Title
Modular Nucleic Acid Scaffolds for Synthesizing Monodisperse and Sequence-Encoded Antibody Oligomers.

Journal
Chem, 8(11)

ISSN
2451-9294

Authors
Winegar, Peter
Figg, C
Teplensky, Michelle
et al.

Publication Date
2022-11-10

DOI
10.1016/j.chempr.2022.07.003

Peer reviewed

HHS Public Access

Author manuscript

Chem. Author manuscript; available in PMC 2023 November 10.

Published in final edited form as:

Chem. 2022 November 10; 8(11): 3018–3030. doi:10.1016/j.chempr.2022.07.003.

Modular Nucleic Acid Scaffolds for Synthesizing Monodisperse and Sequence-Encoded Antibody Oligomers

Peter H. Winegar[1,2,4], **C. Adrian Figg**[1,2,4], **Michelle H. Teplensky**[1,2], **Namrata Ramani**[2,3], **Chad A. Mirkin**[1,2,3,5,*]

[1]Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

[2]International Institute for Nanotechnology, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

[3]Department of Materials Science and Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

[4]These authors contributed equally

[5]Lead contact

SUMMARY

Synthesizing protein oligomers that contain exact numbers of multiple different proteins in defined architectures is challenging. DNA–DNA interactions can be used to program protein assembly into oligomers; however, existing methods require changes to DNA design to achieve different numbers and oligomeric sequences of proteins. Herein, we develop a modular DNA scaffold that uses only six synthetic oligonucleotides to organize proteins into defined oligomers. As a proof-of-concept, model proteins (antibodies) are oligomerized into dimers and trimers, where antibody function is retained. Illustrating the modularity of this technique, dimer and trimer building blocks are then assembled into pentamers containing three different antibodies in an exact stoichiometry and oligomeric sequence. In sum, this report describes a generalizable method for organizing proteins into monodisperse, sequence-encoded oligomers using DNA. This advance will enable studies into how oligomeric protein sequences affect material properties in areas

*Correspondence: chadnano@northwestern.edu.

AUTHOR CONTRIBUTIONS

P.H.W. and C.A.F. contributed equally.

Conceptualization, P.H.W., C.A.F, C.A.M.; Methodology, P.H.W., C.A.F., M.H.T., N.R., C.A.M.; Investigation, P.H.W., C.A.F., M.H.T., N.R., C.A.M.; Writing – Original Draft, P.H.W., C.A.F., C.A.M.; Writing – Review and Editing, P.H.W., C.A.F., M.H.T., N.R., C.A.M.; Funding Acquisition, C.A.M.

SUPPLEMENTAL INFORMATION

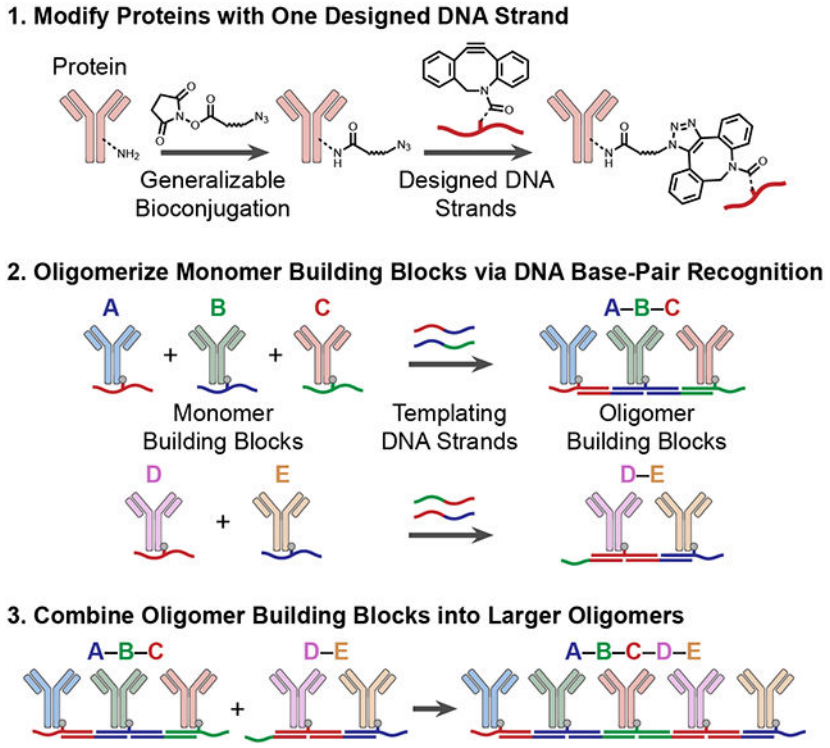
This section should include the titles and (optional) legends of all supplemental items. Document S1 is the main supplemental PDF: Document S1. Supplemental information regarding experimental procedures, results, and discussion, Tables S1-S3, Schemes S1-S5, and Figures S1-S13.

DECLARATION OF INTERESTS

"Synthetic Strategy to Polymerize Protein Into Molecularly Defined Polymers" Mirkin, C. A.; Winegar, P. H.; Figg, C. A., PCT/US2022/021359. March 22, 2022.

spanning pharmaceutical development, cascade catalysis, synthetic photosynthesis, and membrane transport.

Graphical Abstract



eTOC

In nature, many proteins organize into architectures where the exact number and spatial arrangement of each protein can dictate biological function, including catalysis and photosynthesis. To mimic and even surpass such function, the preparation of many different protein architectures is required. However, this is synthetically challenging. Herein, we developed a modular method that uses designed nucleic acid sequences to organize proteins into different monodisperse and sequence-encoded oligomers, including dimers, trimers, and pentamers.

Keywords

biomaterial; protein; antibody; DNA; DNA nanotechnology; polymer; oligomer; sequence-encoded; self-assembly; protein assembly

INTRODUCTION

In Nature, many proteins assemble into defined oligomeric architectures that contain exact numbers and oligomeric sequences of multiple different proteins.[1, 2, 3] Herein, oligomeric sequence of proteins and oligomeric protein sequence are defined as the order of proteins within an oligomeric architecture. This assembly can dictate the biological (*e.g.*, human

IgM antibodies contain five protein subunits),[4] catalytic (*e.g.*, eukaryotic RNA polymerase II contains twelve protein subunits),[5] photophysical (*e.g.*, cyanobacteria photosystem I contains twelve protein subunits),[6] and membrane transport (*e.g.*, *Streptomyces lividans* potassium channel contains four protein subunits)[7] properties of proteins. To mimic and potentially surpass these properties, the modular synthesis of different protein oligomers is needed. A versatile synthetic protein oligomerization method would: (1) provide access to a large number of proteins per oligomer, (2) provide access to any oligomeric sequence of the same or different proteins, (3) be generalizable with regard to proteins, and (4) not require mutations of the amino acid sequence of proteins and recombinant protein expression. A method that meets all four criteria would enable the study of how the identity, number, stoichiometry, oligomeric sequence, and architecture of proteins affects the emergent properties of oligomers. While strategies have been developed to prepare synthetic protein oligomers[8, 9, 10, 11] and study how oligomerization affects protein properties,[12, 13, 14, 15] there is no current method that meets all four listed criteria (Figures 1A and 1B). In this work, we explored the design and synthesis of a single modular nucleic acid scaffold that can be used to organize proteins into a near limitless array of monodisperse and sequence-encoded protein oligomers (Figure 1C).

