT
QEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGA
RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYH
AKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKE
TEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEK
LSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTL
SEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMS
KDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMR
DRARAHVDALRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALE
DLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK
VQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDA
LRTHLAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPV
LESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQK
KWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDE
LRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALE
EYTKKLNTQLPGTGAAALEHHHHHH

[00301] SEQ ID NO: 22 NW90 (assemble ~90 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPY
LDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
LAPYSDELQRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQ

EEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRRQ
 RLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKV
 EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRRQLAARLEALKE
 NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVT
 QEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGA
 RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRRQLAARLEALKENG GARLA EYH
 AKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKE
 TEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEK
 LSPLGEEMRDRARAHVDALRTHLAPYSDELRRQLAARLEALKENG GARLA EYHAKATEHLSTL
 SEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMS
 KDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMR
 DRARAHVDALRTHLAPYSDELRRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALE
 DLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAK
 VQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDA
 LRTHLAPYSDELRRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVL
 ESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQK
 KWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDE
 LRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALE
 EYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYR
 QKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRRQLAARLE
 ALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQ
 LPGTGAAALEHHHHHH

[00302] SEQ ID NO: 23 NW100 (assemble ~100 nm nanodisc)

MGSSHHHHHHENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPY
 LDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTH
 LAPYSDELRRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFK
 VSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQ
 EEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRRQ
 RLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEG LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKV
 EPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRRQLAARLEALKE
 NGGARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVT

QEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGA
 RQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKENG GARLA EYH
 AKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKE
 TEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEK
 LSPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKENG GARLA EYHAKATEHLSTL
 SEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMS
 KDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMR
 DRARAHVDALRTHLAPYSDELQRRLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALE
 DLRQG LLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK
 VQPYLDDFQKKWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDA
 LRTHLAPYSDELQRRLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVL
 ESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQK
 KWQEEMELRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDE
 LRQLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALE
 EYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELR
 QKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLE
 ALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQ
 GTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELRQKVEPLRAE
 LQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEALKENG GAR
 LA EYHAKATEHLSTLSEKAKPALEDLRQG LLPVLESFKVSFLSALEEYTKKLNTQLPGTGAAALE
 HHHHHH

[00303] SEQ ID NO: 5

```

1 mkaavltlav lfltgsgarh fwqqdeppqs pwdrvkdlav vyvdvlkdsg
rdyvvsqfegs
61 algkqlnlkl ldnwdsvtst fsklreqlgp vtqefwdnle keteglrqem
skdleevak
121 vqpylddf qk kwqeemelyr qkveplrael qegarqklhe lqeklsplge
emrdrarahv
181 dalrthlapy sdelrqrlaa rlealkengg arlaeyhaka tehlstlsek
akpaledlrq
241 gllpvlesfk vsflsaleey tkklntq

```

EXAMPLE 5

[00304] A loopable membrane scaffold protein was designed based on SEQ ID NO: 2, with the inclusion of substitution mutations at residues 36, 110, and 175 (SEQ ID NO: 62, NW11_3C). SDS-PAGE analysis of the circularized form of this protein is depicted in Fig. 19.

EXAMPLE 6: A NEW GENERAL-PURPOSE PHOSPHOLAMBAN-BASED SCAFFOLD FOR PEPTIDE IMMUNOGEN CONSTRUCTION

[00305] The methods and compositions described herein relate to antiviral immunology. In some embodiments of any of the aspects, the invention involves the reconstruction of immunogens using the phospholamban (PLN) TM helix as a general stable, low-cost scaffold. Potential applications include subunit vaccine construction as well as the production of diagnostic antibodies.

[00306] It is estimated that 2.3 million people become newly infected annually with the human immunodeficiency virus (HIV), the virus that causes AIDS, one of the most deadly infectious diseases in the history of mankind. A major challenge of HIV-1 vaccine development has been the great diversity of viral isolates. Developing a vaccine that overcomes the diversity of HIV-1 is likely to require a strategy that directs antibody (Ab) responses toward conserved regions of the viral Envelope (Env). However, the generation of neutralizing Abs (NAbs) targeting these regions through vaccination has proven to be very difficult. One conserved region of particular interest is the membrane proximal external region (MPER) of Env located within the gp41 ectodomain. MPER Plays an important role in membrane fusion and contains epitopes of three well-characterized human broadly neutralizing monoclonal Abs; 2F5, 4E10, and 10E8.

[00307] In order to direct the immune response to this region, the phospholamban (PLN, Figs. 20A-20B) transmembrane helix was used as a scaffold for attaching MPER, displaying up to 5 copies per complex on the membrane surface (Figs. 21A-21B). The phospholamban (PLN) TM helix provides a stable pentameric TM scaffold that coincidentally share the first three amino acid residues at the beginning of the PLN-TM region “LFI” with the natural TM sequence of HXB2-gp41 immediately following the MPER segment. Preliminary negative-stain EM images show that pre-formed 10E8 binds to MPER-PLN chimera incorporated in nanodisc (Figs. 22A-22B). Moreover, immunogenicity data indicate that mice immunized with MPER-PLN chimera elicited antibodies specific to MPER.

[00308] Immunized mice received membrane presented MPER-phospholamban fusion protein displayed on nanodiscs (100ug of MPER-phospholamban:nanodisc per mouse). Mice were immunized three times at 3-week intervals via subcutaneous tail-base injection in conjunction with 5ug/mouse CDN adjuvant-loaded liposomes. No MPER-specific IgG antibodies are detected in any naive mice. A MPER-specific antibody response is generated to nanodisc delivered MPER-immunogen (Fig 27). When nanodiscs containing MPER-Phospholamban were administered to mice, high titer immune responses specific to both MPER and human MSP, the membrane scaffold protein that wrap around nanodisc, were generated after two immunizations (Fig. 28)

- [00309] PLN-TM offers the following advantages as a vaccine scaffold:
- a. Stable pentameric scaffold TM that share the first three amino acid residues with the natural TM sequence of HXB2-gp41 immediately following the MPER segment.
 - b. MPER-PLN chimera (~50 AA) can be synthesized by conventional peptide synthesis techniques as well as expressed in *E. coli* or other expression systems.
 - c. Permits mixing and control of how many MPER arms are present per nanodisc thus control distances between each MPER copy.
 - d. Pore formation can be clearly shown by negative-stain EM that permits verification of incorporation and whether there are aggregations or not.
 - e. The hybrid MPER-PLN can be readily incorporated into nanodisc or liposomes to permit preparation of immunogens.
 - f. Mice that were immunized with nanodiscs including PLN-TM displaying MPER elicited MPER- specific Ab responses.

[00310] SEQ ID NO: 113 Phospholamban transmembrane domain
LFINFCLILICLLLIIV

[00311] SEQ ID NO: 114 Phospholamban transmembrane domain
LFVNFCLILICLLLIIV

What is claimed herein is:

1. A vaccine scaffold polypeptide comprising at least the transmembrane domain of a single α -helix transmembrane protein, or a variant thereof.
2. The polypeptide of claim 1, wherein the single α -helix transmembrane protein is selected from the group consisting of:
phospholamban; influenza M2 channel protein; stannin; Vpu; HIV Vpu; parainfluenza 5 fusion protein; parainfluenza fusion protein; E protein; SARS-CoV E; SH protein; and RSV-SH.
3. The polypeptide of any of claims 1-2, wherein the transmembrane domain is the sequence of one of SEQ ID NOs: 95-101, 113, or 114.
4. The polypeptide of any of claims 1-2, wherein the variant of the transmembrane domain is a sequence with at least 90% homology to one of SEQ ID NO: 95-101, 113 or 114.
5. An immunogenic conjugate polypeptide comprising:
 - a. at least one antigen polypeptide conjugated to;
 - b. a vaccine scaffold polypeptide of claims 1-4.
6. The polypeptide of claim 5, wherein the antigen polypeptide is conjugated to the N-terminus of the vaccine scaffold polypeptide.
7. The polypeptide of claim 5 wherein the antigen polypeptide is conjugated to the C-terminus of the vaccine scaffold polypeptide.
8. The polypeptide of any of claims 5-7, wherein the antigen polypeptide is a viral envelope protein, a cell surface protein, or portion thereof.
9. The polypeptide of any of claims 5-8, wherein the antigen polypeptide is an HIV protein or portion thereof.
10. The polypeptide of any of claims 5-9, wherein the antigen polypeptide is the membrane proximal external region (MPER) of Env.
11. The polypeptide of any of claims 5-9, wherein the antigen polypeptide is the 10E8 epitope or the 4E10 epitope.
12. The polypeptide of any of claims 5-11, comprising multiple repeats of the same antigen polypeptide.
13. The polypeptide of any of claims 5-12, comprising multiple different antigen polypeptides.

14. The polypeptide of any of claims 12-13, wherein the antigen polypeptides or repeats thereof are separated by a linker.
15. The polypeptide of claim 14, wherein the linker is hydrophobic.
16. The polypeptide of claim 14, wherein the linker is hydrophilic.
17. The polypeptide of any of claims 14-16, wherein the linker is from 1 to 5 amino acids in length.
18. The polypeptide of any of claims 5-17, wherein the immunogenic polypeptide comprises from 1 to 5 copies of the antigen polypeptide.
19. The polypeptide of any of claims 5-18, wherein the antigen polypeptide and the vaccine scaffold polypeptide are joined by a linker.
20. The polypeptide of claim 19, wherein the linker is hydrophobic.
21. The polypeptide of claim 19, wherein the linker is hydrophilic.
22. The polypeptide of any of claims 19-21, wherein the linker is from 1 to 5 amino acids in length.
23. A nucleic acid encoding the vaccine scaffold polypeptide of any of claims 1-5 or the immunogenic conjugate polypeptide of any of claims 5-22.
24. A polypeptide complex comprising a polymeric complex of the polypeptides of any of claims 1-23.
25. A composition comprising a polypeptide complex of claim 24 and/or the immunogenic conjugate polypeptide of any of claims 5-22 embedded in the phospholipid bilayer of a liposome or nanodisc.
26. The composition of claim 25, wherein the phospholipid bilayer comprises a viral membrane lipid; anionic lipid; or synthetic lipid.
27. The composition of claim 25, wherein the phospholipid bilayer comprises cholesterol.
28. The composition of any of claims 25-27, wherein each liposome or nanodisc comprises 10 or fewer copies of the antigen polypeptide.
29. The composition of any of claims 25-28, wherein each liposome or nanodisc comprises 5 or fewer copies of the antigen polypeptide.
30. A pharmaceutical composition comprising the composition of any of claims 25-29 and a pharmaceutically acceptable carrier.
31. The composition of claim 30, further comprising an adjuvant.

32. A method of inducing an immune response in a subject, the method comprising administering to the subject a composition of any of claims 25-31.
33. The method of claim 32, wherein the immune response comprises immunization.
34. The method of claims 32-33, wherein the subject is further administered an adjuvant.
35. The method of claim 34, wherein the compositions of any of claims 24-30 is admixed with the adjuvant prior to administration.
36. The method of any of claims 32-35, wherein the administration is a prime administration and/or a boost administration.
37. The method of any of claims 32-36, wherein the subject is further administered a second vaccine.
38. The method of any of claims 32-37, further comprising a step of isolating from the subject:
 - a. an antibody that specifically binds to the antigen polypeptide;
 - b. a nucleic acid encoding an antibody that specifically binds to the antigen polypeptide; or
 - c. a cell comprising a nucleic acid sequence encoding an antibody that specifically binds to the antigen polypeptide.

