Self-assembling protein nanostructures displaying paramyxovirus and/or pneumovirus F proteins and their use

Cross Reference

This application claims priority to U.S. Provisional Patent Application Serial No. 62/481331 filed April 4, 2017, incorporated by reference herein in its entirety.

Background

Molecular self- and co-assembly of proteins into highly ordered, symmetric supramolecular complexes is an elegant and powerful means of patterning matter at the atomic scale. Recent years have seen advances in the development of self-assembling biomaterials, particularly those composed of nucleic acids. DNA has been used to create, for example, nanoscale shapes and patterns, molecular containers, and three-dimensional macroscopic crystals. Methods for designing self-assembling proteins have progressed more slowly, yet the functional and physical properties of proteins make them attractive as building blocks for the development of advanced functional materials.

Summary

In one example, nanostructures are provided comprising:

(a) a plurality of first assemblies, each first assembly comprising a plurality of identical first polypeptides;

(b) a plurality of second assemblies, each second assembly comprising a plurality of identical second polypeptides, wherein the second polypeptide differs from the first polypeptide;

wherein the plurality of first assemblies non-covalently interact with the plurality of second assemblies to form a nanostructure; and

wherein the nanostructure displays multiple copies of one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, on an exterior of the nanostructure.

In one aspect, the present disclosure refers to a nanostructure, comprising:

(a) a plurality of first assemblies, each first assembly comprising a plurality of identical first polypeptides, wherein the first polypeptides comprise an amino acid sequence having at least 75% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51;

(b) a plurality of second assemblies, each second assembly comprising a plurality of identical second polypeptides, wherein the second polypeptides comprise an amino acid sequence having at least 75% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51, wherein the second polypeptide differs from the first polypeptide;

wherein the plurality of first assemblies non-covalently interact with the plurality of second assemblies to form a nanostructure; and

wherein the nanostructure displays multiple copies of one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, on an exterior of the nanostructure.

In one example, (a) the first polypeptides comprise a polypeptide having at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or

100% identity along its full length to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51; and

(b) the second polypeptides comprise a polypeptide having at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51.

In another example, the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, comprise a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS: 53, 61-68, and 101.

In various examples:

(a) the first polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of T33-31A (SEQ ID NO:51) and the second polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of T33-09B/T33-31B (SEQ ID NO:44);

(b) the first polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of T33-15B (SEQ ID NO:46) and the second polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of T33-15A (SEQ ID NO:45);

(c) the first polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of a polypeptide selected from the group consisting of I53-50A (SEQ ID NO:7), I53-50A.1 (SEQ ID NO:29), I53-50A.1NegT2 (SEQ ID NO:30), and I53-50A.1PosT1 (SEQ ID NO:31), and the second polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of a polypeptide selected from the group consisting of I53-50B (SEQ ID NO:8), I53-50B.1 (SEQ ID NO:32), I53-50B.1NegT2 (SEQ ID NO:33), and I53-50B.4PosT1 (SEQ ID NO:34); or

(d) the first polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of I32-28A (SEQ ID NO:21) and the second polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of I32-28B (SEQ ID NO:22).

In one example, the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, are expressed as a fusion protein with the first polypeptides. In one such example, the plurality of first assemblies each comprise identical first polypeptides; in another such example, the plurality of first assemblies in total comprise two or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof. In another example, only a subset of the first polypeptides comprise a fusion protein with an F protein or antigenic fragment thereof. In a further example, each first assembly comprises a homotrimer of the first polypeptide.

In another example, the fusion protein comprises an amino acid linker positioned between the first polypeptide and the paramyxovirus and/or pneumovirus F proteins, or antigenic fragment thereof. In one such example, the fusion protein comprises an amino acid linker positioned between the first polypeptide and the paramyxovirus F proteins, or antigenic fragment thereof.

In one example the amino acid linker sequence comprises one or more trimerization domain; in another example the amino acid linker sequence comprises a Gly-Ser linker.