Protein oligomers are frequently prepared using techniques from molecular biology, including genetic engineering and recombinant expression of mutated proteins (Figure 1A). Generally, fusion proteins are designed via genetic engineering and recombinantly expressed to achieve the desired oligomerization structure. Three common methods for oligomerization using fusion proteins include: direct expression of protein oligomers as a single polypeptide[16, 17, 18]; fusion of proteins to a subunit of a multimeric protein (*e.g.*, streptavidin) that assembles into supramolecular constructs (Figure 1A, i)[19, 20, 21, 22]; or fusion of proteins to a subunit that can selectively form covalent bonds with a complementary group (*e.g.*, SpyTag/SpyCatcher,[23] SnoopTag/Snoopcatcher,[24] SNAPtag,[25] HALOtag,[26] or cutinase,[27] Figure 1A, ii).[24, 28, 29, 30, 31] Furthermore, recent advances in the design of protein–protein interfaces enable sophisticated control over synthetic protein architectures (Figure 1A, iii).[32, 33, 34, 35, 36, 37] Each of these methods requires mutations of the amino acid sequence of proteins and recombinant protein expression. However, many proteins are challenging to prepare via recombinant expression (*e.g.*, proteins with post-translational modifications,[38] proteins with disulfide bonds,[39] toxic proteins,[40] or proteins that aggregate[41]), potentially limiting the scope of proteins that can be oligomerized through these methods.

Chemical approaches to assemble proteins are another powerful method to control oligomerization (Figure 1B). The amino acid sequence of proteins can be mutated to incorporate (un)natural amino acids at defined positions for interactions such as electrostatic,[42] supramolecular host–guest binding,[43, 44, 45, 46] metal coordination,[47, 48, 49] or covalent linking.[50, 51] However, without extensive chemical design, modification of protein amino acid sequences and/or recombinant protein expression, it is challenging to access monodisperse and sequence-encoded oligomers that are larger than dimers or trimers. Chemical approaches to directly oligomerize proteins post-expression without modifying the amino acid sequence of proteins or using an assembly template (*vide infra*) were expanded with the introduction of bioorthogonal “click” reactions (Figure 1B, i).[52, 53, 54] While these

reactions have been used to oligomerize therapeutically relevant proteins (*e.g.*, antibodies), it is challenging to achieve oligomers larger than dimers or trimers.[55]

Proteins can be oligomerized using attachment to chemical scaffolds (*e.g.*, polymers[56, 57, 58, 59, 60] or DNA[61, 62, 63, 64, 65], Figure 1B, ii). Templated assembly of proteins using DNA is one of the most promising and versatile approaches to organize proteins into oligomers larger than dimers or trimers. The utility of this approach is a result of the programmability of nucleic acids where specific, defined assemblies can be accessed solely based on DNA sequence design.[61, 62, 63, 64, 65] For example, proteins have been covalently[66, 67, 68] or noncovalently[67, 69] modified with oligonucleotides and the resulting constructs have been organized via DNA–DNA interactions into a multitude of protein oligomers with one-,[70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81] two-,[82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104] and three-dimensional[105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119] architectures. However, in each of these systems, DNA design must be changed to synthesize defined protein oligomers that contain different numbers or oligomeric sequences of proteins. For example, a protein tetramer can be readily synthesized using a DNA tetrahedron scaffold[82] or a four-arm Holliday DNA junction scaffold.[83, 84, 85] However, to synthesize a protein pentamer, every DNA sequence must be redesigned to account for the one additional protein. Modular multi-protein constructs can be realized on large DNA origami structures,[86, 87, 105] but most of the composition (*e.g.*, >80%) of these constructs is DNA instead of protein. Molecular constructs that minimize the required amount of DNA such that most of the chemical properties are dictated by the identity and organization of proteins are inaccessible using these techniques.[62] Together, these limitations significantly hamper any studies where access to libraries of different protein oligomers with discrete stoichiometries and oligomeric sequences of proteins could provide insight into how protein–protein interactions and cooperativity can be exploited for enhanced properties of oligomers.

We hypothesized that a single set of designed DNA strands could be used as a modular scaffold to organize proteins into oligomers with exact stoichiometries and oligomeric sequences (Figure 1C). This DNA design would enable different proteins to be precisely organized into an expansive array of monodisperse, sequence-encoded oligomers (Figure 1C, i). Herein, we tested our hypothesis by designing a modular six-strand DNA scaffold and using it to oligomerize commercially available and therapeutically relevant proteins (*i.e.*, antibodies). The scaffold consists of three distinct DNA strands that can be conjugated to proteins and three distinct DNA strands that template the assembly of DNA-modified proteins into oligomers via DNA–DNA interactions (Scheme S1). Importantly, each of the six DNA strands contains two distinct binding domains and the sites for attachment to proteins can be located anywhere on the DNA strands (Scheme S3). Using the designed DNA scaffold, monodisperse, sequence-encoded monomer, dimer, and trimer building blocks are synthesized. Next, these building blocks are used to access a larger oligomer (*i.e.*, pentamer) that contains a defined number and oligomeric sequence of proteins. Importantly, the foundational examples shown herein are a fraction of the possible oligomeric sequences that are accessible using the modular six-strand DNA scaffold (Figure 1C, ii, Table S3). For example, if five different proteins are used, there are, in principle, 3,125 different accessible pentameric sequences. Overall, this generalizable synthetic route

will enable future investigations into how the identity, stoichiometry, oligomeric sequence, and architecture of proteins in oligomers affect the properties of these constructs.

RESULTS

Three commercially available IgG antibodies commonly used as checkpoint inhibitors (*i.e.*, **anti-mouse-PD-1** (**A**), **anti-mouse-TIGIT** (**B**), and **anti-mouse-CTLA-4** (**C**)) were chosen for the sequence-encoded oligomerization of proteins using a modular DNA scaffold. To install a single DNA strand onto either **A**, **B**, or **C**, each antibody was reacted with 2 equiv. of an oligo(ethylene glycol) molecule containing an *N*-hydroxysuccinimide activated ester and an azide (NHS–PEG_{12}–N_3) for 45 min (Figure 2A). This chemistry targets the primary amines (*e.g.*, ε-amines on lysines and α-amines on N-termini)[120] on both the Fc and Fab regions of the antibody and was chosen because it is generalizable with regard to proteins. Although the exact location of conjugation cannot be controlled, we expected that the number of azide modifications per antibody would be controlled by low numbers of equivalents of NHS–PEG_{12}–N_3 added. We hypothesized that the low number of primary amines modified would not inhibit the target binding characteristics of antibodies. After purification by size-exclusion chromatography (SEC), the azide on the surface of each antibody underwent a strain-promoted azide-alkyne cycloaddition reaction with 5 equiv. of DNA strands containing dibenzocyclooctyne (DBCO) and a fluorophore (*i.e.*, Cyanine 3 (Cy3), Cyanine 5 (Cy5), or fluorescein (FITC)) and two distinct 20 base nucleic acid sequences (Table S1, Schemes S1A and S2, *i.e.*, S2–DBCO–Cy3–S3, S4–DBCO–Cy5–S5, or S6–DBCO–FITC–S1). After 16 h, roughly 25–30% of antibodies were modified with one DNA strand (Figure S1). Next, unreacted DNA was removed from the reaction mixture using SEC. Anion exchange chromatography was used to isolate antibodies that were functionalized with a single DNA strand from unreacted antibodies and antibodies that were functionalized with multiple DNA strands (Figure 2A). Three different protein–DNA conjugates (*i.e.*, S2–**A**–Cy3–S3, S4–**B**–Cy5–S5, and S6–**C**–FITC–S1) were prepared and confirmed to contain a single DNA functionalization via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Figure 2B) and SEC (Figure S2).