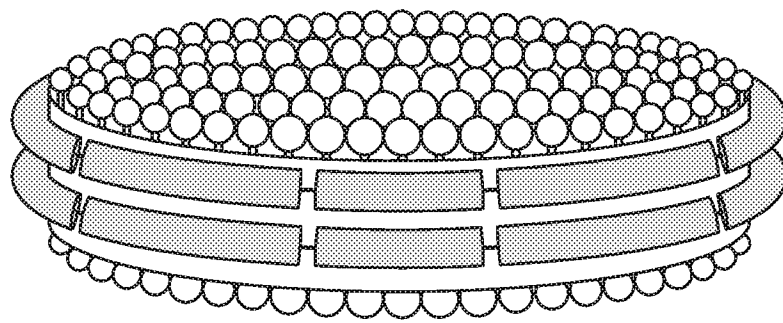


FIG. 1

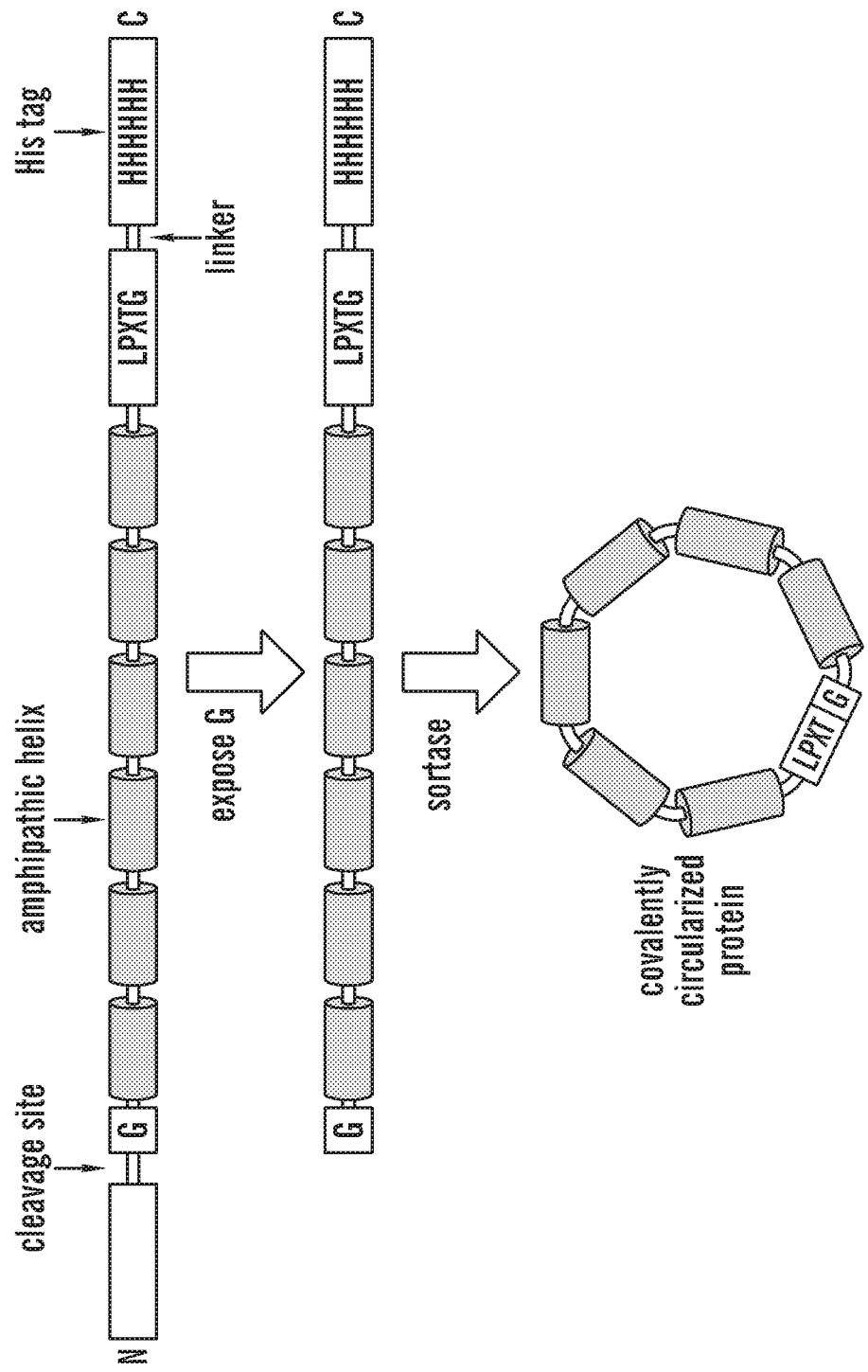


FIG. 2A

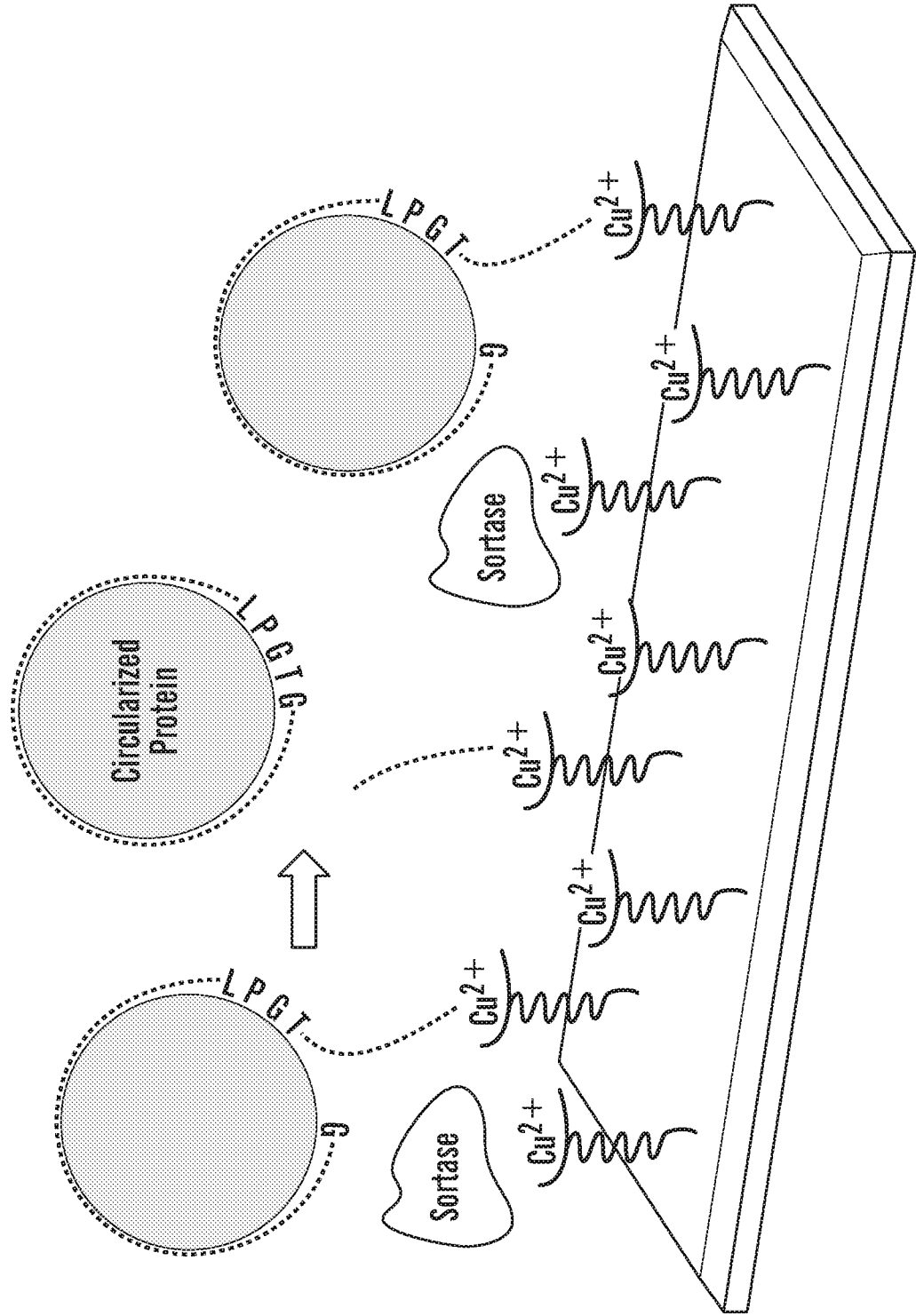


FIG. 2B

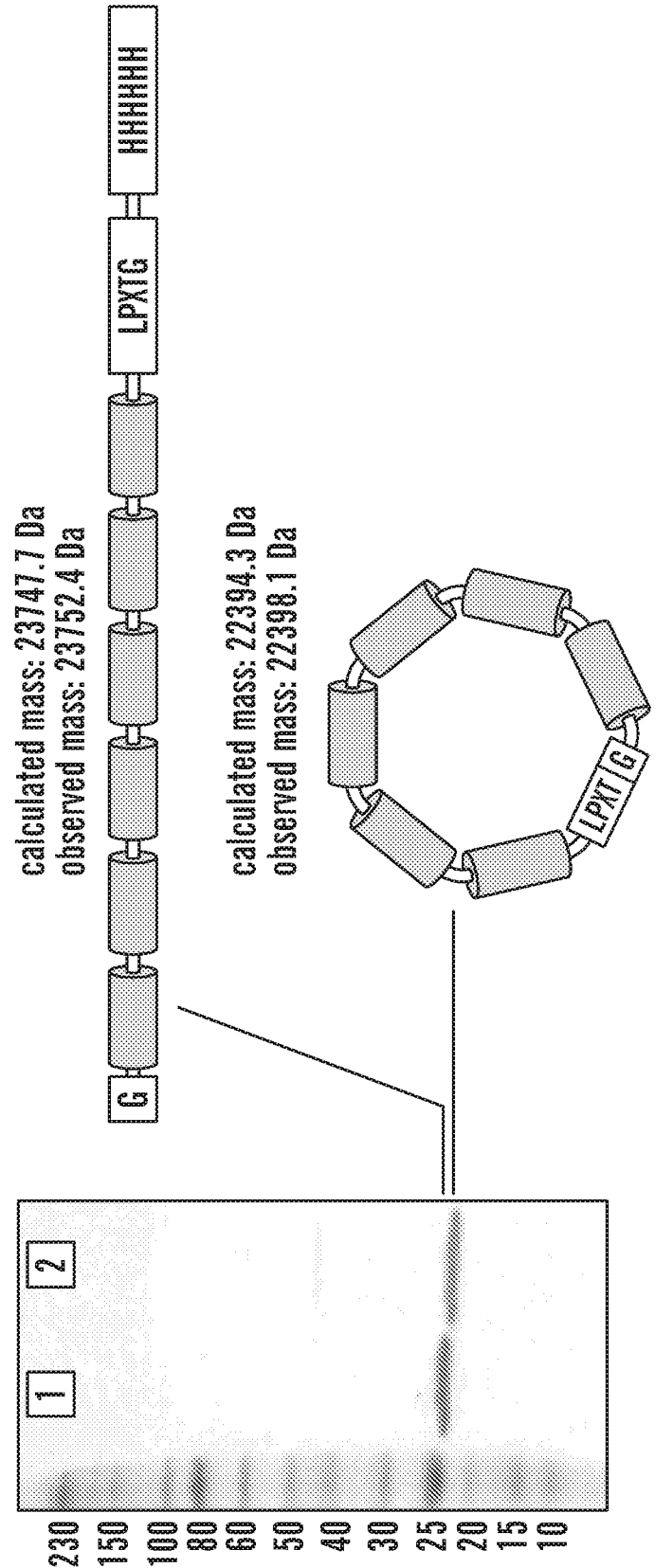


FIG. 2C

seq	#	b	y	#
K	1	120.102	...	15
L	2	242.186	1450.754	14
N	3	356.229	1337.670	13
T	4	457.277	1223.627	12
Q	5	585.335	1122.579	11
L	6	698.420	994.520	10
P	7	795.472	881.436	9
G	8	852.494	784.384	8
T	9	953.541	727.362	7
G	10	1010.563	626.314	6
S	11	1097.595	569.293	5
T	12	1198.643	482.261	4
F	13	1345.711	381.213	3
S	14	1432.743	234.145	2
K	15	...	147.113	1

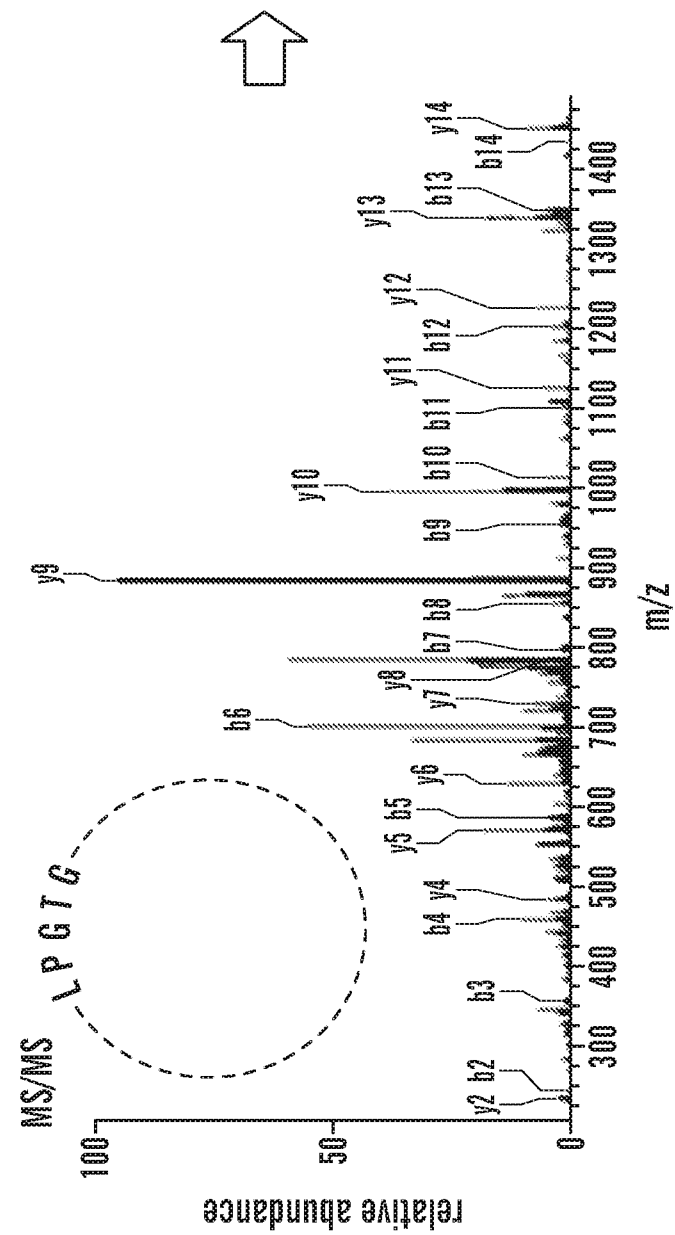
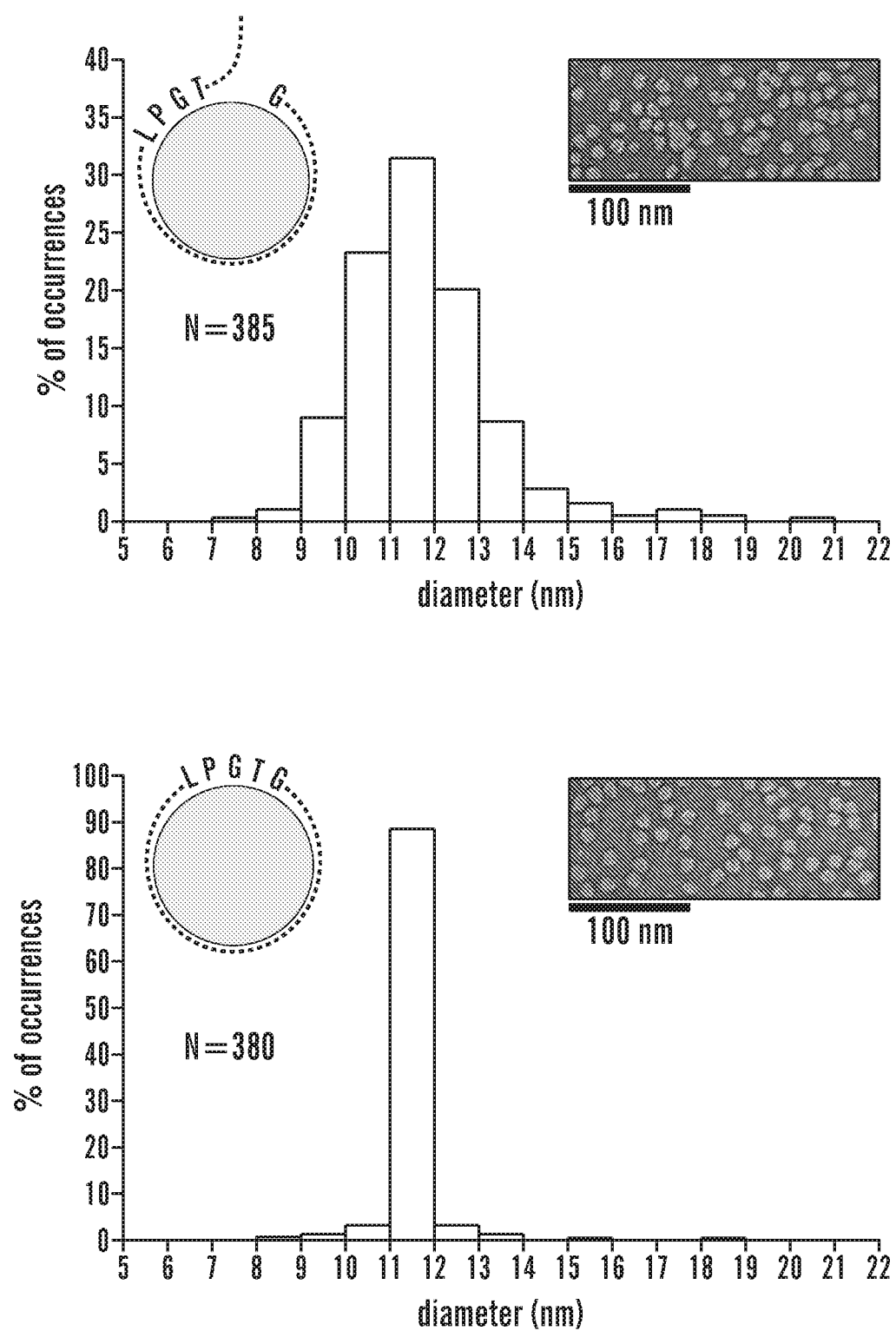