In various examples, the first polypeptides comprise or consist of first polypeptides having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along their full length to the amino acid sequence of a polypeptide selected from the group consisting of DS-Cav1-foldon-T33-31A (SEQ ID NO:69), DS-Cav1-T33-31A (SEQ ID NO:70), DS-Cav1-foldon-T33-15B (SEQ ID NO:71), DS-Cav1-T33-15B (SEQ ID NO:72), DS-Cav1-foldon-I53-50A (SEQ ID NO:73), DS-Cav1-I53-50A (SEQ ID NO:74), and DS-Cav1-I32-28A (SEQ ID NO:75). In other examples,

(a) when each first polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of DS-Cav1-foldon-T33-31A (SEQ ID NO:69) or DS-Cav1-T33-31A (SEQ ID NO:70), each second polypeptide has at least 75%, 80%, 85%, 90%, 91%,

92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of T33-31B (SEQ ID NO:44);

(b) when each first polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of DS-Cav1-foldon-T33-15B (SEQ ID NO:71) or DS-Cav1-T33-15B (SEQ ID NO:72), each second polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of T33-15A (SEQ ID NO:45);

(c) when each first polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of DS-Cav1-foldon-I53-50A (SEQ ID NO:73) or DS-Cav1-I53-50A (SEQ ID NO:74), each second polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of a polypeptide selected from the group consisting ofI53-50B (SEQ ID NO:8), I53-50B.1 (SEQ ID NO:32), I53-50B.1NegT2 (SEQ ID NO:33), or I53-50B.4PosT1 (SEQ ID NO:34); or

(d) when each first polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of DS-Cav1-I32-28A (SEQ ID NO:75), each second polypeptide has at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of I32-28B.

In one example, the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of DS-Cav1(SEQ ID NO:53). In one such example, each first polypeptide comprises a fusion polypeptide of a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of DS-Cav1 linked via an amino acid linker to a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of SEQ ID NO:7 (I53-50A). In another example, the amino acid linker comprises a Gly-Ser linker. In a further example, each fusion protein comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid

sequence selected from the group consisting of SEQ ID NOS:69-100. In a specific example, each fusion protein comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of SEQ ID NO:76 (F10). In other examples, each second polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence selected from the group consisting of I53-50B (SEQ ID NO:8), I53-50B.1 (SEQ ID NO:32), I53-50B.1NegT2 (SEQ ID NO:33), or I53-50B.4PosT1 (SEQ ID NO:34). In a specific example, each second polypeptide comprises a polypeptide having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along its full length to the amino acid sequence of I53-50B.4PosT1 (SEQ ID NO:34).

In other examples, recombinant expression nucleic acids expressing the first polypeptide fusions, recombinant expression vectors comprising the recombinant nucleic acids linked to a promoter, and recombinant host cells comprising the recombinant expression vectors are provided. In one aspect, the present disclosure refers to a recombinant nucleic acid encoding the first polypeptide fusion protein disclosed herein. In one aspect, the present disclosure refers to a recombinant expression vector comprising the recombinant nucleic acid disclosed herein operatively linked to a promoter. In one aspect, the present disclosure refers to a recombinant host cell, comprising one or more recombinant expression vectors capable of expressing the first polypeptides and the second polypeptides disclosed herein, or the vector disclosed herein.

Also provided are immunogenic compositions comprising the nanostructure of any example or combination of examples disclosed herein, and a pharmaceutically acceptable carrier. In one aspect, the present disclosure refers to an immunogenic composition comprising the nanostructure disclosed herein, and a pharmaceutically acceptable carrier. In one example, the immunogenic compositions may further comprise an adjuvant.

In other examples, methods for generating an immune response to RSV F protein in a subject, or for treating or limiting a RSV infection in a subject are provided, comprising administering to the subject in need thereof an effective amount of the nanostructure or immunogenic composition of example or combination of examples disclosed herein to generate the immune response or to treat or prevent RSV infection in the subject.

Also provided are processes assembling the nanostructures of any example or combination of examples disclosed herein, comprising mixing two or more nanostructures components in aqueous conditions to drive spontaneous assembly of the desired nanostructures. In one aspect, the present disclosure refers to a process for assembling the nanostructures disclosed herein in vitro, comprising mixing two or more nanostructure components in aqueous conditions to drive spontaneous assembly of the desired nanostructure. In one aspect, the present disclosure refers to a nanostructure produced by the process disclosed herein.

Brief Description of the Drawings

The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same become better understood by reference to

We claim:

1. A nanostructure, comprising:

(a) a plurality of first assemblies, each first assembly comprising a plurality of identical first polypeptides, wherein the first polypeptides comprise an amino acid sequence having at least 75% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51;

(b) a plurality of second assemblies, each second assembly comprising a plurality of identical second polypeptides, wherein the second polypeptides comprise an amino acid sequence having at least 75% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51, wherein the second polypeptide differs from the first polypeptide;

wherein the plurality of first assemblies non-covalently interact with the plurality of second assemblies to form a nanostructure; and

wherein the nanostructure displays multiple copies of one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, on an exterior of the nanostructure.