Protein oligomers were synthesized by mixing the purified protein–DNA conjugates (Figure 3A, 3E: lanes 1–3) with template DNA strands (Table S1, Scheme S1B, *i.e.*, S1'–S2', S3'–S4', or S5'–S6'). The template strands were designed as complements to two 20 base nucleic acid sequences on different antibody–DNA conjugates (Table S2, Scheme S1C). For example, the S5' DNA sequence on the template strand S5'–S6' is complementary to the S5 DNA sequence on S4–**B**–Cy5–S5 and the S6' DNA sequence is complementary to the S6 DNA sequence on S6–**C**–FITC–S1. Equal amounts of the **B**–DNA conjugate, **C**–DNA conjugate, S5'–S6' template strand, and S1'–S2' template strand were mixed to synthesize a protein dimer with the oligomeric sequence S4–**B**–**C**–S2' (Figure 3B, Scheme S5A) at an assembly yield of 68% (Figure S5A, D). Oligomers that contain greater than two antibodies were not observed in the assembly mixture because there are no DNA sequences that are complementary to either the S4 or S2' DNA sequences. Protein dimers were isolated from unreacted monomers and template strands in the assembly mixture using SEC purification and characterized with agarose gel electrophoresis (Figure 3E: lane 4). Importantly, the agarose gel showed a single band for dimers with only the expected

Cy5 and FITC fluorescence and lower electrophoretic mobility than either antibody–DNA conjugate alone. Therefore, monodisperse and sequence-encoded protein dimers with the oligomeric sequence S4–**B**–**C**–S2' were successfully synthesized.

Next, the synthesis of a sequence-encoded protein trimer was targeted. Equal amounts of the **A**–DNA conjugate, **B**–DNA conjugate, **C**–DNA conjugate, S3'–S4' template strand, and S5'–S6' template strand were mixed to synthesize a protein trimer with the oligomeric sequence S2–**A**–**B**–**C**–S1 (Figure 3C, Scheme S5B) at an assembly yield of 27% (Figure S5B, D). Oligomers that contain greater than three antibodies or trimers containing alternative oligomeric antibody sequences were not observed in the assembly mixture because there are no DNA sequences that are complementary to either the S2 or S1 DNA sequences. Protein trimers were isolated from the assembly mixture using SEC purification and characterized with agarose gel electrophoresis (Figure 3E: lane 5). The agarose gel showed a single band for the trimers with the expected Cy3, Cy5, and FITC dye fluorescence as well as lower electrophoretic mobility on an agarose gel than the dimers. Therefore, these results indicate that monodisperse and sequence-encoded protein trimers with the oligomeric sequence S2–**A**–**B**–**C**–S1 were successfully synthesized. Importantly, no disassembly of S4–**B**–**C**–S2' or S2–**A**–**B**–**C**–S1 oligomers were observed over 10 days of storage at 4 °C.

To ensure that this synthetic technique is generalizable, different protein oligomers were synthesized, including a protein dimer with the oligomeric sequence **A**–**B** (Figure S13) and a protein trimer with the oligomeric sequence, **A**–**B**–**B** (Figure S6). Furthermore, another antibody, **anti-human-PD-1** (**D**), was functionalized with a single DNA strand of S2–DBCO–Cy3–S3, S4–DBCO–Cy5–S5, or S6–DBCO–FITC–S1 (Figure S3 and S4), and the resulting constructs were organized into protein dimers with the oligomeric sequence **D**–**D** (Figures S7, S8, and S10) and trimers with the oligomeric sequence **D**–**D**–**D** (Figure S7). Analytical SEC analysis of antibody–DNA conjugates, dimers, and trimers show a single peak for each sample with decreases in retention time as degree of oligomerization increases (Figures S6C and S7C).

The target binding characteristics of human antibodies after functionalization with DNA and oligomerization with the modular DNA scaffold were investigated using antigen binding and checkpoint inhibitor activity cellular assays. We studied **D**, **D**–DNA conjugates, **D**–**D** dimers, and **D**–**D**–**D** trimers using these assays and confirmed that antigen binding and checkpoint inhibitor activity were retained in each sample (Figures S11 and S12, see the supplemental information (SI) for additional details). Importantly, an antibody dimer, **A**–**B**, exhibited minimal degradation in the cellular media used in these experiments (Figure S13, see the SI for additional details).

Finally, a protein dimer and trimer were used as building blocks to synthesize a monodisperse and sequence-encoded protein pentamer where three different antibodies are organized into a precise oligomeric sequence. The S4–**B**–**C**–S2' protein dimer and S2–**A**–**B**–**C**–S1 protein trimer were mixed together at a 1:1 ratio and the specific binding between the S2' DNA sequence on the dimer and the S2 DNA sequence on the trimer leads to the synthesis of a sequence-encoded protein pentamer with the oligomeric sequence S4–**B**–**C**–**A**–**B**–**C**–S1 (Figure 3D, Scheme S5C) at an assembly yield of 58% (Figure

S5C, D). Oligomers that contain greater than five antibodies were not observed in the assembly mixture. Protein pentamers were isolated from other unreacted dimers, trimers, and template strands in the assembly mixture using SEC purification and characterized with agarose gel electrophoresis (Figure 3E: lane 6). The pentamers showed a single band with the expected Cy3, Cy5, and FITC dye fluorescence as well as lower electrophoretic mobility on an agarose gel than the trimers. Therefore, monodisperse and sequence-encoded protein pentamers with the oligomeric sequence S4–**B**–**C**–**A**–**B**–**C**–S1 were successfully synthesized. This is the first reported monodisperse antibody pentamer that contains different antibodies in a predefined oligomeric sequence.

DISCUSSION

The designed set of six DNA strands was used as a modular scaffold to organize proteins into oligomers with programmed identity, stoichiometry, and oligomeric sequence. This scaffold provided access to monomer, dimer, and trimer building blocks that could be modularly combined independent from the identity of proteins. Therefore, all of the criteria for a versatile synthetic protein oligomerization method were met: (1) providing access to a large number of proteins per oligomer, (2) providing access to any oligomeric sequence of the same or different proteins, (3) being generalizable with regard to proteins, and (4) not requiring mutations of the amino acid sequence of proteins and recombinant protein expression.

Established chemistry was used to functionalize a primary amine (*e.g.*, ε-amine on a lysine or α-amine on a N-terminus) on proteins with a single DNA strand.[120] Nearly all proteins contain primary amines, so this approach is generalizable with regard to proteins, including proteins that are commercially available, isolated from natural sources, or recombinantly expressed. Many other covalent[66, 67] and noncovalent[67, 69] bioconjugation methods could also be used to modify proteins with one of the DNA strands reported here. By taking advantage of these approaches, nearly any protein can be modified with a single DNA strand and organized into monodisperse, sequence-encoded oligomers using the designed DNA scaffold.