FIG. 2D

**FIG. 2E**

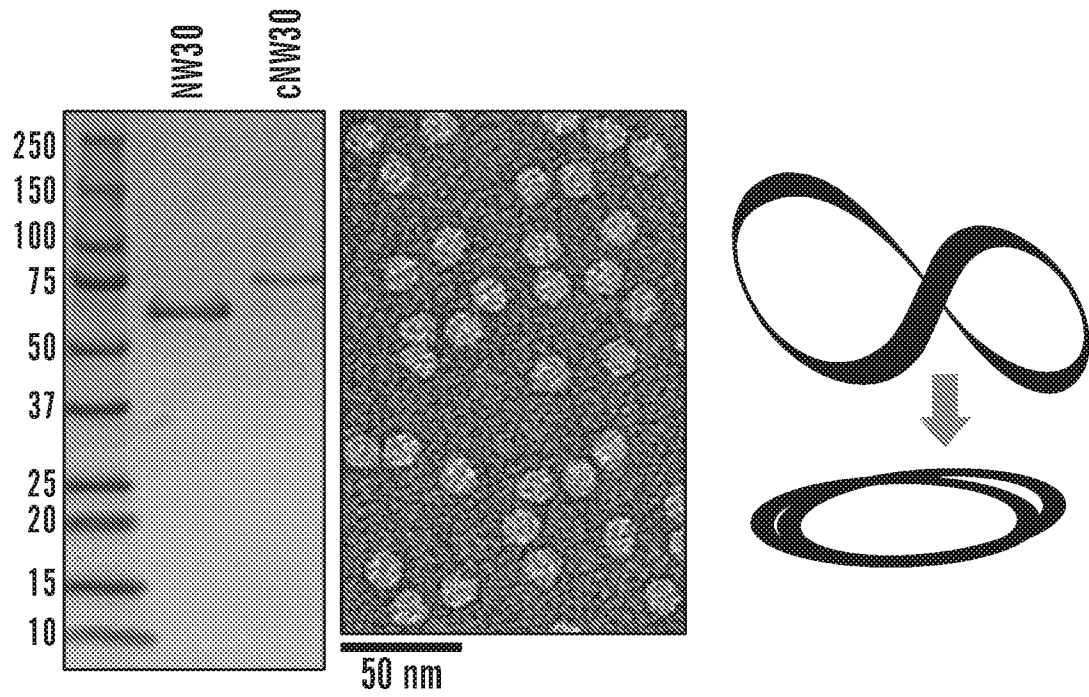


FIG. 3A

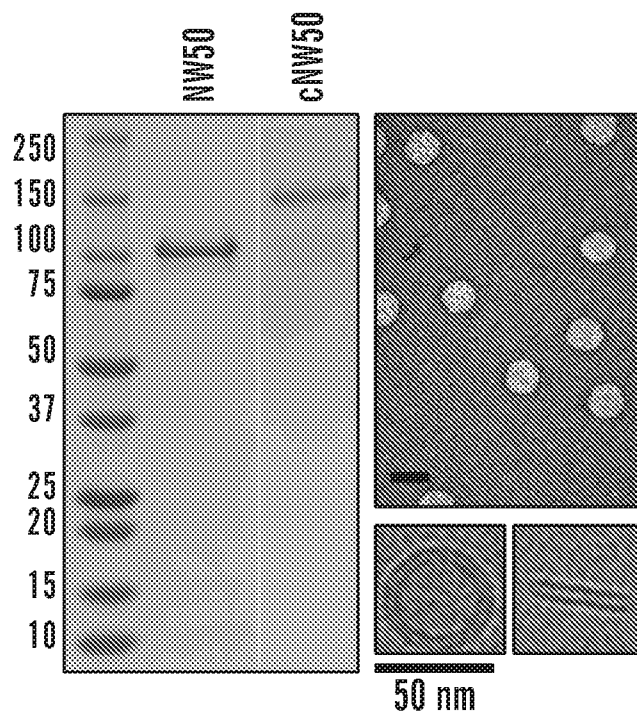


FIG. 3B



FIG. 3C

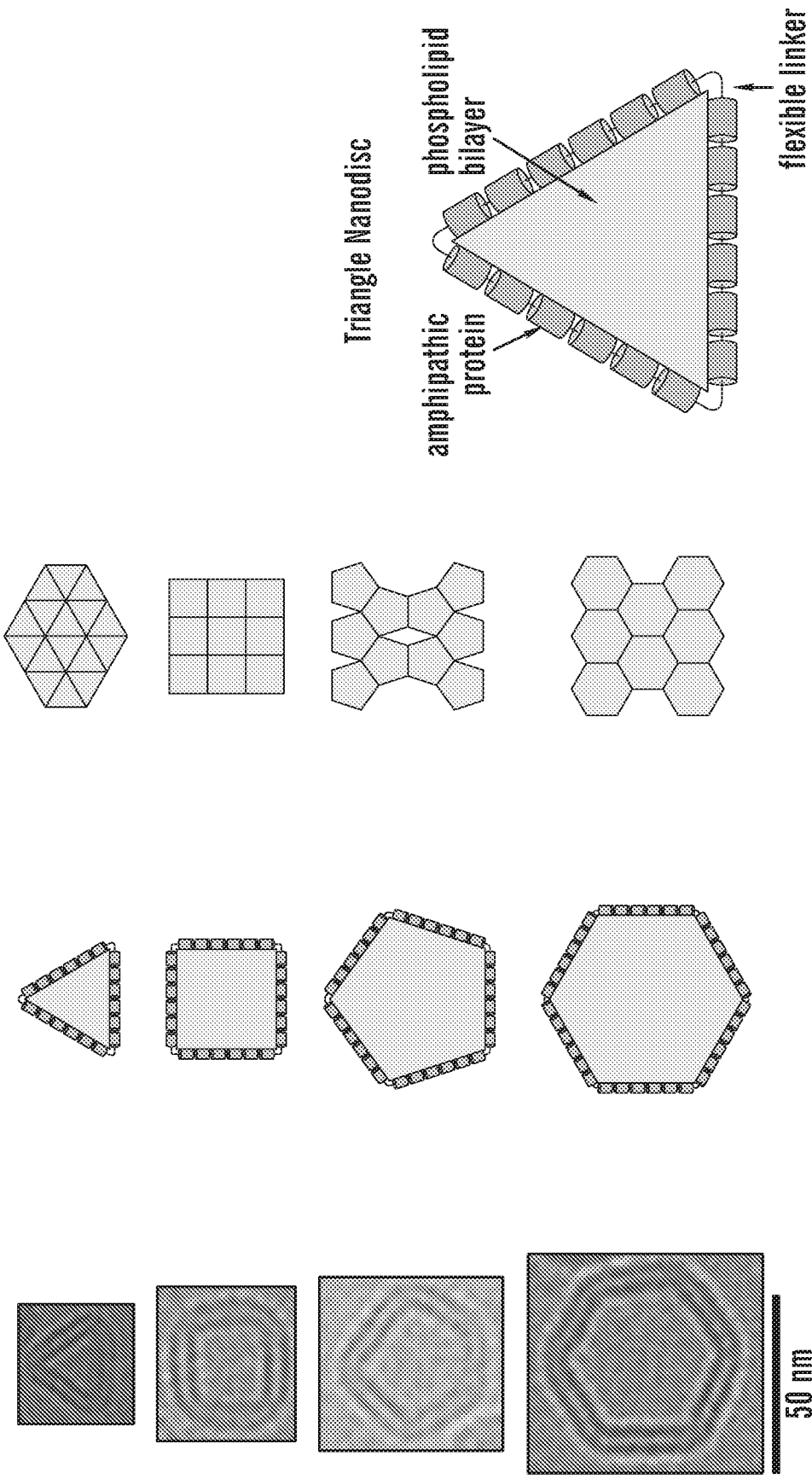


FIG. 4D

FIG. 4C

FIG. 4B

FIG. 4A

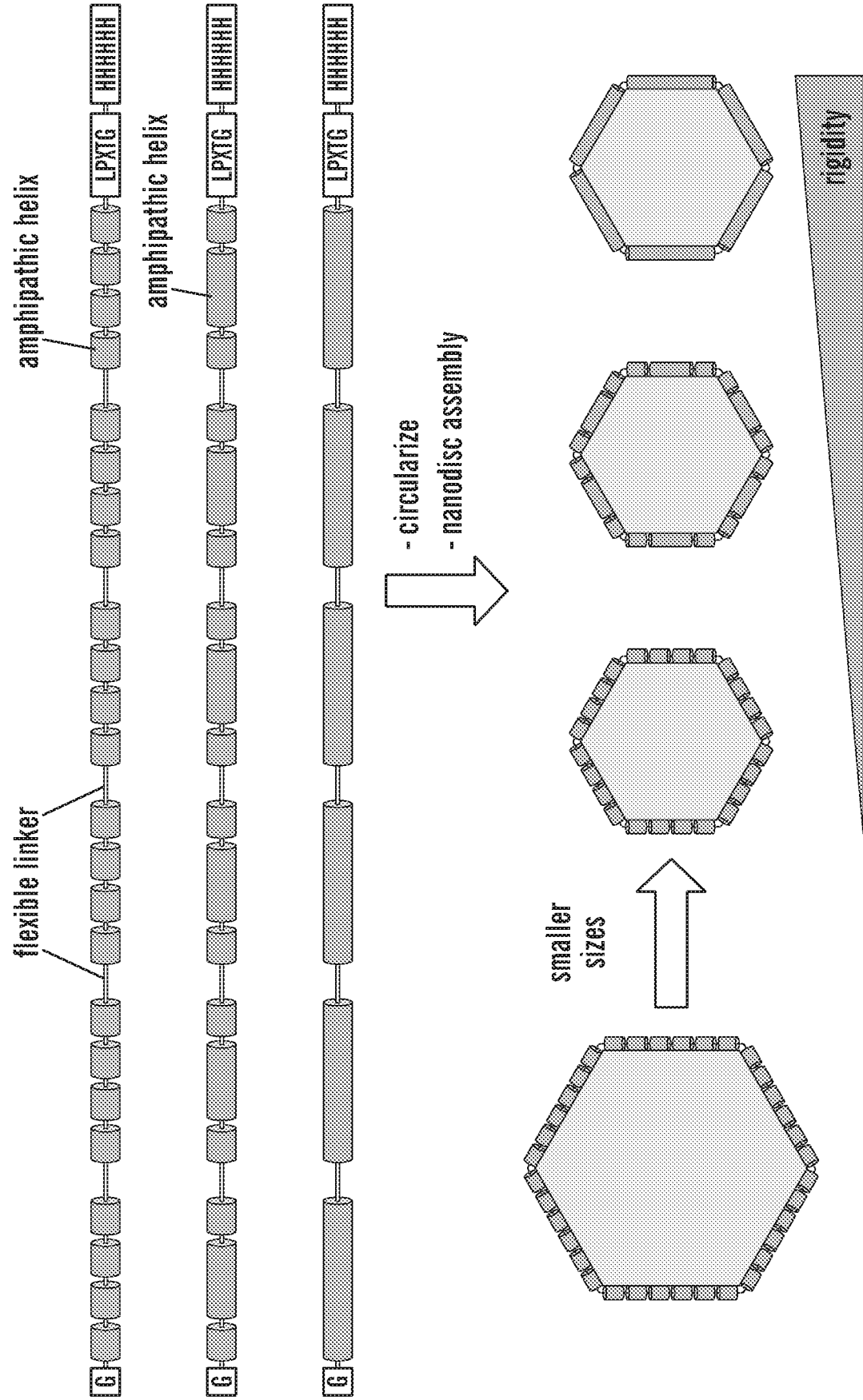
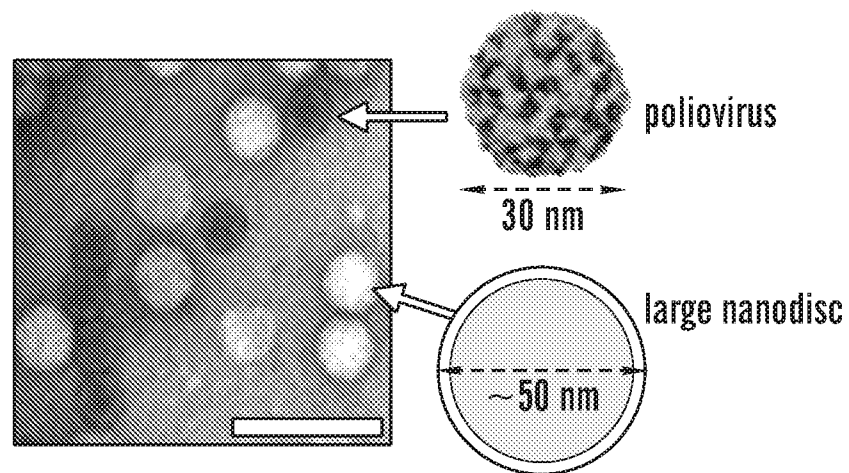
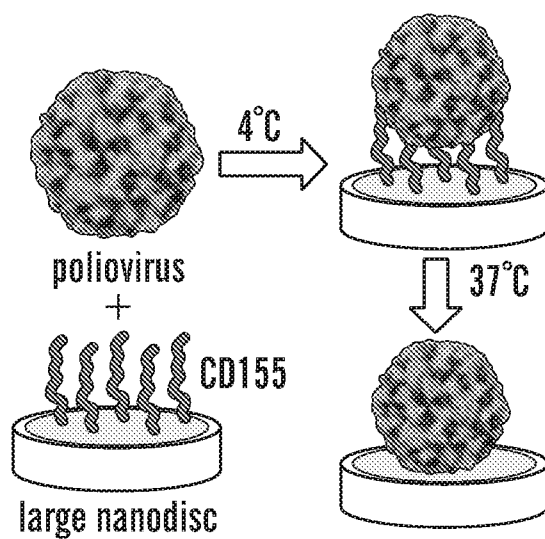
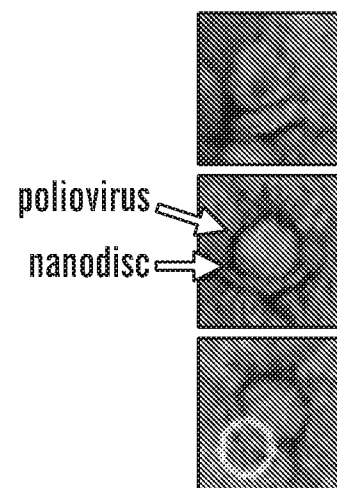


FIG. 5

**FIG. 6A****FIG. 6B****FIG. 6C**

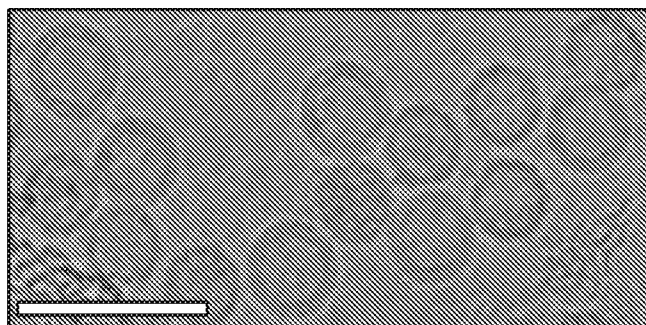


FIG. 6D

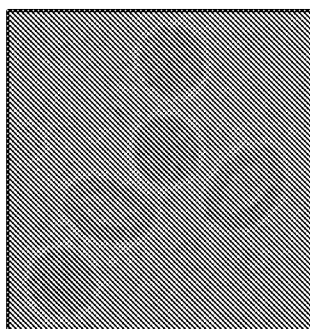


FIG. 6E

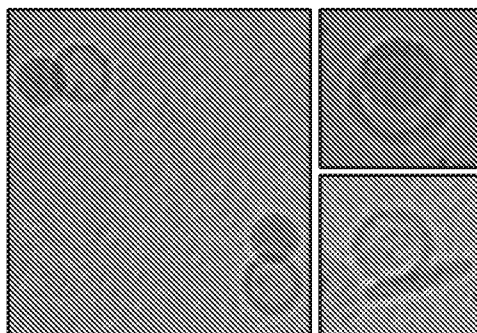


FIG. 6F

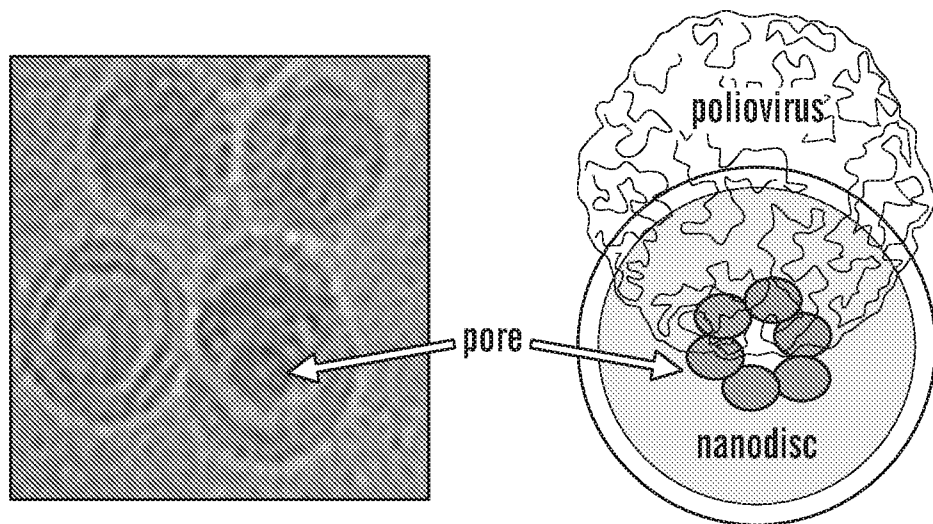


FIG. 6G

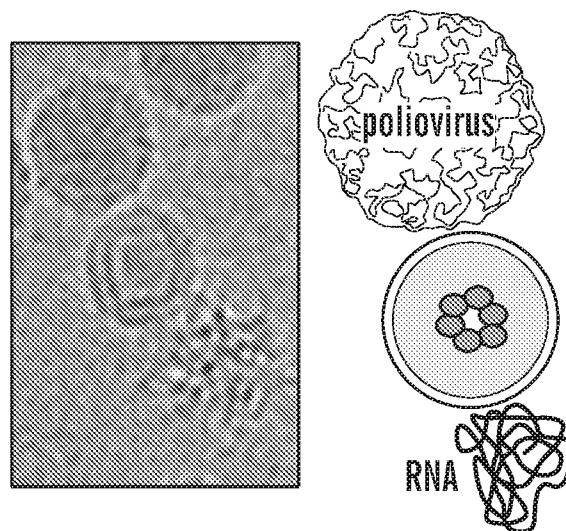


FIG. 6H

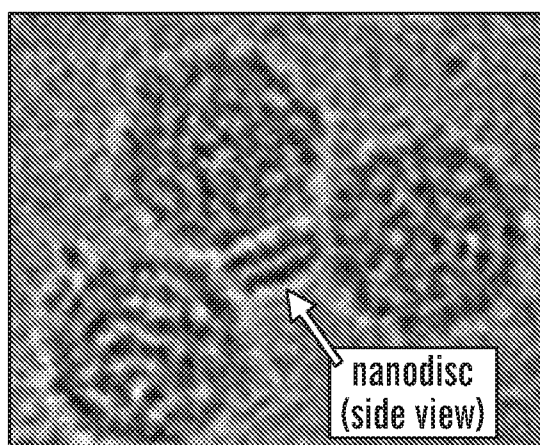


FIG. 6I

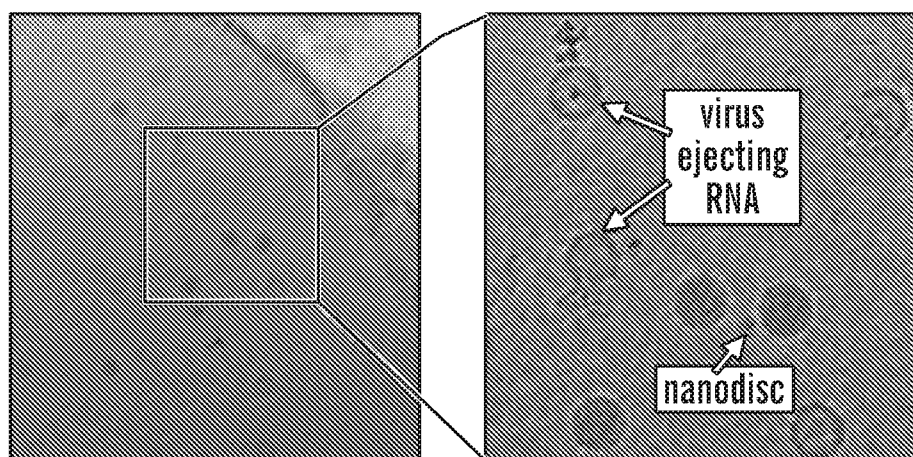
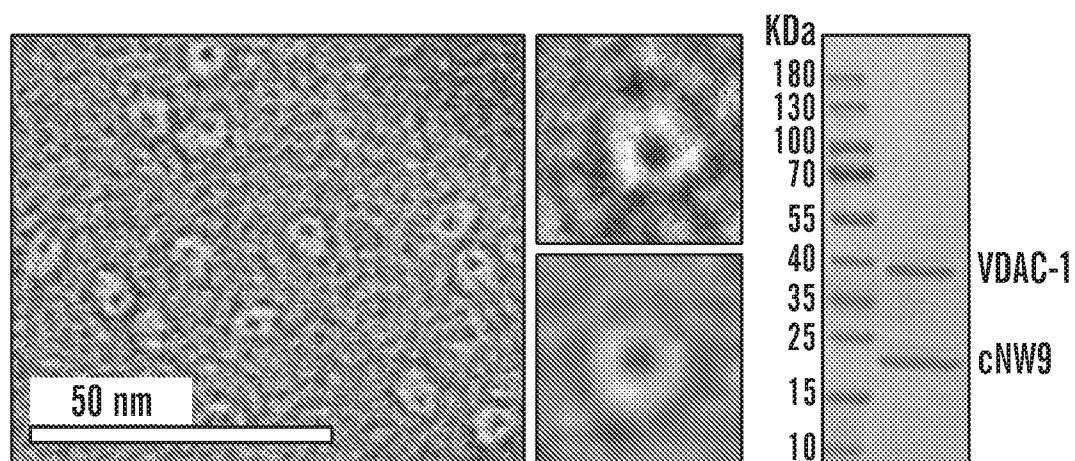
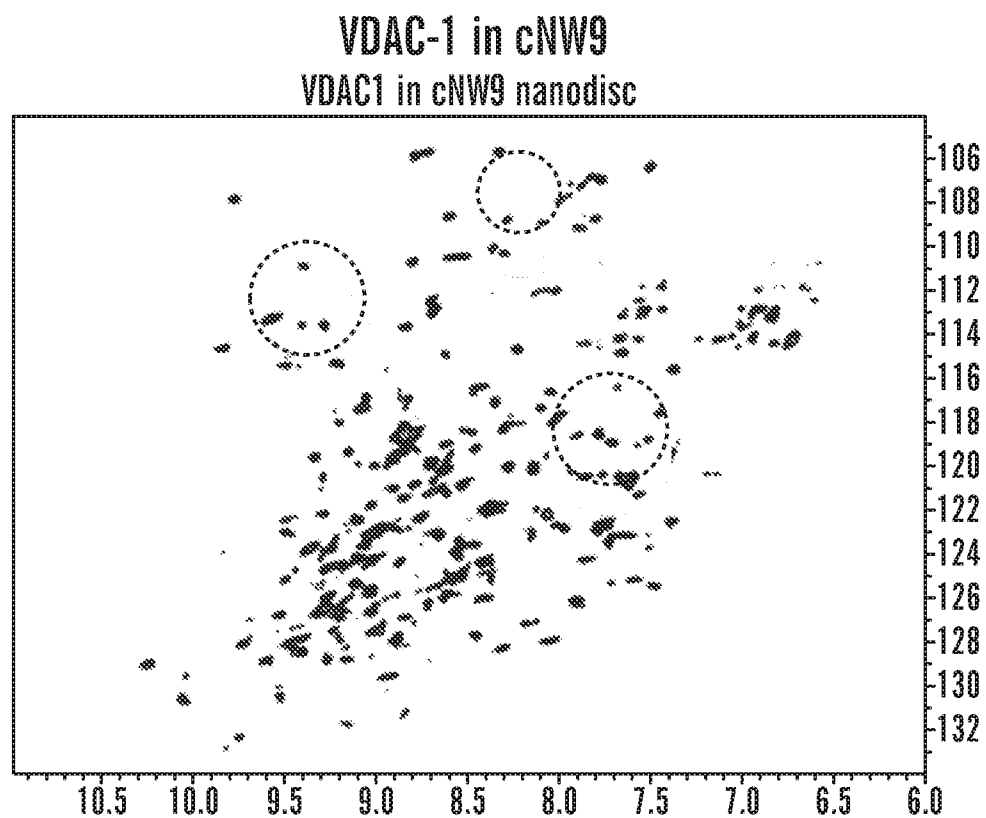