2. The nanostructure of claim 1, wherein:

(a) the first polypeptides comprise an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51; and

(b) the second polypeptides comprise an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:1-51.

3. The nanostructure of claim 1 or 2, wherein each of the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, comprises

(a) an amino acid sequence having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence of a polypeptide selected from the group consisting of SEQ ID NOS:53, 61-68, and 101;

(b) an amino acid sequence having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence of

an RSV F protein or mutant thereof selected from the group consisting of SEQ ID NO:53 and 61-64, wherein the protein includes one or more of the following residues: 67I, 149C, 458C, 46G, 465Q, 215P, 92D, and 487Q; or

(c) an amino acid sequence having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence of an MPV F protein or mutant thereof selected from the group consisting of SEQ ID NO:65-68 and 101, wherein the protein includes one or more of the following residues: 113C, 120C, 339C, 160F, 177L, 185P, and 426C.

4. The nanostructure of any one of claims 1-3, wherein the first polypeptides and the second polypeptides comprise an amino acid sequence having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequences selected from the following pairs:

SEQ ID NO:1 and SEQ ID NO:2 (I53-34A and I53-34B);
SEQ ID NO:3 and SEQ ID NO:4 (I53-40A and I53-40B);
SEQ ID NO:3 and SEQ ID NO:24 (I53-40A and I53-40B.1);
SEQ ID NO:23 and SEQ ID NO:4 (I53-40A.1 and I53-40B);
SEQ ID NO:35 and SEQ ID NO:36 (I53-40A genus and I53-40B genus);
SEQ ID NO:5 and SEQ ID NO:6 (I53-47A and I53-47B);
SEQ ID NO:5 and SEQ ID NO:27 (I53-47A and I53-47B.1);
SEQ ID NO:5 and SEQ ID NO:28 (I53-47A and I53-47B.1NegT2);
SEQ ID NO:25 and SEQ ID NO:6 (I53-47A.1 and I53-47B);
SEQ ID NO:25 and SEQ ID NO:27 (I53-47A.1 and I53-47B.1);
SEQ ID NO:25 and SEQ ID NO:28 (I53-47A.1 and I53-47B.1NegT2);
SEQ ID NO:26 and SEQ ID NO:6 (I53-47A.1NegT2 and I53-47B);
SEQ ID NO:26 and SEQ ID NO:27 (I53-47A.1NegT2 and I53-47B.1);
SEQ ID NO:26 and SEQ ID NO:28 (I53-47A.1NegT2 and I53-47B.1NegT2);
SEQ ID NO:37 and SEQ ID NO:38 (I53-47A genus and I53-47B genus);
SEQ ID NO:7 and SEQ ID NO:8 (I53-50A and I53-50B);
SEQ ID NO:7 and SEQ ID NO:32 (I53-50A and I53-50B.1);
SEQ ID NO:7 and SEQ ID NO:33 (I53-50A and I53-50B.1NegT2);
SEQ ID NO:7 and SEQ ID NO:34 (I53-50A and I53-50B.4PosT1);
SEQ ID NO:29 and SEQ ID NO:8 (I53-50A.1 and I53-50B);
SEQ ID NO:29 and SEQ ID NO:32 (I53-50A.1 and I53-50B.1);