While modular multi-protein constructs can be prepared using large DNA origami scaffolds, proteins comprise less than 20% of the mass of these constructs.[86, 87, 105] This large amount of DNA compared to protein means that most of the solution properties and interactions of these constructs are dictated by the DNA scaffold instead of by the proteins. In cases where protein binding interactions are integral to function (*e.g.*, antibody–antigen binding), this large amount of DNA may affect target recognition and accessibility. In contrast, using the modular DNA scaffold reported in this work, proteins make up most of the mass of oligomeric constructs. For example, proteins comprise 84%, 86%, and 85% of the mass of the sequence-encoded **B**–**C** dimer, **A**–**B**–**C** trimer, and **B**–**C**–**A**–**B**–**C** pentamer, respectively. While other protein assembly techniques using nucleic acids also provide access to oligomers mostly comprised of proteins (*e.g.*, oligomerization using a DNA tetrahedron scaffold[82] or a four-arm Holliday DNA junction scaffold[83, 84, 85]), they lack modularity to access different numbers of proteins per construct.

In principle, the modular DNA scaffold described herein provides access to vast numbers of different oligomeric sequences and sizes. For example, 3 different proteins could be oligomerized into trimeric constructs with 27 different oligomeric sequences, including homooligomers of one protein, and hetero-oligomers of 2 or 3 proteins (Figure 1c, ii). Oligonucleotides in the DNA scaffold interact through Watson–Crick–Franklin base pairing to form a right-handed double helix. Therefore, two oligomers with reversed oligomeric sequences of proteins (*e.g.*, A–A–B and B–A–A) form different structures and are considered as different sequences. Likewise, 2 different proteins could be oligomerized into dimeric constructs with 4 different oligomeric sequences, 4 different proteins could be oligomerized into tetrameric constructs with 256 different oligomeric sequences, and 5 different proteins could be oligomerized into pentameric constructs with 3,125 different oligomeric sequences (Table S3). Furthermore, each oligomer building block synthesized using this method inherently contains living chain ends where more units could be added to access larger oligomers (*e.g.*, hexamers, heptamers, and octamers, Scheme S4). Considering the growing number of discovered proteins, the foundational oligomers synthesized in this work illustrate the unlimited number of protein oligomers that could be accessed via a single modular DNA scaffold.

CONCLUSION

In conclusion, this work shows how monodisperse, sequence-encoded protein oligomers can be synthesized using generalizable bioconjugation chemistry with a judiciously designed DNA scaffold. This versatile protein oligomerization approach is powerful and useful because oligomers with different stoichiometries and oligomeric sequences can be synthesized without the need to redesign the proteins or the DNA scaffold. Importantly, this synthetic advance will enable subsequent studies to understand the fundamental relationships between protein oligomer structures and properties, which have significant implications for many fields (*e.g.*, therapeutics, catalysis, photosynthesis, and membrane transport).

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact—Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Chad A. Mirkin (chadnano@northwestern.edu).

Materials availability—All materials generated in this study can be obtained upon reasonable request from the corresponding author with a valid materials transfer agreement and an assumption of costs.

Data and code availability—This study did not generate any datasets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This material is based upon work supported by the following awards: National Science Foundation DMR-2104353, Air Force Office of Scientific Research FA9550-16-1-0150, and National Cancer Institute of the National Institutes of Health U54CA199091. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. M.H.T. gratefully acknowledges support from Northwestern University's Cancer Nanotechnology Training Program Award T32CA186897. This work made use of the IMSERC MS facility at Northwestern University, which has received support from the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-2025633), the State of Illinois, and the International Institute for Nanotechnology (IIN).

REFERENCES*

1. Ahnert SE, Marsh JA, Hernández H, Robinson CV, and Teichmann SA (2015). Principles of Assembly Reveal a Periodic Table of Protein Complexes. Science 350, aaa2245. 10.1126/science.aaa2245. [PubMed: 26659058]
2. Marsh JA, and Teichmann SA Structure, Dynamics, Assembly, and Evolution of Protein Complexes. Annu. Rev. Biochem 84, 551–575. 10.1146/annurev-biochem-060614-034142.
3. Pieters BJGE, van Eldijk MB, Nolte RJM, and Mecinović J (2016). Natural Supramolecular Protein Assemblies. Chem. Soc. Rev 45, 24–39. 10.1039/C5CS00157A. [PubMed: 26497225]
4. Perkins SJ, Nealis AS, Sutton BJ, and Feinstein A (1991). Solution Structure of Human and Mouse Immunoglobulin M by Synchrotron X-ray Scattering and Molecular Graphics Modelling: A Possible Mechanism for Complement Activation. J. Mol. Biol 221, 1345–1366. 10.1016/0022-2836(91)90937-2. [PubMed: 1942055]
5. Bushnell DA, and Kornberg RD (2003). Complete, 12-Subunit RNA Polymerase II at 4.1-Å Resolution: Implications for the Initiation of Transcription. Proc. Natl. Acad. Sci. U. S. A 100, 6969–6973. 10.1073/pnas.1130601100. [PubMed: 12746498]
6. Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, and Krauß N (2001). Three-Dimensional Structure of Cyanobacterial Photosystem I at 2.5 Å Resolution. Nature 411, 909–917. 10.1038/35082000. [PubMed: 11418848]
7. Doyle DA, Cabral JM, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT, and MacKinnon R (1998). The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. Science 280, 69–77. 10.1126/science.280.5360.69. [PubMed: 9525859]
8. Zhu J, Avakyan N, Kakkis A, Hoffnagle AM, Han K, Li Y, Zhang Z, Choi TS, Na Y, Yu C-J, et al. (2021). Protein Assembly by Design. Chem. Rev 121, 13701–13796. 10.1021/acs.chemrev.1c00308. [PubMed: 34405992]
9. Luo Q, Hou C, Bai Y, Wang R, and Liu J (2016). Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. Chem. Rev 116, 13571–13632. 10.1021/acs.chemrev.6b00228. [PubMed: 27587089]
10. King NP, and Lai Y-T (2013). Practical Approaches to Designing Novel Protein Assemblies. Curr. Opin. Struct. Biol 23, 632–638. 10.1016/j.sbi.2013.06.002. [PubMed: 23827813]
11. Permana D, Putra HE, and Djaenudin D (2022). Designed Protein Multimerization and Polymerization for Functionalization of Proteins. Biotechnol. Lett 44, 341–365. 10.1007/s10529-021-03217-8. [PubMed: 35083582]
12. Nuñez-Prado N, Compte M, Harwood S, Álvarez-Méndez A, Lykkemark S, Sanz L, and Álvarez-Vallina L (2015). The Coming of Age of Engineered Multivalent Antibodies. Drug Discov. 20, 588–594. 10.1016/j.drudis.2015.02.013.
13. Zhang G, Quin MB, and Schmidt-Dannert C (2018). Self-Assembling Protein Scaffold System for Easy in Vitro Coimmobilization of Biocatalytic Cascade Enzymes. ACS Catal. 8, 5611–5620. 10.1021/acscatal.8b00986.
14. Li X, Qiao S, Zhao L, Liu S, Li F, Yang F, Luo Q, Hou C, Xu J, and Liu J (2019). Template-Free Construction of Highly Ordered Monolayered Fluorescent Protein Nanosheets: A Bioinspired Artificial Light-Harvesting System. ACS Nano 13, 1861–1869. 10.1021/acsnano.8b08021. [PubMed: 30747517]

15. Aydin S (2015). A Short History, Principles, and Types of ELISA, and Our Laboratory Experience with Peptide/Protein Analyses Using ELISA. Peptides 72, 4–15. 10.1016/j.peptides.2015.04.012. [PubMed: 25908411]