FIG. 6J



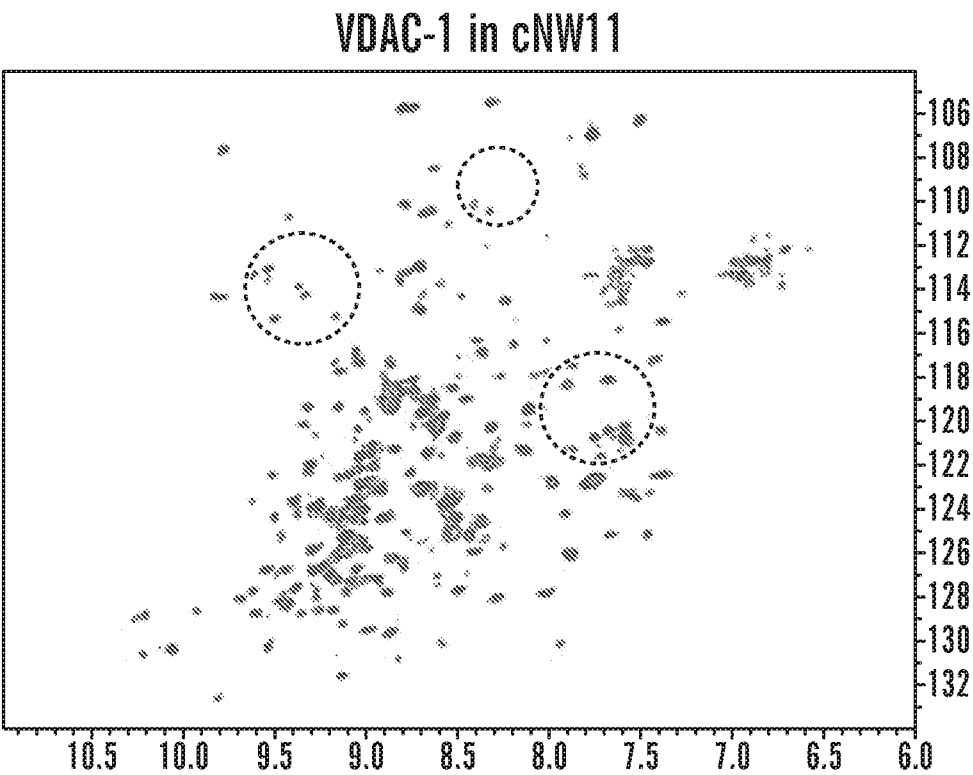


FIG. 7C

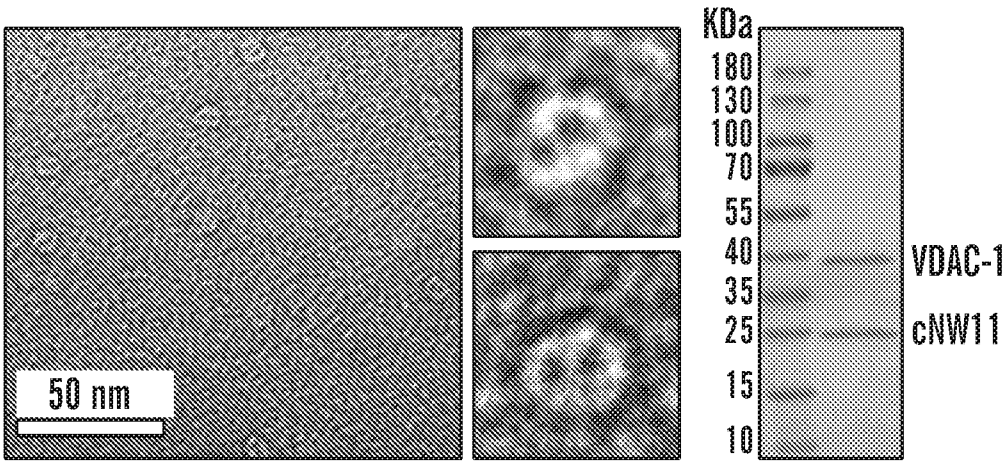


FIG. 7D

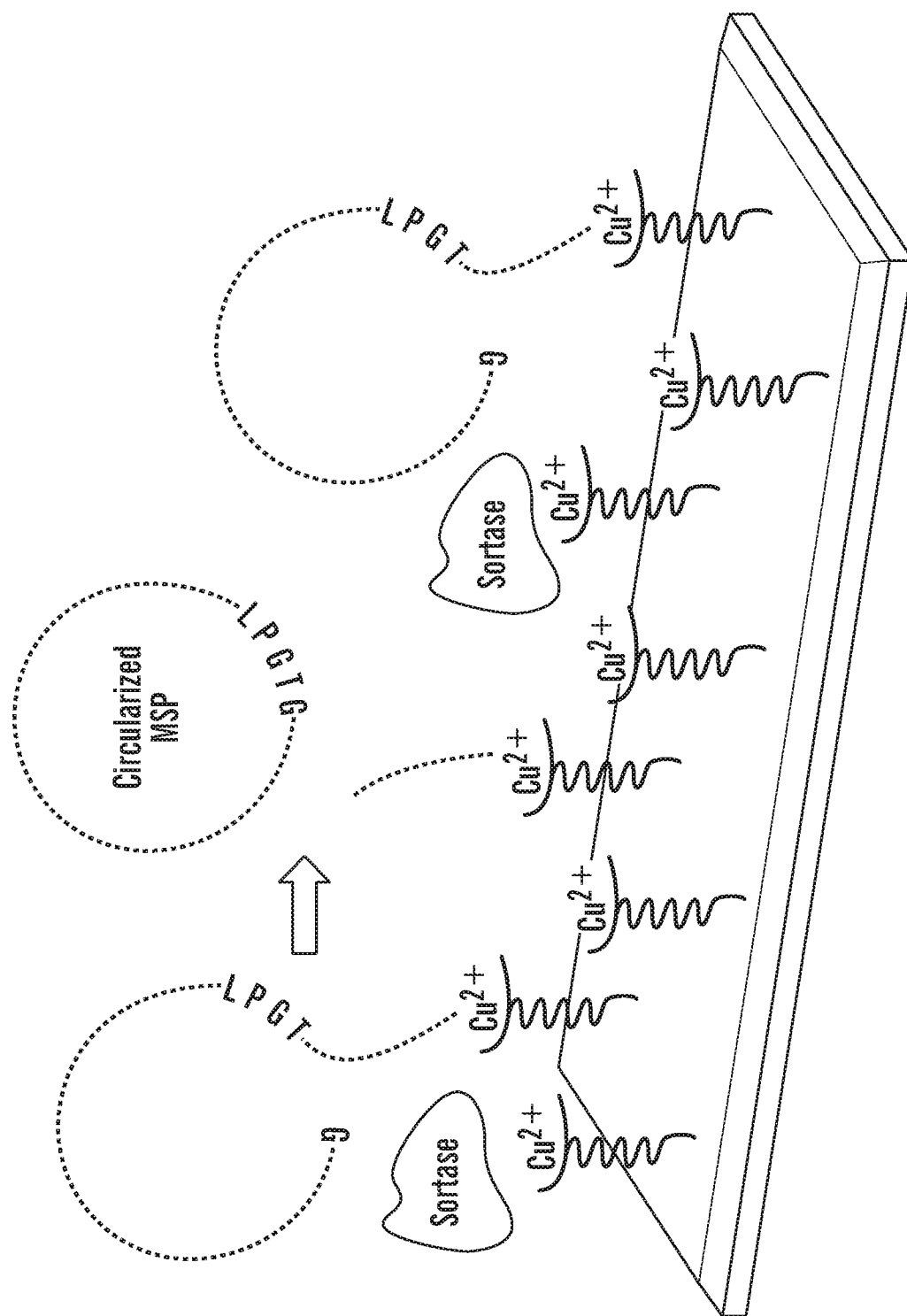


FIG. 8

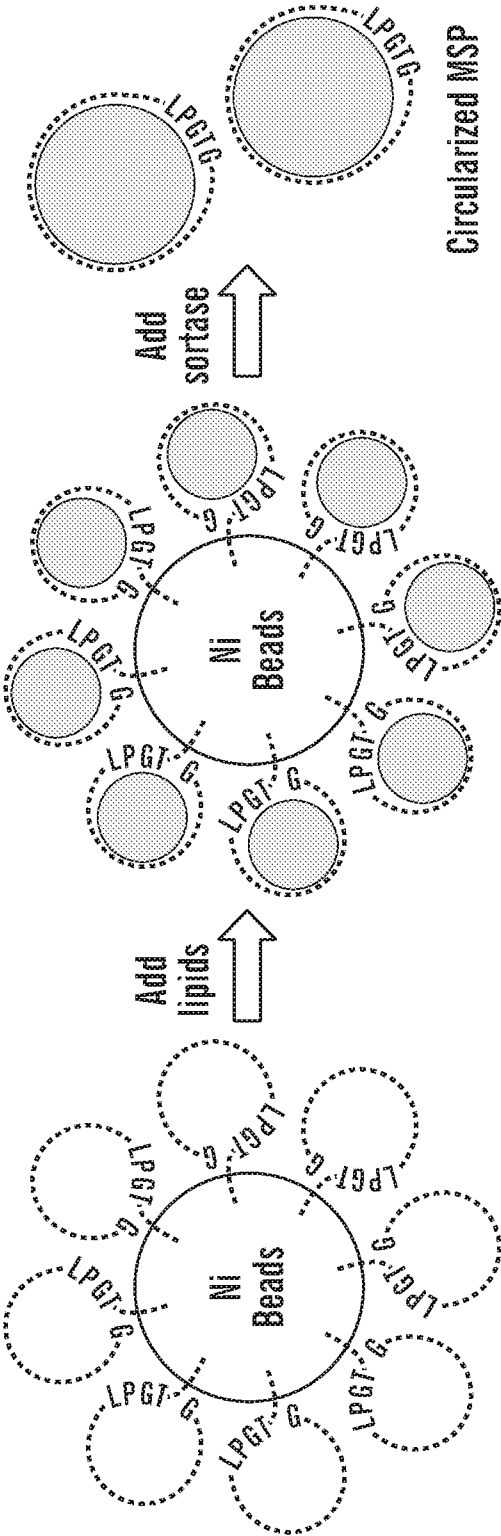


FIG. 9A

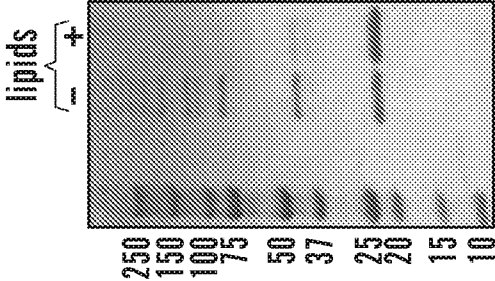


FIG. 9B

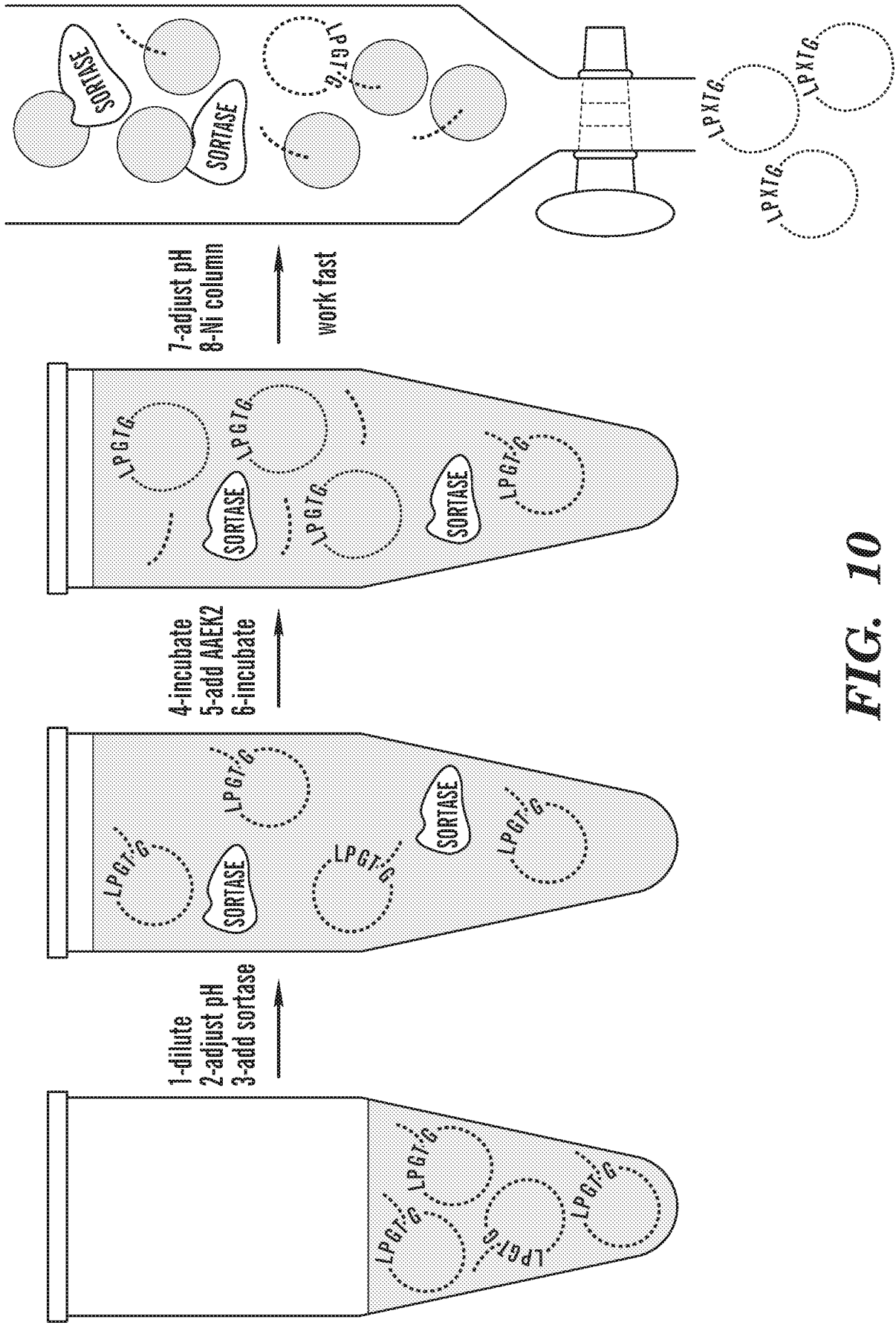


FIG. 10

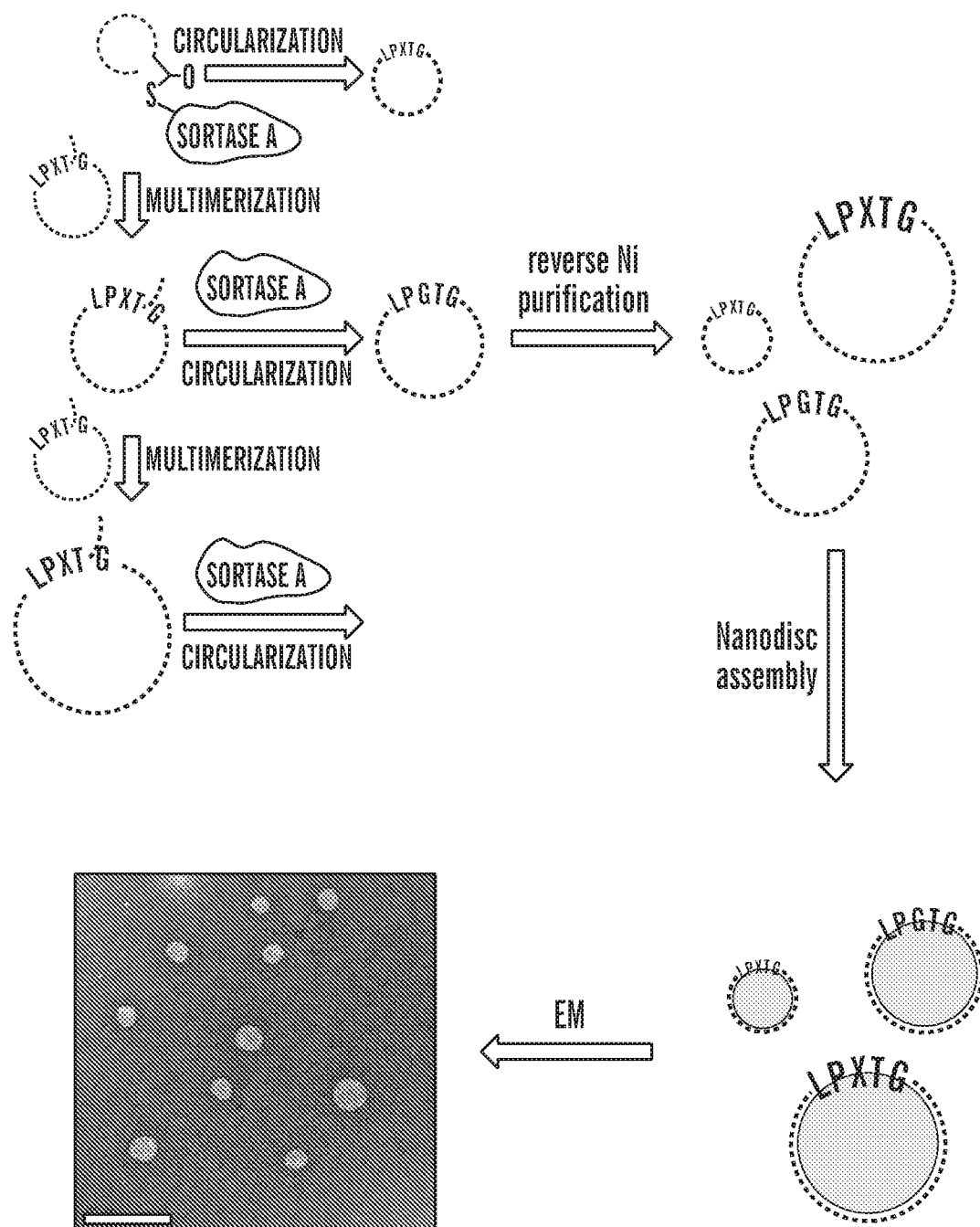


FIG. 11

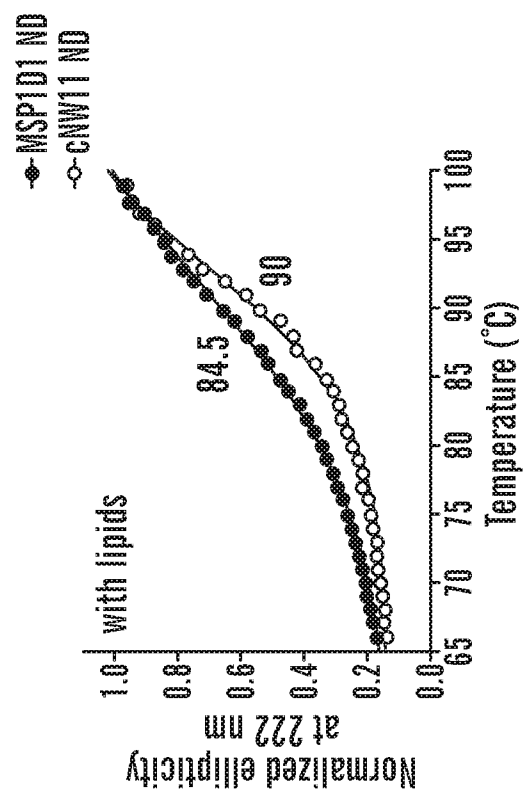


FIG. 12B

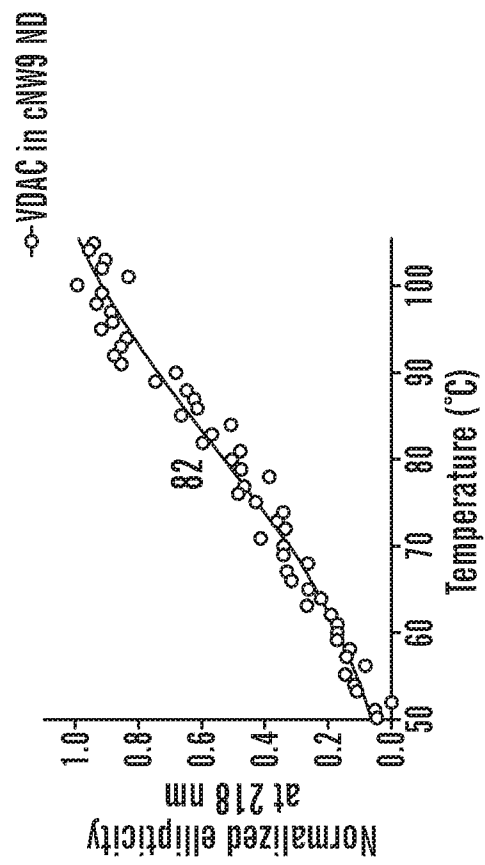


FIG. 12D

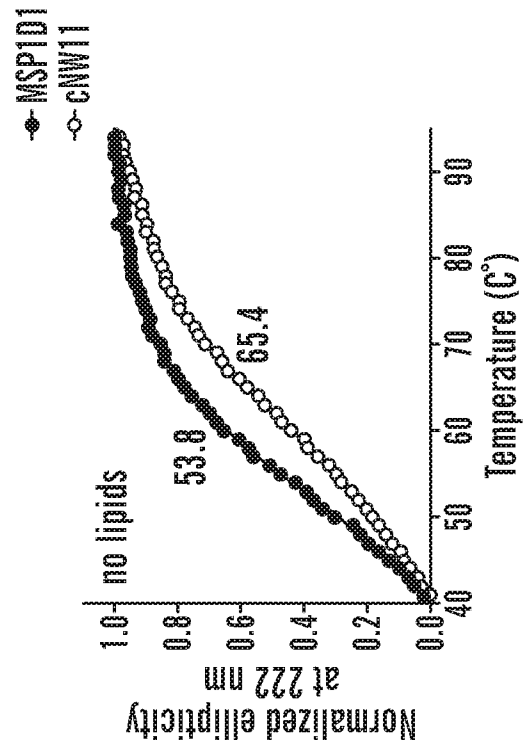


FIG. 12A

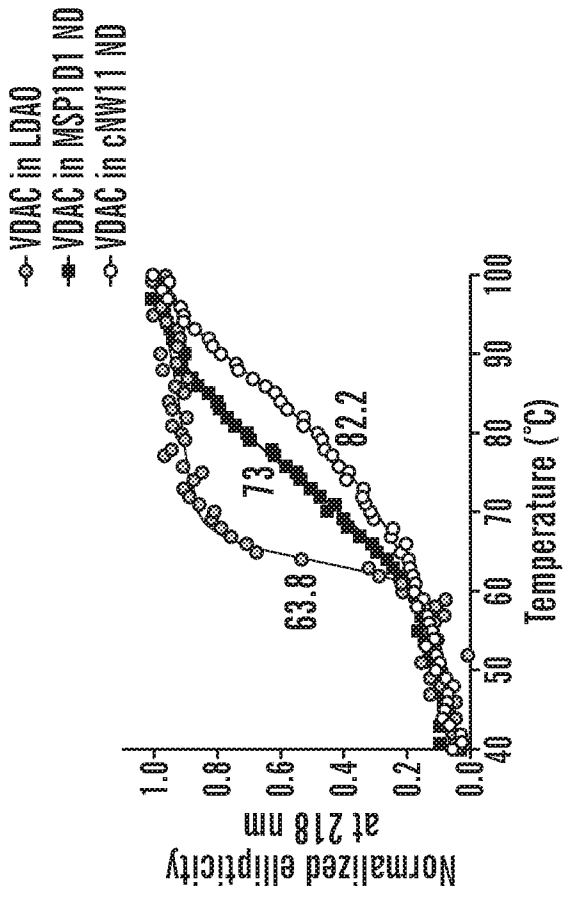
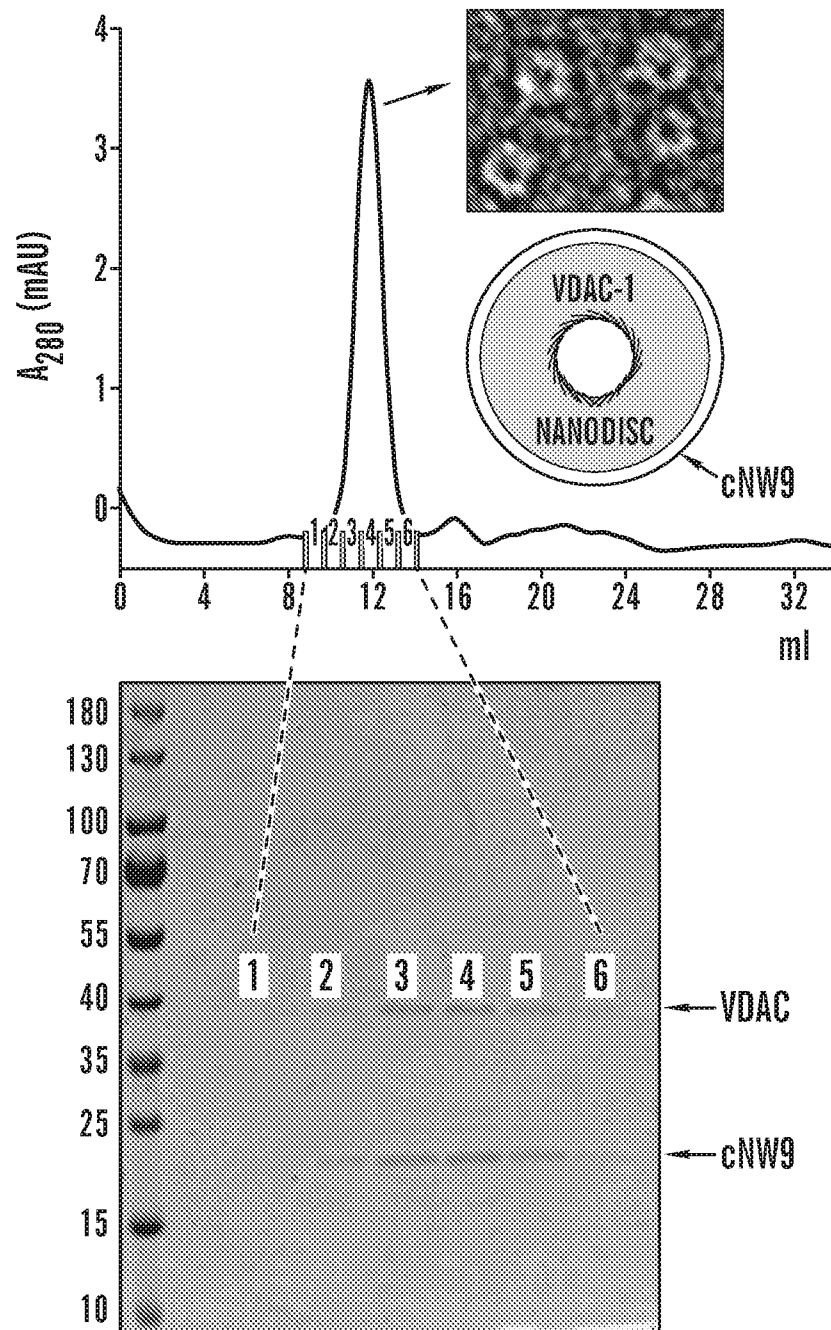


FIG. 12C

**FIG. 12E**

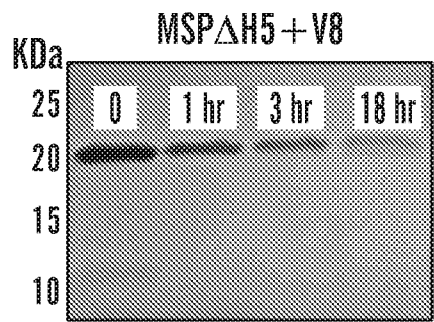


FIG. 13A

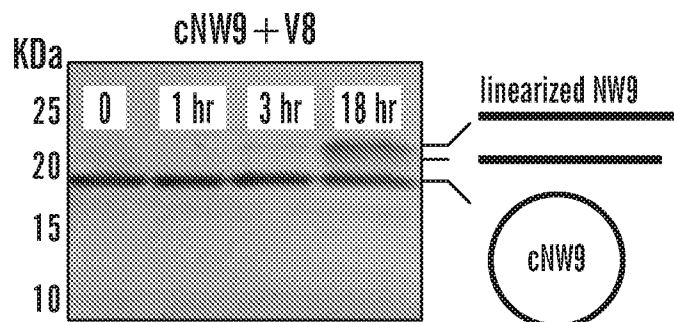


FIG. 13B

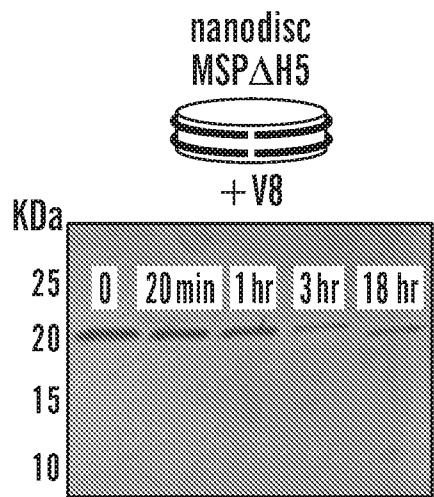


FIG. 13C

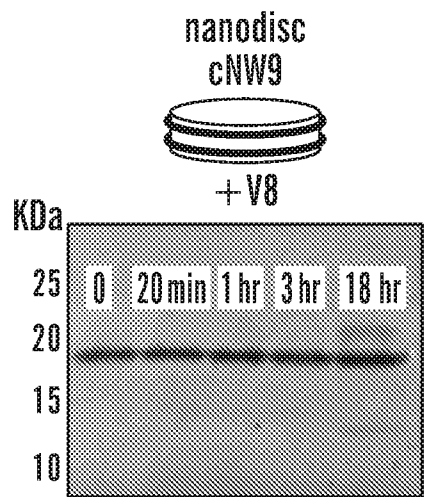


FIG. 13D

NW9 (amino acid sequence after TEV cleavage)
QSTFSKLR^QEQLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQE
EMELYRQKVEPLGEEMRDRAHVDALRTHLAPYSDELQRQLAARLEALKENGGARLA
EYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQLPGITGGAA
ALEHHHHHH



FIG. 14

NW30 (amino acid sequence after TEV cleavage)
GSTFSKLRQLGPVTOEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQE
EMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYS
ELRQRLAARLEALKENGARLAIEYHAKATEHLSTLSEKAKPALEDLRQGLIPVLESEFKVS
FLSALEEYTKLNTQGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKK
WQEEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAP
YSDELQRRLAARLEALKENGARLAIEYHAKATEHLSTLSEKAKPALEDLRQGLIPVLESEF
KVSFLSALEEYTKLNTQGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDF
QKKWQEEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRT
HLAPYSDELQRRLAARLEALKENGARLAIEYHAKATEHLSTLSEKAKPALEDLRQGLIPV
LESFKVSFLSALEEYTKLNTQLPGIGAAALEHHHHH