SEQ ID NO:29 and SEQ ID NO:33 (I53-50A.1 and I53-50B.1NegT2);
SEQ ID NO:29 and SEQ ID NO:34 (I53-50A.1 and I53-50B.4PosT1);
SEQ ID NO:30 and SEQ ID NO:8 (I53-50A.1NegT2 and I53-50B);
SEQ ID NO:30 and SEQ ID NO:32 (I53-50A.1NegT2 and I53-50B.1);
SEQ ID NO:30 and SEQ ID NO:33 (I53-50A.1NegT2 and I53-50B.1NegT2);
SEQ ID NO:30 and SEQ ID NO:34 (I53-50A.1NegT2 and I53-50B.4PosT1);
SEQ ID NO:31 and SEQ ID NO:8 (I53-50A.1PosT1 and I53-50B);
SEQ ID NO:31 and SEQ ID NO:32 (I53-50A.1PosT1 and I53-50B.1);
SEQ ID NO:31 and SEQ ID NO:33 (I53-50A.1PosT1 and I53-50B.1NegT2);
SEQ ID NO:31 and SEQ ID NO:34 (I53-50A.1PosT1 and I53-50B.4PosT1);
SEQ ID NO:39 and SEQ ID NO:40 (I53-50A genus and I53-50B genus);
SEQ ID NO:9 and SEQ ID NO:10 (I53-51A and I53-51B);
SEQ ID NO:11 and SEQ ID NO:12 (I52-03A and I52-03B);
SEQ ID NO:13 and SEQ ID NO:14 (I52-32A and I52-32B);
SEQ ID NO:15 and SEQ ID NO:16 (I52-33A and I52-33B)
SEQ ID NO:17 and SEQ ID NO:18 (I32-06A and I32-06B);
SEQ ID NO:19 and SEQ ID NO:20 (I32-19A and I32-19B);
SEQ ID NO:21 and SEQ ID NO:22 (I32-28A and I32-28B);
SEQ ID NO:23 and SEQ ID NO:24 (I53-40A.1 and I53-40B.1);
SEQ ID NO:41 and SEQ ID NO:42 (T32-28A and T32-28B);
SEQ ID NO:43 and SEQ ID NO:44 (T33-09A and T33-09B);
SEQ ID NO:45 and SEQ ID NO:46 (T33-15A and T33-15B);
SEQ ID NO:47 and SEQ ID NO:48 (T33-21A and T33-21B);
SEQ ID NO:49 and SEQ ID NO:50 (T33-28A and T32-28B); and
SEQ ID NO:51 and SEQ ID NO:44 (T33-31B and T33-09B (also referred to as T33-31B)).

5. The nanostructure of any one of claims 1-4, wherein the first polypeptides comprise polypeptides having at least 95%, 96%, 97%, 98%, 99%, or 100% identity along their full length to the amino acid sequence of I53-50A (SEQ ID NO: 7) and the second polypeptides comprise polypeptides having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity along their full length to the amino acid sequence of I53-50B (SEQ ID NO: 34).

6. The nanostructure of any one of claims 1-5, wherein the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof, are expressed as a fusion protein with the first polypeptides.

7. The nanostructure of claim 6, wherein
(a) the plurality of first assemblies each comprise identical fusion proteins; or
(b) the plurality of first assemblies in total comprise two or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof.

8. The nanostructure of claim 7, wherein only a subset of the first polypeptides comprise a fusion protein with an F protein or antigenic fragment thereof.

9. The nanostructure of any one of claims 1-8, wherein each first assembly comprises a homotrimer of the first polypeptide.

10. The nanostructure of any one of claims 1-9, wherein the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof comprises an amino acid sequence having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence of DS-Cav1 (SEQ ID NO:53).

11. The nanostructure of any one of claims 1-10, wherein the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragments thereof comprises an amino acid sequence having at least 95% identity to the amino acid sequence of DS-Cav1 (SEQ ID NO:53).

12. The nanostructure of any one of claims 5-11, wherein each fusion protein comprises an amino acid linker positioned between the first polypeptide and the one or more paramyxovirus and/or pneumovirus F proteins, or antigenic fragment thereof.

13. The nanostructure of claim 12, wherein the amino acid linker sequence comprises one or more trimerization domain.

14. The nanostructure of claim 12 or 13, wherein the amino acid linker sequence (a) comprises the amino acid sequence GYIPEAPRDGQAYVRKDGEWVLLSTFL (SEQ ID NO:54) or (b) comprises a Gly-Ser linker.

15. The nanostructure of any one of claims 6-14, wherein the fusion protein comprises an amino acid sequence having at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the amino acid sequence selected from the group consisting of SEQ ID NOS:69-100.

16. The nanostructure of claim 11, wherein each first polypeptide comprises a fusion protein of DS-Cav1 (SEQ ID NO:53) linked to SEQ ID NO:7 (I53-50A) via an amino acid linker.

17. The nanostructure of claim 16, wherein the amino acid linker comprises a Gly-Ser linker and/or a helical extension domain.

18. A recombinant nucleic acid encoding the first polypeptide fusion protein as recited in any one of claims 6-17.

19. A recombinant expression vector comprising the recombinant nucleic acid of claim 18 operatively linked to a promoter.

20. A recombinant host cell, comprising one or more recombinant expression vectors capable of expressing the first polypeptides and the second polypeptides of any one of claims 1-17, or the vector of claim 19.

21. An immunogenic composition comprising the nanostructure of any one of claims 1-17, and a pharmaceutically acceptable carrier.

22. The immunogenic composition of claim 21, further comprising an adjuvant.

23. A process for assembling the nanostructures of any one of claims 1-17 in vitro, comprising mixing two or more nanostructure components in aqueous conditions to drive spontaneous assembly of the desired nanostructure.

24. A nanostructure produced by the process of claim 23.