16. Paloni JM, Miller EA, Sikes HD, and Olsen BD (2018). Improved Ordering in Low Molecular Weight Protein–Polymer Conjugates Through Oligomerization of the Protein Block. Biomacromolecules 19, 3814–3824. 10.1021/acs.biomac.8b00928. [PubMed: 30132651]

17. Structural Genomics Consortium, A.e.F.d.M., Biologiques, B.S.G.C., China Structural Genomics, Consortium, I.C.f.S.a.F.I, Israel Structural, Proteomics Center, J.C.f.S.G., Midwest Center for, Structural Genomics, N.Y.S.G.R.C.f.S., Genomics, N.S.G.C., Oxford Protein, Production Facility, P.S.P.F., Max Delbrück Center for, Molecular Medicine, R.S.G.P.I., and SPINE2-, and Complexes (2008). Protein Production and Purification. Nat. Methods 5, 135–146. 10.1038/nmeth.f.202. [PubMed: 18235434]

18. Lapenta F, Aupič J, Strmšek Ž, and Jerala R (2018). Coiled Coil Protein Origami: From Modular Design Principles Towards Biotechnological Applications. Chem. Soc. Rev 47, 3530–3542. 10.1039/C7CS00822H. [PubMed: 29400389]

19. Halin C, Gafner V, Villani ME, Borsi L, Berndt A, Kosmehl H, Zardi L, and Neri D (2003). Synergistic Therapeutic Effects of a Tumor Targeting Antibody Fragment, Fused to Interleukin 12 and to Tumor Necrosis Factor α. Cancer Res. 63, 3202–3210. [PubMed: 12810649]

20. Kipriyanov SM, Breitling F, Little M, and Dübel S (1995). Single-Chain Antibody Streptavidin Fusions: Tetrameric Bifunctional scFv-Complexes with Biotin Binding Activity and Enhanced Affinity to Antigen. Hum. Antibodies 6, 93–101. 10.3233/HAB-1995-6303.

21. Zhang J, Tanha J, Hirama T,Khieu NH, To R, Tong-Sevinc H, Stone E, Brisson J-R, and Roger Mackenzie C (2004). Pentamerization of Single-domain Antibodies from Phage Libraries: A Novel Strategy for the Rapid Generation of High-Avidity Antibody Reagents. J. Mol. Biol 335, 49–56. 10.1016/j.jmb.2003.09.034. [PubMed: 14659739]

22. Kim YE, Kim Y.-n., Kim JA, Kim HM, and Jung Y (2015). Green Fluorescent Protein Nanopolygons as Monodisperse Supramolecular Assemblies of Functional Proteins with Defined Valency. Nat. Commun 6, 7134. 10.1038/ncomms8134. [PubMed: 25972078]

23. Zakeri B, Fierer JO, Celik E, Chittock EC, Schwarz-Linek U, Moy VT, and Howarth M (2012). Peptide Tag Forming a Rapid Covalent Bond to a Protein, through Engineering a Bacterial Adhesin. Proc. Natl. Acad. Sci. U. S. A 109, E690–E697. 10.1073/pnas.1115485109. [PubMed: 22366317]

24. Veggiani G, Nakamura T, Brenner MD, Gayet RV, Yan J, Robinson CV, and Howarth M (2016). Programmable Polyproteams Built Using Twin Peptide Superglues. Proc. Natl. Acad. Sci. U. S. A 113, 1202–1207. 10.1073/pnas.1519214113. [PubMed: 26787909]

25. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, and Johnsson K (2003). A General Method for the Covalent Labeling of Fusion Proteins with Small Molecules In Vivo. Nat. Biotechnol 21, 86–89. 10.1038/nbt765. [PubMed: 12469133]

26. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, et al. (2008). HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. ACS Chem. Biol 3, 373–382. 10.1021/cb800025k. [PubMed: 18533659]

27. Hodneland CD, Lee Y-S, Min D-H, and Mrksich M (2002). Selective Immobilization of Proteins to Self-Assembled Monolayers Presenting Active Site-Directed Capture Ligands. Proc. Natl. Acad. Sci. U. S. A 99, 5048–5052. 10.1073/pnas.072685299. [PubMed: 11959956]

28. Modica JA, Iderzorig T, and Mrksich M (2020). Design and Synthesis of Megamolecule Mimics of a Therapeutic Antibody. J. Am. Chem. Soc 142, 13657–13661. 10.1021/jacs.0c05093. [PubMed: 32706963]

29. Modica JA, Skarpathiotis S, and Mrksich M (2012). Modular Assembly of Protein Building Blocks To Create Precisely Defined Megamolecules. ChemBioChem 13, 2331–2334. 10.1002/cbic.201200501. [PubMed: 23070998]

30. Modica JA, Lin Y, and Mrksich M (2018). Synthesis of Cyclic Megamolecules. J. Am. Chem. Soc 140, 6391–6399. 10.1021/jacs.8b02665. [PubMed: 29723476]

31. Kimmel BR, Modica JA, Parker K, Dravid V, and Mrksich M (2020). Solid-Phase Synthesis of Megamolecules. J. Am. Chem. Soc 142, 4534–4538. 10.1021/jacs.9b12003. [PubMed: 32105451]