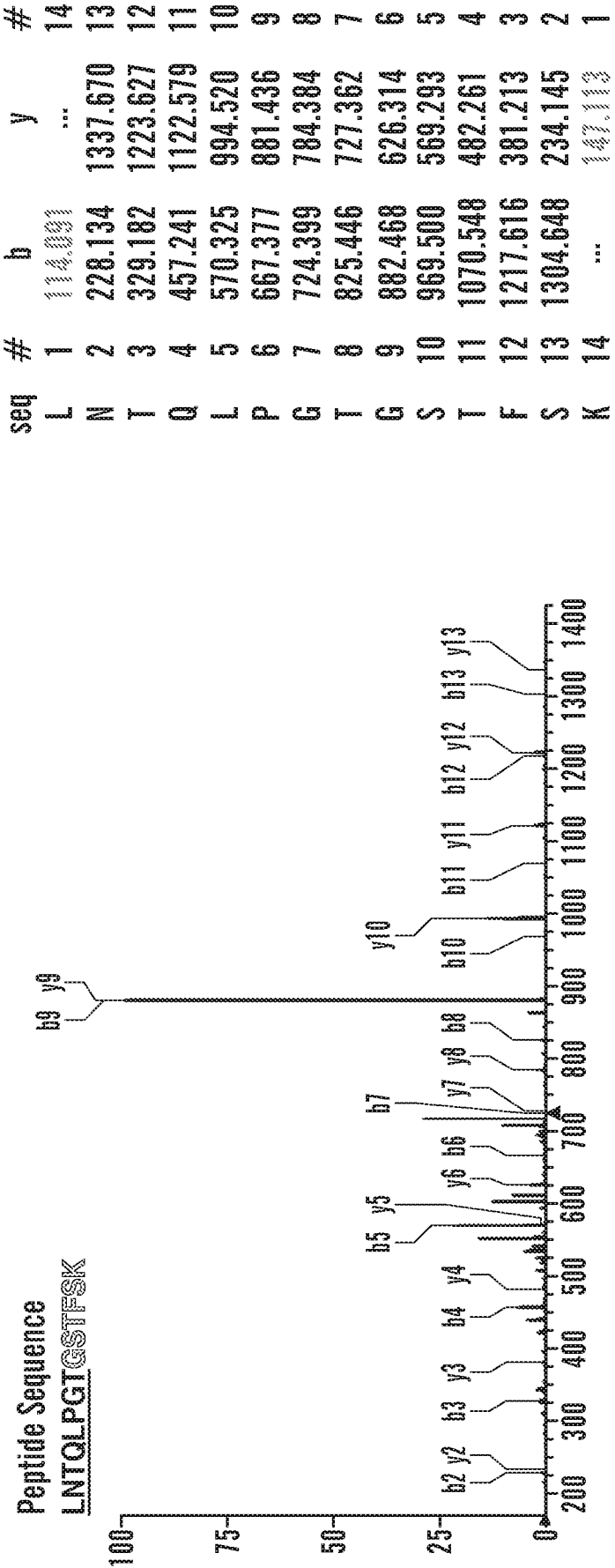


FIG. 15

NW50 (amino acid sequence after TEV cleavage)

GSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEM
ELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRO
RLAARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALE
EYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMEL
YRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRORL
AARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELY
RQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRORLA
ARLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYR
QKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRORLAA
RLEALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEY
TKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQ
KVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRORLAARL
EALKENG GARLA EYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKL
NTQLPGTGAAALEHHHHHH

FIG. 16

seq	#	b	y	#
L	1	114.001	...	14
N	2	228.134	1337.070	13
T	3	329.182	1223.627	12
Q	4	457.241	1122.579	11
L	5	570.325	994.520	10
P	6	667.377	881.436	9
G	7	724.399	784.384	8
T	8	825.446	727.362	7
G	9	882.468	626.314	6
S	10	969.500	500.293	5
T	11	1070.548	402.261	4
F	12	1217.616	381.213	3
S	13	1304.648	234.145	2
K	14	...	147.113	1

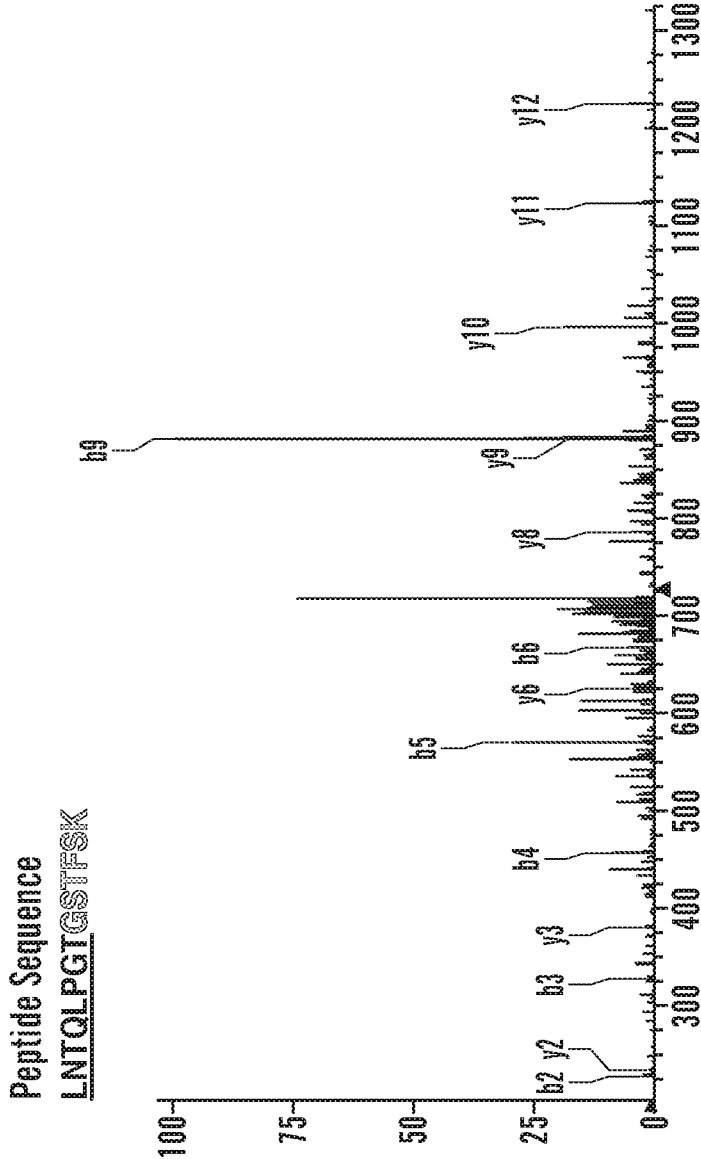
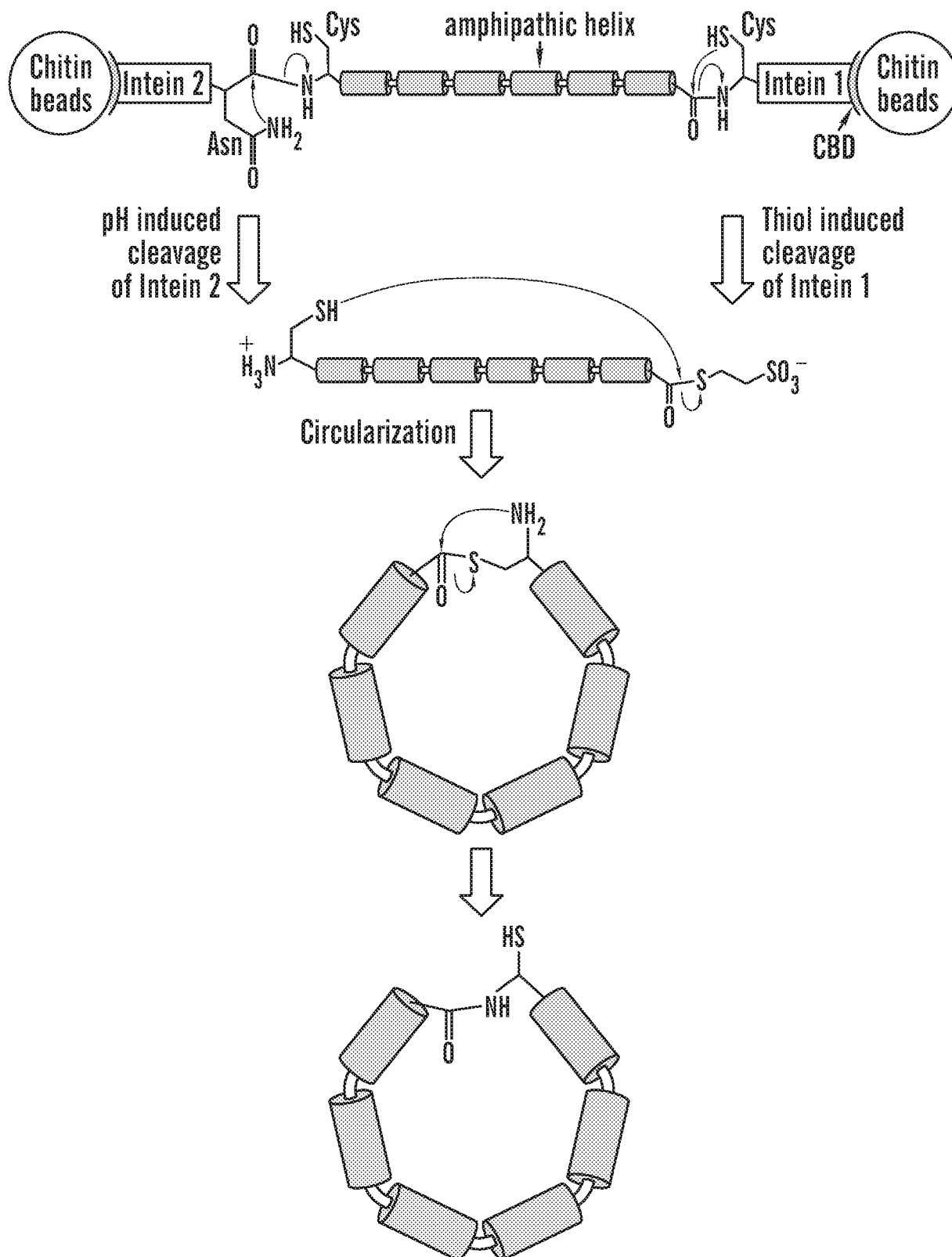
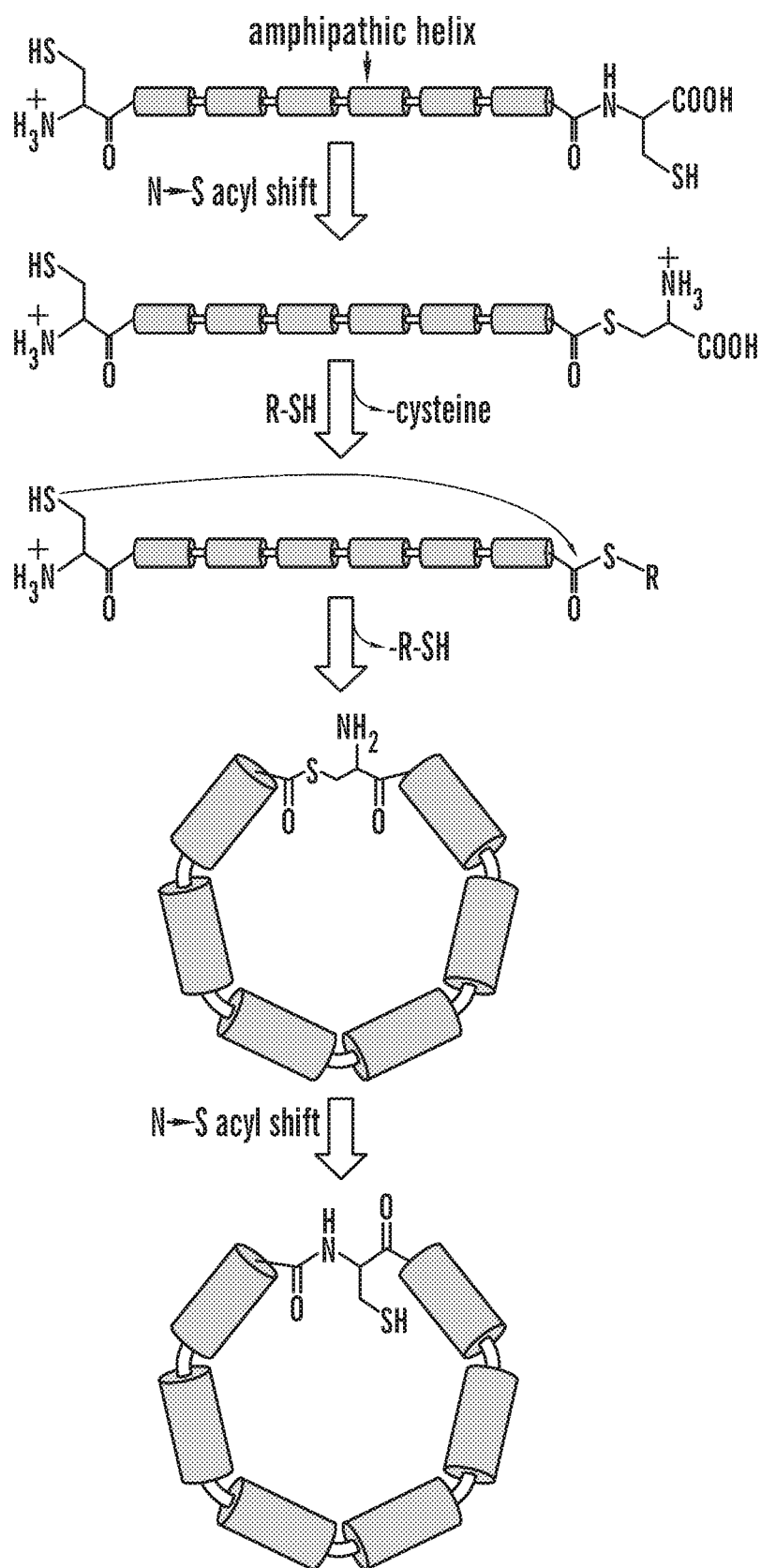


FIG. 16 (cont.)

**FIG. 17**

**FIG. 18**

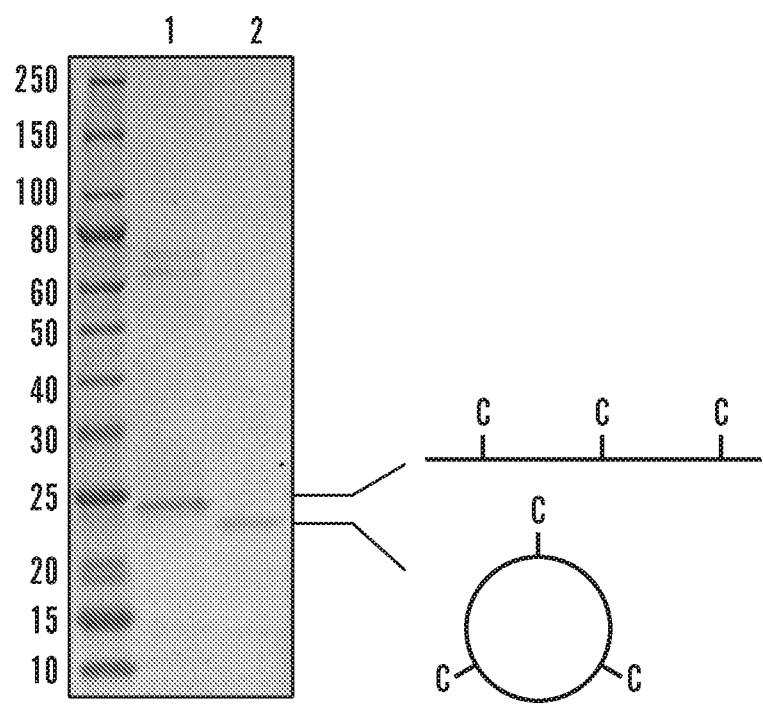


FIG. 19

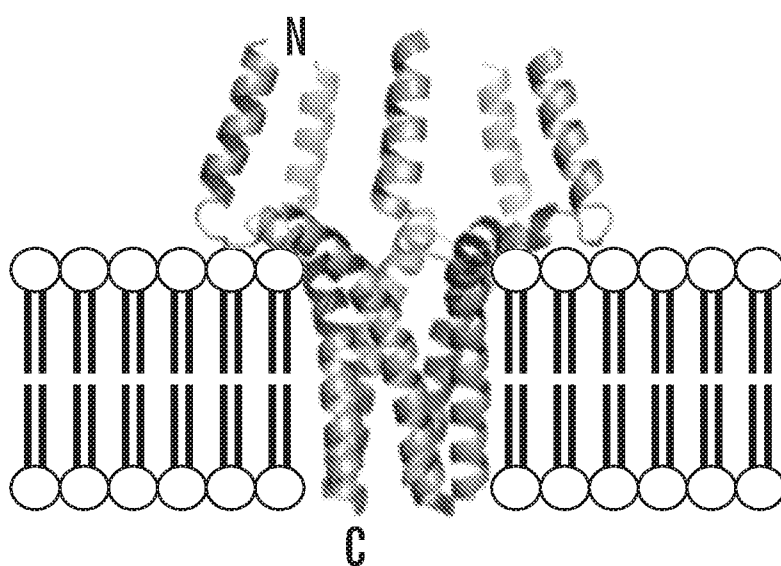


FIG. 20A

	Cytoplasmic domain	Transmembrane domain	Luminal domain
Human	MEKVQYLTRSAIRRASTIEMPQQARQKLQNLF	INFCLILICLLLCIIIVMLL	
Chimpanzee	MEKVQ H LTRSAIRRASTIEMPQQARQKLQNLF	INFCLILICLLLCIIIVMLL	
Mouse	MEKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	
Rat	MEKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	
Rabbit	MEKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	
Pig	MDKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	
Rhesus monkey	MEKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	
Bovine	MDKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	
Chicken	MEKVQY I TRSA L RRAST LEVNP QARQ RLQEL VNF	CLILICLLLCIIIVMLL	
Horse	MEKVQYLTRSAIRRASTIEMPQQARQ N LQNLF	INFCLILICLLLCIIIVMLL	