32. Lai Y-T, Reading E, Hura GL, Tsai K-L, Laganowsky A, Asturias FJ, Tainer JA, Robinson CV, and Yeates TO (2014). Structure of a Designed Protein Cage that Self-Assembles into a Highly Porous Cube. Nat. Chem 6, 1065–1071. 10.1038/nchem.2107. [PubMed: 25411884]
33. Divine R, Dang HV, Ueda G, Fallas JA, Vulovic I, Sheffler W, Saini S, Zhao YT, Raj IX, Morawski PA, et al. (2021). Designed Proteins Assemble Antibodies into Modular Nanocages. Science 372, eabd9994. 10.1126/science.abd9994. [PubMed: 33795432]
34. Boyken Scott E, Chen Z, Groves B, Langan Robert A, Oberdorfer G, Ford A, Gilmore Jason M, Xu C, DiMaio F, Pereira Jose H, et al. (2016). De Novo Design of Protein Homo-Oligomers with Modular Hydrogen-Bond Network–Mediated Specificity. Science 352, 680–687. 10.1126/science.aad8865. [PubMed: 27151862]
35. Bale JB, Gonen S, Liu Y, Sheffler W, Ellis D, Thomas C, Cascio D, Yeates TO, Gonen T, King NP, et al. (2016). Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes. Science 353, 389–394. 10.1126/science.aaf8818. [PubMed: 27463675]
36. Chen Z, Johnson MC, Chen J, Bick MJ, Boyken SE, Lin B, De Yoreo JJ, Kollman JM, Baker D, and DiMaio F (2019). Self-Assembling 2D Arrays with de Novo Protein Building Blocks. J. Am. Chem. Soc 141, 8891–8895. 10.1021/jacs.9b01978. [PubMed: 31050411]
37. Sahtoe DD, Praetorius F, Courbet A, Hsia Y, Wicky BIM, Edman NI, Miller LM, Timmermans BJR, Decarreau J, Morris HM, et al. (2022). Reconfigurable Asymmetric Protein Assemblies through Implicit Negative Design. Science 375, eabj7662. 10.1126/science.abj7662. [PubMed: 35050655]
38. Conibear AC (2020). Deciphering Protein Post-Translational Modifications Using Chemical Biology Tools. Nat. Rev. Chem 4, 674–695 10.1038/s41570-020-00223-8.
39. de Marco A (2009). Strategies for Successful Recombinant Expression of Disulfide Bond-Dependent Proteins in Escherichia Coli. Microb. Cell Fact. 8, 26. 10.1186/1475-2859-8-26.
40. Rosano GL, and Ceccarelli EA (2014). Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. Front. Microbiol 5, 411–421. 10.3389/fmicb.2014.00172. [PubMed: 25157245]
41. Singh A, Upadhyay V, Upadhyay AK, Singh SM, and Panda AK (2015). Protein Recovery from Inclusion Bodies of Escherichia Coli Using Mild Solubilization Process. Microb. Cell Fact 14, 41. 10.1186/s12934-015-0222-8. [PubMed: 25889252]
42. Simon AJ, Zhou Y, Ramasubramani V, Glaser J, Pothukuchy A, Gollihar J, Gerberich JC, Leggere JC, Morrow BR, Jung C, et al. (2019). Supercharging Enables Organized Assembly of Synthetic Biomolecules. Nat. Chem 11, 204–212. 10.1038/s41557-018-0196-3. [PubMed: 30643229]
43. Engilberge S, Rennie ML, Dumont E, and Crowley PB (2019). Tuning Protein Frameworks via Auxiliary Supramolecular Interactions. ACS Nano 13, 10343–10350. 10.1021/acsnano.9b04115. [PubMed: 31490058]
44. Rennie ML, Fox GC, Pérez J, and Crowley PB (2018). Auto-regulated Protein Assembly on a Supramolecular Scaffold. Angew. Chem 130, 13960–13965. 10.1002/ange.201807490.
45. McGovern RE, McCarthy AA, and Crowley PB (2014). Protein Assembly Mediated by Sulfonatocalix[4]arene. Chem. Commun 50, 10412–10415. 10.1039/C4CC04897K.
46. Bai Y, Luo Q, and Liu J (2016). Protein Self-Assembly Via Supramolecular Strategies. Chem. Soc. Rev 45, 2756–2767. 10.1039/C6CS00004E. [PubMed: 27080059]
47. Churchfield LA, and Tezcan FA(2019). Design and Construction of Functional Supramolecular Metalloprotein Assemblies. Acc. Chem. Res 52, 345–355. 10.1021/acs.accounts.8b00617. [PubMed: 30698941]
48. Golub E, Subramanian RH, Esselborn J, Alberstein RG, Bailey JB, Chiong JA, Yan X, Booth T, Baker TS, and Tezcan FA (2020). Constructing Protein Polyhedra Via Orthogonal Chemical Interactions. Nature 578, 172–176. 10.1038/s41586-019-1928-2. [PubMed: 31969701]
49. Zhang L, Bailey JB, Subramanian RH, Groisman A, and Tezcan FA (2018). Hyperexpandable, Self-Healing Macromolecular Crystals with Integrated Polymer Networks. Nature 557, 86–91. 10.1038/s41586-018-0057-7. [PubMed: 29720635]
50. Mogilevsky CS, Lobba MJ, Brauer DD, Marmelstein AM, Maza JC, Gleason JM, Doudna JA, and Francis MB (2021). Synthesis of Multi-Protein Complexes through Charge-Directed Sequential

Activation of Tyrosine Residues. J. Am. Chem. Soc 143, 13538–13547. 10.1021/jacs.1c03079. [PubMed: 34382787]

51. Koniev O, and Wagner A (2015). Developments and Recent Advancements in the Field of Endogenous Amino Acid Selective Bond Forming Reactions for Bioconjugation. Chem. Soc. Rev 44, 5495–5551. 10.1039/C5CS00048C. [PubMed: 26000775]
52. Sletten EM, and Bertozzi CR (2009). Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. Angew. Chem 48, 6974–6998. 10.1002/anie.200900942. [PubMed: 19714693]
53. Kolb HC, Finn MG, and Sharpless KB (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew. Chem. Int. Ed 40, 2004–2021. 10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.
54. Devaraj NK (2018). The Future of Bioorthogonal Chemistry. ACS Cent. Sci 4, 952–959. 10.1021/acscentsci.8b00251. [PubMed: 30159392]
55. Szijj P, and Chudasama V (2021). The Renaissance of Chemically Generated Bispecific Antibodies. Nat. Rev. Chem 5, 78–92. 10.1038/s41570-020-00241-6.
56. Griffith BR, Allen BL, Rapraeger AC, and Kiessling LL (2004). A Polymer Scaffold for Protein Oligomerization. J. Am. Chem. Soc 126, 1608–1609. 10.1021/ja037646m. [PubMed: 14871072]
57. Broyer RM, Grover GN, and Maynard HD (2011). Emerging Synthetic Approaches for Protein–Polymer Conjugations. Chem. Commun 47, 2212–2226. 10.1039/C0CC04062B.
58. Tao L, Kaddis CS, Loo RRO, Grover GN, Loo JA, and Maynard HD (2009). Synthesis of Maleimide-End-Functionalized Star Polymers and Multimeric Protein-Polymer Conjugates. Macromolecules 42, 8028–8033. 10.1021/ma901540p. [PubMed: 21544227]
59. Tao L, Kaddis CS, Ogorzalek Loo RR, Grover GN, Loo JA, and Maynard HD (2009). Synthetic Approach to Homodimeric Protein–Polymer Conjugates. Chem. Commun 2148–2150. 10.1039/B822799C.
60. Heredia KL, Grover GN, Tao L, and Maynard HD (2009). Synthesis of Heterotelechelic Polymers for Conjugation of Two Different Proteins. Macromolecules 42, 2360–2367. 10.1021/ma8022712. [PubMed: 25378715]
61. McMillan JR, Hayes OG, Winegar PH, and Mirkin CA (2019). Protein Materials Engineering with DNA. Acc. Chem. Res 52, 1939–1948. 10.1021/acs.accounts.9b00165. [PubMed: 31199115]
62. Stephanopoulos N (2019). Peptide–Oligonucleotide Hybrid Molecules for Bioactive Nanomaterials. Bioconj. Chem 30, 1915–1922. 10.1021/acs.bioconjchem.9b00259
63. Saccà B, and Niemeyer CM (2011). Functionalization of DNA Nanostructures with Proteins. Chem. Soc. Rev 40, 5910–5921. 10.1039/C1CS15212B. [PubMed: 21975573]
64. Seeman NC, and Sleiman HF (2017). DNA Nanotechnology. Nat. Rev. Mater 3, 17068. 10.1038/natrevmats.2017.68.
65. Pinheiro AV, Han D, Shih WM, and Yan H (2011). Challenges and Opportunities for Structural DNA Nanotechnology. Nat. Nanotechnol 6, 763–772. 10.1038/nnano.2011.187. [PubMed: 22056726]
66. Trads JB, Tørring T, and Gothelf KV (2017). Site-Selective Conjugation of Native Proteins with DNA. Acc. Chem. Res 50, 1367–1374. 10.1021/acs.accounts.6b00618. [PubMed: 28485577]
67. Sunasee R, and Narain R (2014). Covalent and Noncovalent Bioconjugation Strategies. In Chemistry of Bioconjugates, Narain R, ed. (John Wiley & Sons, Inc.), pp. 1–75.
68. Yang YR, Liu Y, and Yan H (2015). DNA Nanostructures as Programmable Biomolecular Scaffolds. Bioconj. Chem 26, 1381–1395. 10.1021/acs.bioconjchem.5b00194
69. Hudson WH, and Ortlund EA (2014). The Structure, Function and Evolution of Proteins that Bind DNA and RNA. Nat. Rev. Mol. Cell Biol 15, 749–760. 10.1038/nrm3884. [PubMed: 25269475]
70. McMillan JR, and Mirkin CA (2018). DNA-Functionalized, Bivalent Proteins. J. Am. Chem. Soc 140, 6776–6779. 10.1021/jacs.8b03403. [PubMed: 29799197]
71. Kashiwagi D, Sim S, Niwa T, Taguchi H, and Aida T (2018). Protein Nanotube Selectively Cleavable with DNA: Supramolecular Polymerization of “DNA-Appended Molecular Chaperones”. J. Am. Chem. Soc 140, 26–29. 10.1021/jacs.7b09892. [PubMed: 29226681]
72. McMillan JR, Hayes OG, Remis JP, and Mirkin CA (2018). Programming Protein Polymerization with DNA. J. Am. Chem. Soc 140, 15950–15956. 10.1021/jacs.8b10011. [PubMed: 30407003]

73. Figg CA, Winegar PH, Hayes OG, and Mirkin CA (2020). Controlling the DNA Hybridization Chain Reaction. J. Am. Chem. Soc 142, 8596–8601. 10.1021/jacs.0c02892. [PubMed: 32356981]
74. Coyle MP, Xu Q, Chiang S, Francis MB, and Groves JT (2013). DNA-Mediated Assembly of Protein Heterodimers on Membrane Surfaces. J. Am. Chem. Soc 135, 5012–5016. 10.1021/ja3101215. [PubMed: 23530555]
75. Li H, Park SH, Reif JH, LaBean TH, and Yan H (2004). DNA-Templated Self-Assembly of Protein and Nanoparticle Linear Arrays. J. Am. Chem. Soc 126, 418–419. 10.1021/ja0383367. [PubMed: 14719910]
76. Mou Y, Yu J-Y, Wannier TM, Guo C-L, and Mayo SL (2015). Computational Design of Co-Assembling Protein–DNA Nanowires. Nature 525, 230–233. 10.1038/nature14874. [PubMed: 26331548]
77. Lim S, Kim J, Kim Y, Xu D,and Clark DS (2020). CRISPR/Cas-Directed Programmable Assembly of Multi-Enzyme Complexes. Chem. Commun 56, 4950–4953. 10.1039/D0CC01174F.
78. Wilner OI, Shimron S, Weizmann Y, Wang Z-G, and Willner I (2009). Self-Assembly of Enzymes on DNA Scaffolds: En Route to Biocatalytic Cascades and the Synthesis of Metallic Nanowires. Nano Lett 9, 2040–2043. 10.1021/nl900302z. [PubMed: 19323557]
79. Cheglakov Z, Weizmann Y, Braunschweig AB, Wilner OI, and Willner I (2008). Increasing the Complexity of Periodic Protein Nanostructures by the Rolling- Circle-Amplified Synthesis of Aptamers. Angew. Chem. Int. Ed 47, 126–130. 10.1002/anie.200703688.
80. Erkelenz M, Kuo C-H, and Niemeyer CM (2011). DNA-Mediated Assembly of Cytochrome P450 BM3 Subdomains. J. Am. Chem. Soc 133, 16111–16118. 10.1021/ja204993s. [PubMed: 21919448]
81. Marczynke M, Gröger K, and Seitz O (2017). Selective Binders of the Tandem Src Homology 2 Domains in Syk and Zap70 Protein Kinases by DNA-Programmed Spatial Screening. Bioconj. Chem 28, 2384–2392. 10.1021/acs.bioconjchem.7b00382
82. Setyawati MI, Kutty RV, and Leong DT (2016). DNA Nanostructures Carrying Stoichiometrically Definable Antibodies. Small 12, 5601–5611. 10.1002/smll.201601669. [PubMed: 27571230]
83. Kazane SA, Axup JY, Kim CH, Ciobanu M, Wold ED, Barluenga S, Hutchins BA, Schultz PG, Winssinger N, and Smider VV (2013). Self-Assembled Antibody Multimers through Peptide Nucleic Acid Conjugation. J. Am. Chem. Soc 135, 340–346. 10.1021/ja309505c. [PubMed: 23210862]
84. Meyer R, and Niemeyer CM (2011). Orthogonal Protein Decoration of DNA Nanostructures. Small 7, 3211–3218. 10.1002/smll.201101365. [PubMed: 21953812]
85. Pan L, Cao C, Run C, Zhou L, and Chou JJ (2020). DNA- Mediated Assembly of Multispecific Antibodies for T Cell Engaging and Tumor Killing. Adv. Sci 7, 1900973. 10.1002/advs.201900973.
86. Saccà B, Meyer R, Erkelenz M, Kiko K, Arndt A, Schroeder H, Rabe KS, and Niemeyer CM (2010). Orthogonal Protein Decoration of DNA Origami. Angew. Chem. Int. Ed 49, 9378–9383. 10.1002/anie.201005931.
87. Nguyen TM, Nakata E, Saimura M, Dinh H, and Morii T (2017). Design of Modular Protein Tags for Orthogonal Covalent Bond Formation at Specific DNA Sequences. J. Am. Chem. Soc 139, 8487–8496. 10.1021/jacs.7b01640. [PubMed: 28521084]
88. Liang SI, McFarland JM, Rabuka D, and Gartner ZJ (2014). A Modular Approach for Assembling Aldehyde-Tagged Proteins on DNA Scaffolds. J. Am. Chem. Soc 136, 10850–10853. 10.1021/ja504711n. [PubMed: 25029632]
89. Nielsen TB, Thomsen RP, Mortensen MR, Kjems J, Nielsen PF, Nielsen TE, Kodal ALB, Cló E, and Gothelf KV (2019). Peptide-Directed DNA-Templated Protein Labelling for The Assembly of a Pseudo-IgM. Angew. Chem. Int. Ed 58, 9068–9072. 10.1002/anie.201903134.
90. Chhabra R, Sharma J, Ke Y, Liu Y, Rinker S, Lindsay S, and Yan H(2007). Spatially Addressable Multiprotein Nanoarrays Templated by Aptamer-Tagged DNA Nanoarchitectures. J. Am. Chem. Soc 129, 10304–10305. 10.1021/ja072410u. [PubMed: 17676841]
91. Aslan H, Krissanaprasit A, Besenbacher F, Gothelf KV, and Dong M (2016). Protein Patterning by a DNA Origami Framework. Nanoscale 8, 15233–15240. 10.1039/C6NR03199D. [PubMed: 27487933]

92. Yan H, Park SH, Finkelstein G, Reif JH, and LaBean TH (2003). DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. Science 301, 1882–1884. 10.1126/science.1089389. [PubMed: 14512621]

93. Xu Y, Jiang S, Simmons CR, Narayanan RP, Zhang F, Aziz A-M, Yan H, and Stephanopoulos N (2019). Tunable Nanoscale Cages from Self-Assembling DNA and Protein Building Blocks. ACS Nano 13, 3545–3554. 10.1021/acsnano.8b09798. [PubMed: 30835439]

94. Li S, Jiang Q, Liu S, Zhang Y, Tian Y, Song C, Wang J, Zou Y, Anderson GJ, Han J-Y, et al. (2018). A DNA Nanorobot Functions as a Cancer Therapeutic in Response to a Molecular Trigger In Vivo. Nat. Biotechnol 36, 258–264. 10.1038/nbt.4071. [PubMed: 29431737]

95. Yang YR, Fu J, Wootten S, Qi X, Liu M, Yan H, and Liu Y (2018). 2D Enzyme Cascade Network with Efficient Substrate Channeling by Swinging Arms. ChemBioChem 19, 212–216. 10.1002/cbic.201700613. [PubMed: 29178416]

96. Praetorius F, and Dietz H (2017). Self-Assembly of Genetically Encoded DNA-Protein Hybrid Nanoscale Shapes. Science 355, eaam5488. 10.1126/science.aam5488. [PubMed: 28336611]

97. Ke G, Liu M, Jiang S, Qi X, Yang YR, Wootten S, Zhang F, Zhu Z, Liu Y, and Yang CJ (2016). Directional Regulation of Enzyme Pathways Through the Control of Substrate Channeling on a DNA Origami Scaffold. Angew. Chem 128, 7609–7612. 10.1002/ange.201603183.