FIG. 20B

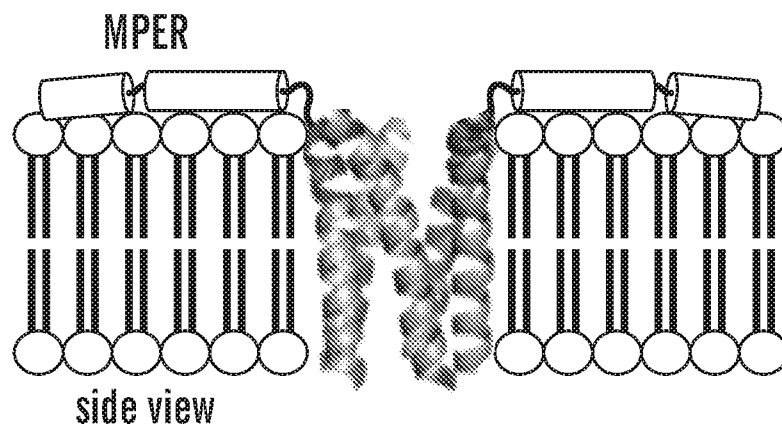
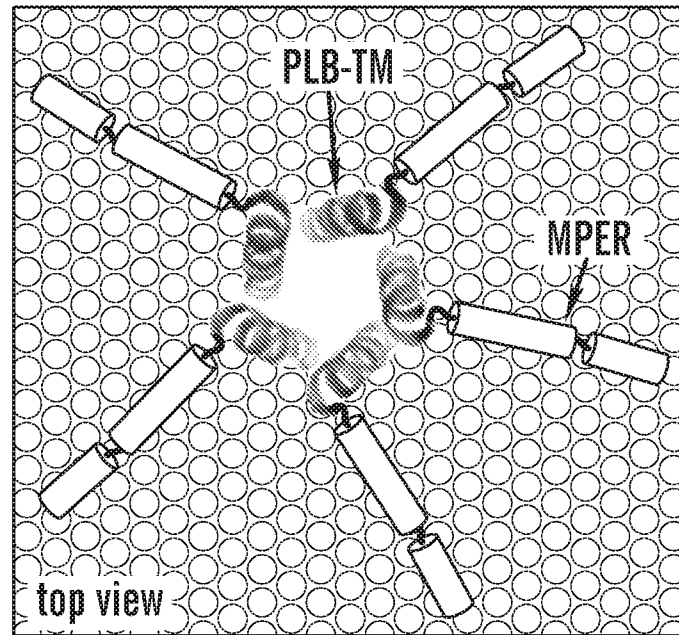


FIG. 21A

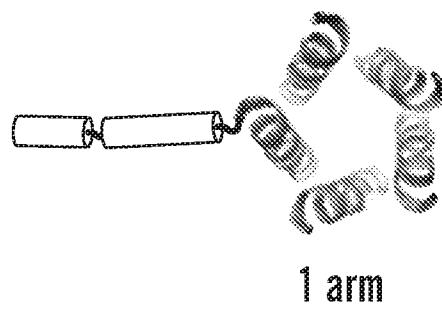
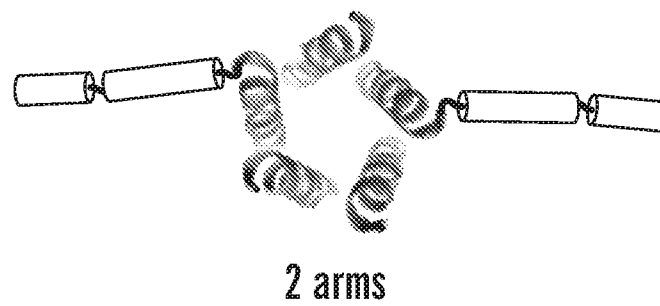
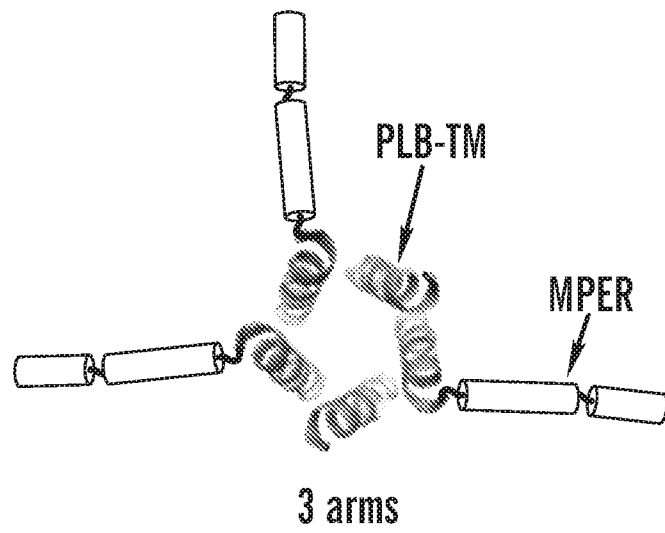


FIG. 21B

PLB-MPER in nanodisc + 10E8 fab

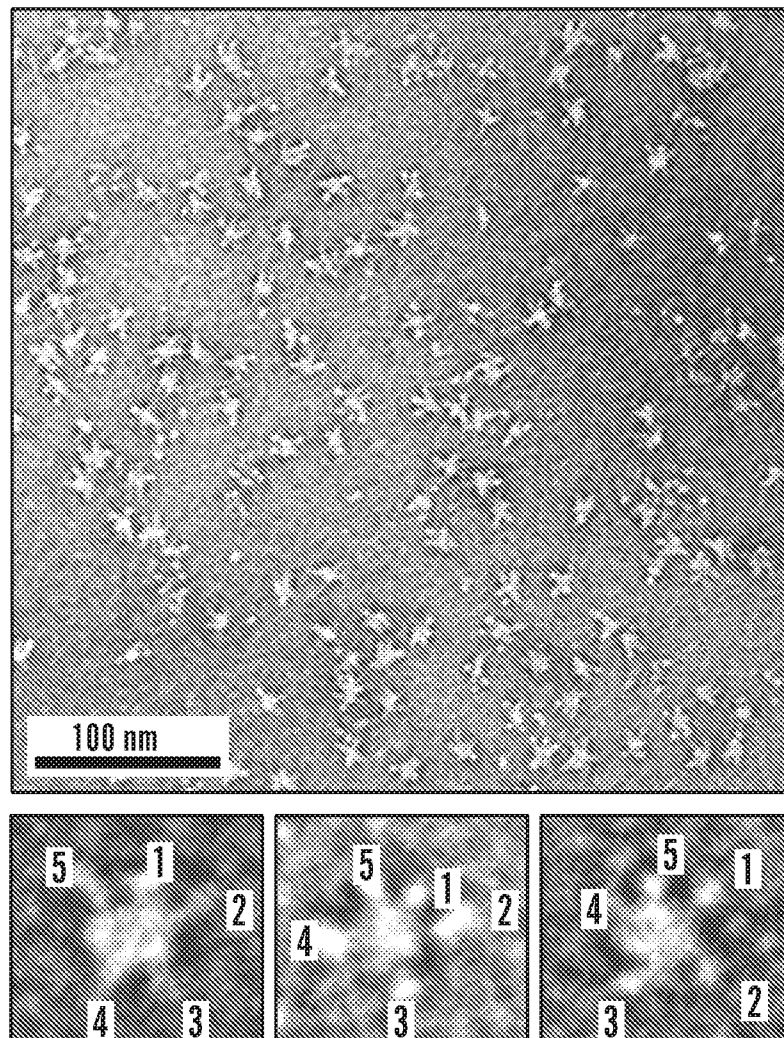


FIG. 22A

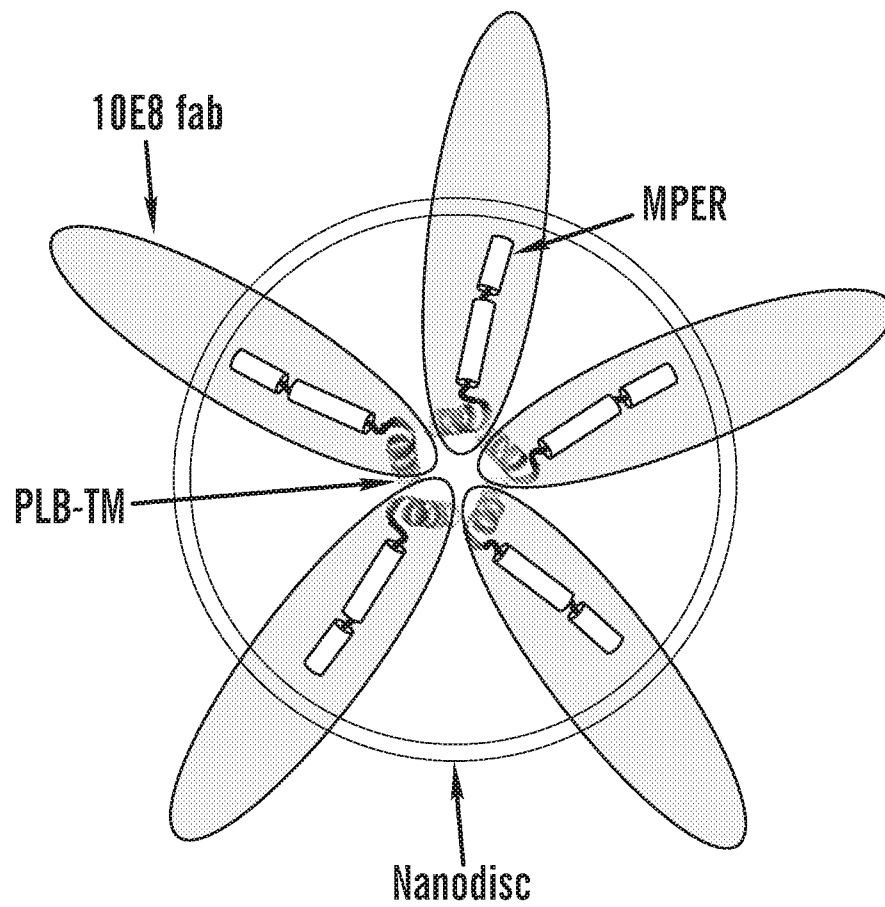


FIG. 22B

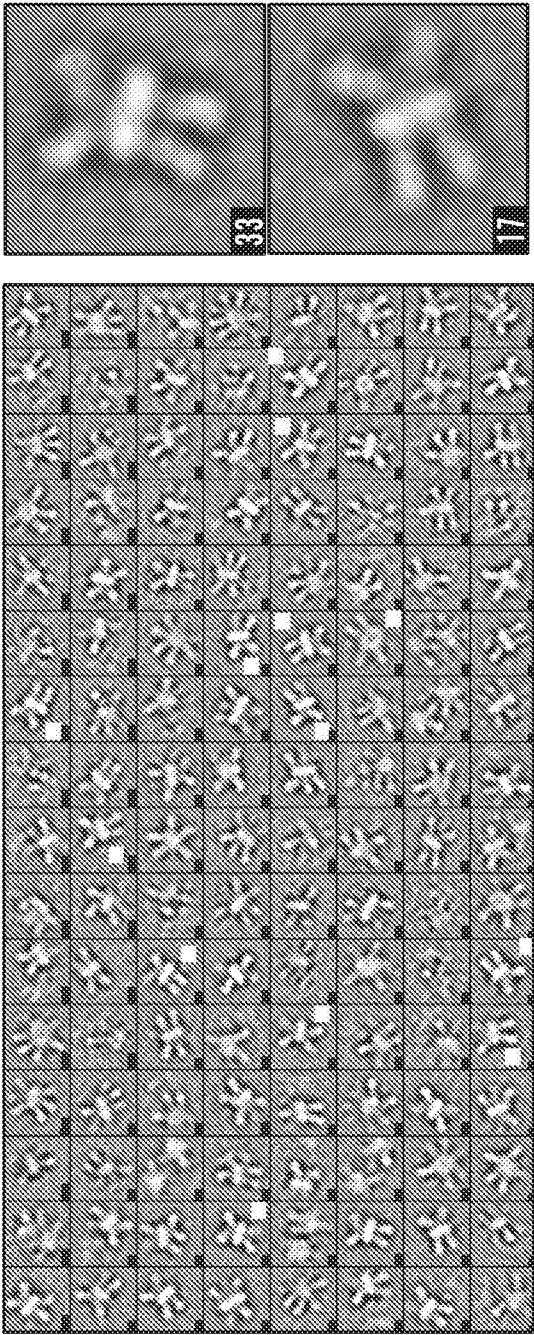


FIG. 23A

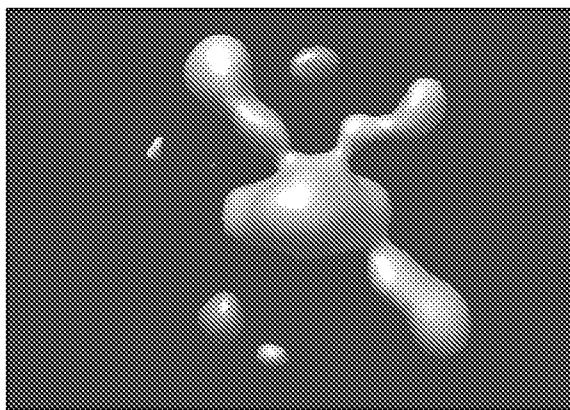


FIG. 23B

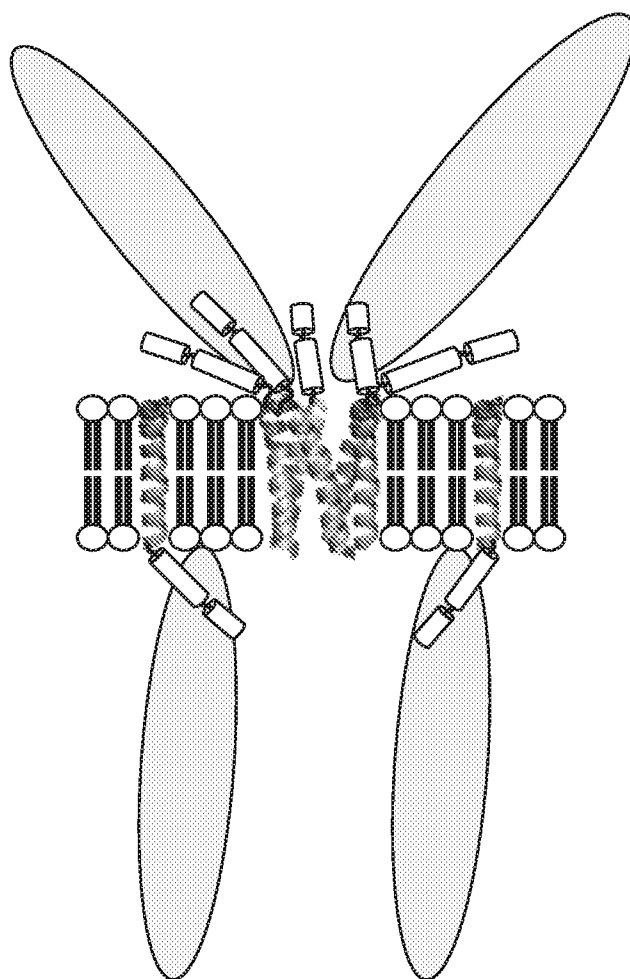


FIG. 23C

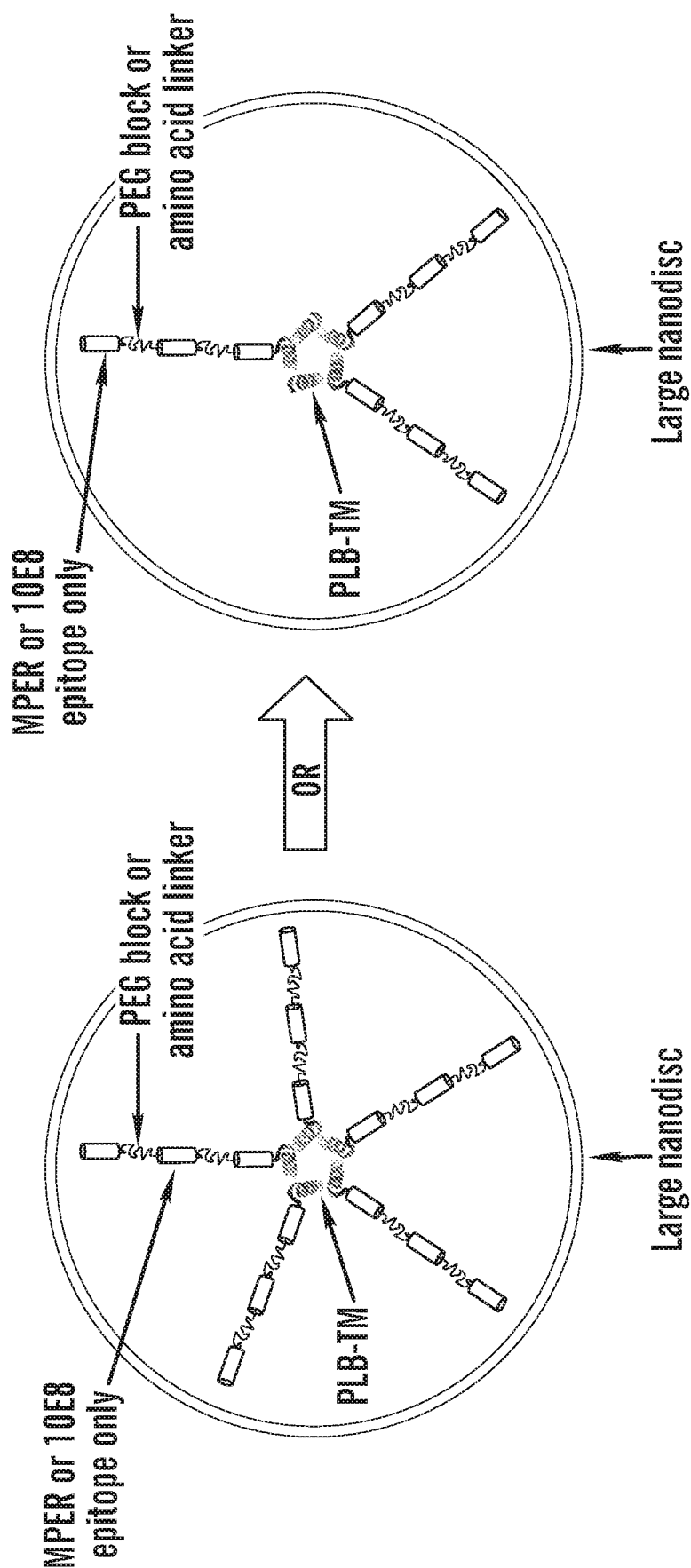


FIG. 24

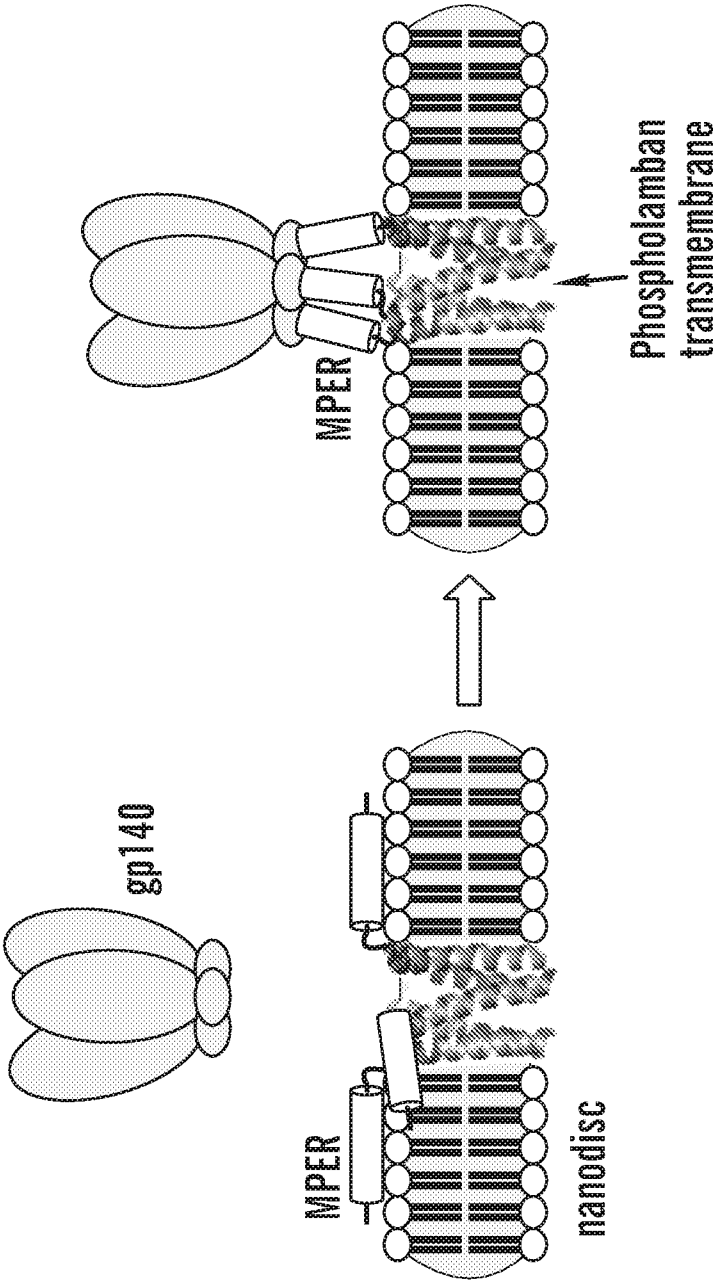


FIG. 25A

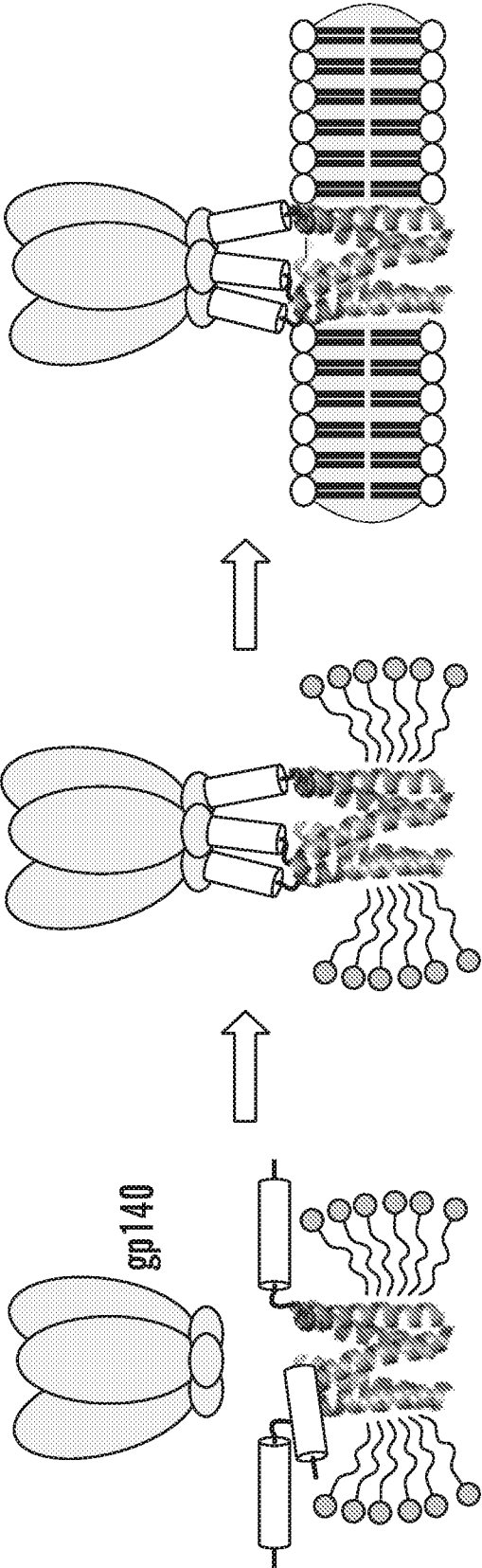


FIG. 25B

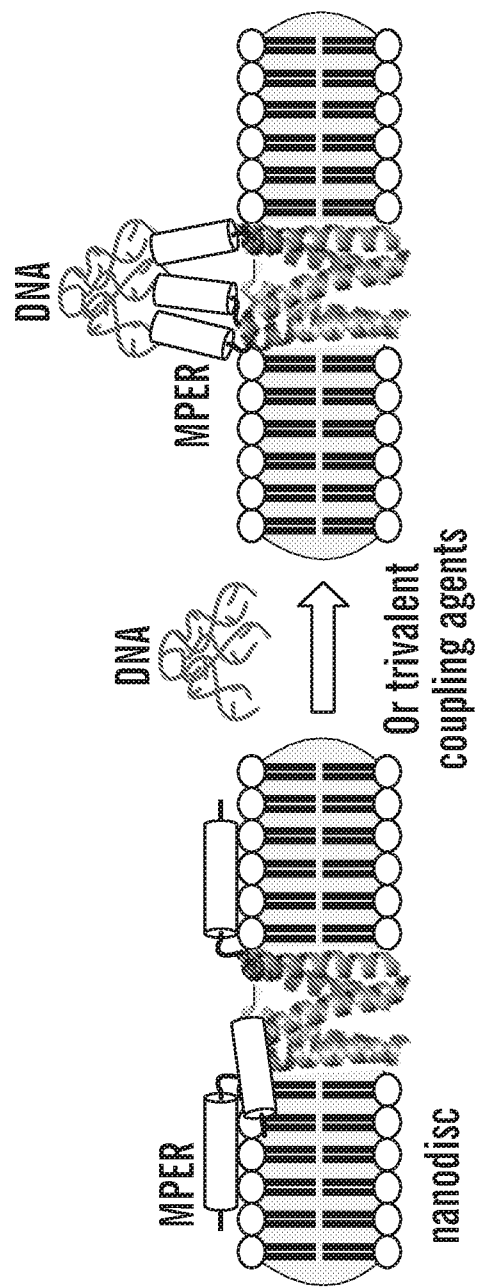


FIG. 26A

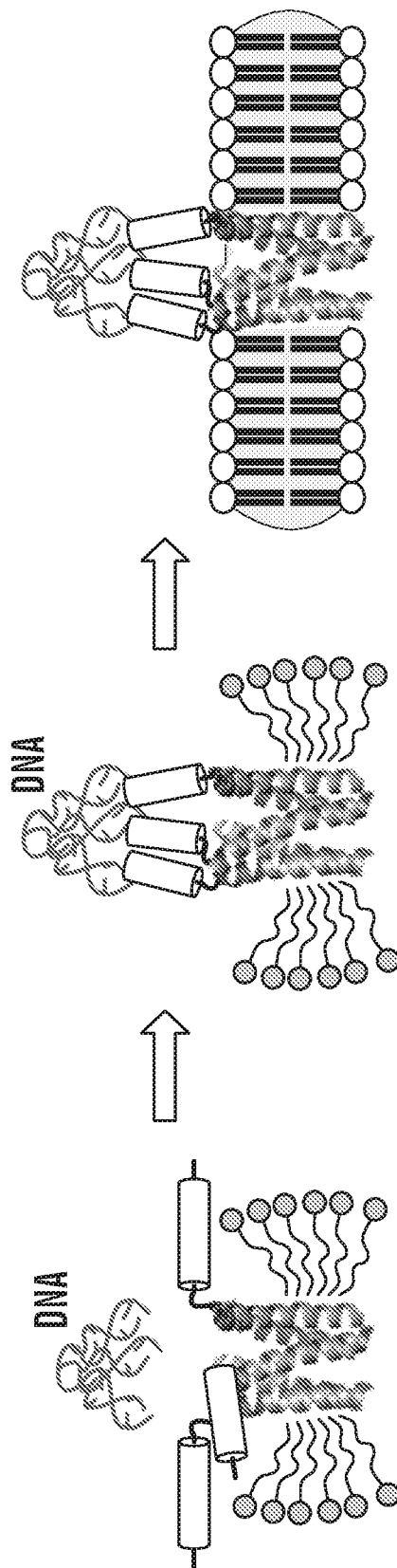
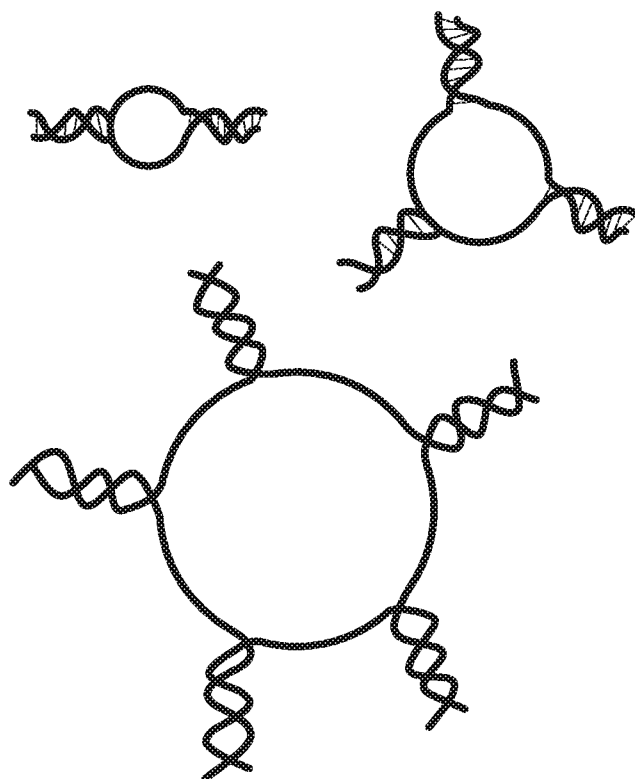
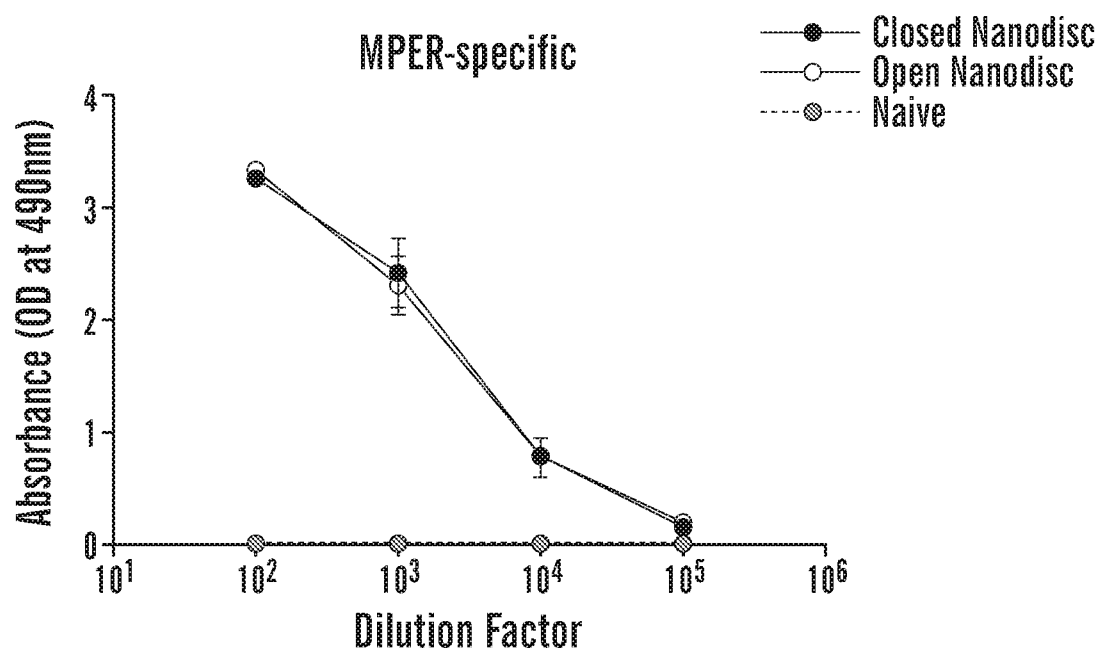
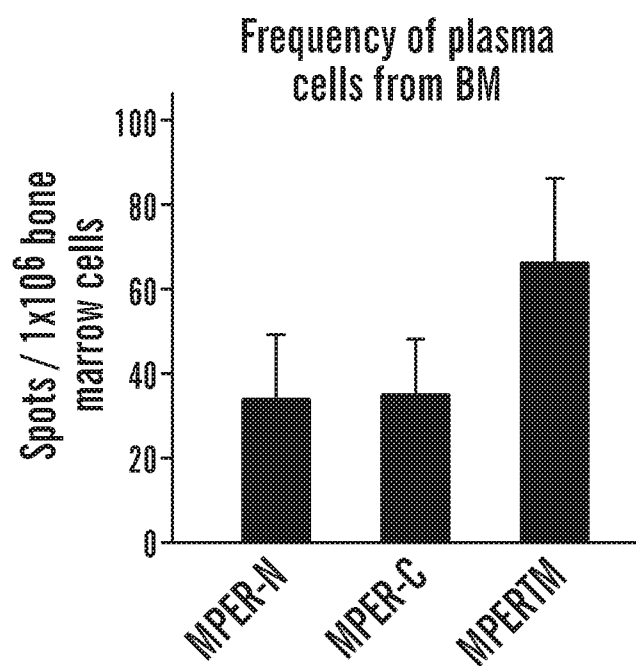
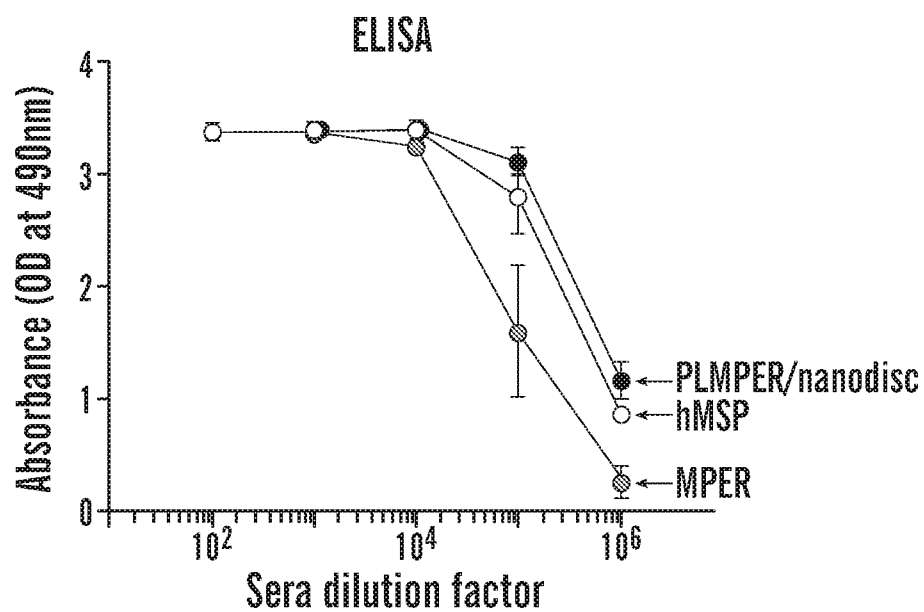


FIG. 26B

**FIG. 26C****FIG. 27**

**FIG. 28**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/42308

A. CLASSIFICATION OF SUBJECT MATTER
 IPC(8) - A61K 39/00, C12N 7/00 (2017.01)
 CPC - A61K 39/12, A61K 39/145, C12N 15/09

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	US 2011/0305723 A1 (RENNER et al.) 15 December 2011 (15.12.2011) abstract, para [0159], [0167], [0168], [0309], SEQ ID NO: 213	1-4
A	US 2013/0337002 A1 (LANGE et al.) 19 December 2013 (19.12.2013) whole document, SEQ ID NO: 9	1-4



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

"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

22 September 2017

Date of mailing of the international search report

05 OCT 2017

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
 P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/42308

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.: 5-38
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:

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.:

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.