98. Wong NY, Xing H, Tan LH, and Lu Y (2013). Nano-Encrypted Morse Code: A Versatile Approach to Programmable and Reversible Nanoscale Assembly and Disassembly. J. Am. Chem. Soc 135, 2931–2934. 10.1021/ja3122284. [PubMed: 23373425]

99. Liu M, Fu J, Qi X, Wootten S, Woodbury NW, Liu Y, and Yan H (2016). A Three-Enzyme Pathway with an Optimised Geometric Arrangement to Facilitate Substrate Transfer. ChemBioChem 17, 1097–1101. 10.1002/cbic.201600103. [PubMed: 26995014]

100. Williams BAR, Lund K, Liu Y, Yan H, and Chaput JC (2007). Self-Assembled Peptide Nanoarrays: An Approach to Studying Protein–Protein Interactions. Angew. Chem. Int. Ed 46, 3051–3054. 10.1002/anie.200603919.

101. Fu J, Yang YR, Dhakal S, Zhao Z, Liu M, Zhang T, Walter NG, and Yan H (2016). Assembly of Multienzyme Complexes on DNA Nanostructures. Nat. Protoc 11, 2243–2273. 10.1038/nprot.2016.139. [PubMed: 27763626]

102. Niemeyer CM, Koehler J, and Wuerdemann C (2002). DNA-Directed Assembly of Bienzymic Complexes from In Vivo Biotinylated NAD(P)H:FMN Oxidoreductase and Luciferase. ChemBioChem 3, 242–245. 10.1002/1439-7633(20020301)3:2/3<242::Aid-cbic242>3.0.Co;2-f. [PubMed: 11921405]

103. Wilner OI, Weizmann Y, Gill R, Lioubashevski O, Freeman R, and Willner I (2009). Enzyme Cascades Activated on Topologically Programmed DNA Scaffolds. Nat. Nanotechnol 4, 249–254. 10.1038/nnano.2009.50. [PubMed: 19350036]

104. Wacker R, and Niemeyer CM (2004). DDI-μFIA—A Readily Configurable Microarray-Fluorescence Immunoassay Based on DNA-Directed Immobilization of Proteins. ChemBioChem 5, 453–459. 10.1002/cbic.200300788. [PubMed: 15185368]

105. Douglas SM, Bachelet I, and Church GM (2012). A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. Science 335, 831–834. 10.1126/science.1214081. [PubMed: 22344439]

106. Lacroix A, Edwardson TGW, Hancock MA, Dore MD, and Sleiman HF (2017). Development of DNA Nanostructures for High-Affinity Binding to Human Serum Albumin. J. Am. Chem. Soc 139, 7355–7362. 10.1021/jacs.7b02917. [PubMed: 28475327]

107. Hayes OG, Partridge BE, and Mirkin CA (2021). Encoding Hierarchical Assembly Pathways of Proteins with DNA. Proc. Natl. Acad. Sci. U. S. A 118, e2106808118. 10.1073/pnas.2106808118. [PubMed: 34593642]

108. Brodin JD, Auyeung E, and Mirkin CA (2015). DNA-Mediated Engineering of Multicomponent Enzyme Crystals. Proc. Natl. Acad. Sci. U. S. A 112, 4564–4569. 10.1073/pnas.1503533112. [PubMed: 25831510]

109. McMillan JR, Brodin JD, Millan JA, Lee B, Olvera de la Cruz M, and Mirkin CA (2017). Modulating Nanoparticle Superlattice Structure Using Proteins with Tunable Bond Distributions. J. Am. Chem. Soc 139, 1754–1757. 10.1021/jacs.6b11893. [PubMed: 28121437]

110. Wang S-T, Minevich B, Liu J, Zhang H, Nykypanchuk D, Byrnes J, Liu W, Bershadsky L, Liu Q, Wang T, et al. (2021). Designed and Biologically Active Protein Lattices. Nat. Commun 12, 3702. 10.1038/s41467-021-23966-4. [PubMed: 34140491]

111. Hayes OG, McMillan JR, Lee B, and Mirkin CA (2018). DNA-Encoded Protein Janus Nanoparticles. J. Am. Chem. Soc 140, 9269–9274. 10.1021/jacs.8b05640. [PubMed: 29989807]

112. Winegar PH, Hayes OG, McMillan JR, Figg CA, Focia PJ, and Mirkin CA (2020). DNA-Directed Protein Packing within Single Crystals. Chem 6, 1007–1017. 10.1016/j.chempr.2020.03.002. [PubMed: 33709040]

113. Partridge BE, Winegar PH, Han Z, and Mirkin CA (2021). Redefining Protein Interfaces within Protein Single Crystals with DNA. J. Am. Chem. Soc 143, 8925–8934. 10.1021/jacs.1c04191. [PubMed: 34096291]

114. Subramanian RH, Smith SJ, Alberstein RG, Bailey JB, Zhang L, Cardone G, Suominen L, Chami M, Stahlberg H, Baker TS, et al. (2018). Self-Assembly of a Designed Nucleoprotein Architecture through Multimodal Interactions. ACS Cent. Sci 4, 1578–1586. 10.1021/acscentsci.8b00745. [PubMed: 30555911]

115. Kashiwagi D, Shen HK, Sim S, Sano K, Ishida Y, Kimura A, Niwa T, Taguchi H, and Aida T (2020). Molecularly Engineered “Janus GroEL”: Application to Supramolecular Copolymerization with a Higher Level of Sequence Control. J. Am. Chem. Soc 142, 13310–13315. 10.1021/jacs.0c05937. [PubMed: 32691585]

116. Strable E, Johnson JE, and Finn MG (2004). Natural Nanochemical Building Blocks: Icosahedral Virus Particles Organized by Attached Oligonucleotides. Nano Lett. 4, 1385–1389. 10.1021/nl0493850.

117. Wang MX, Brodin JD, Millan JA, Seo SE, Girard M, Olvera de la Cruz M, Lee B, and Mirkin CA (2017). Altering DNA-Programmable Colloidal Crystallization Paths by Modulating Particle Repulsion. Nano Lett. 17, 5126–5132. 10.1021/acs.nanolett.7b02502. [PubMed: 28731353]

118. Zhang C, Tian C, Guo F, Liu Z, Jiang W, and Mao C (2012). DNA-Directed Three-Dimensional Protein Organization. Angew. Chem. Int. Ed 51, 3382–3385. 10.1002/anie.201108710.

119. Engelen W, Sigl C, Kadletz K, Willner EM, and Dietz H (2021). Antigen-Triggered Logic-Gating of DNA Nanodevices. J. Am. Chem. Soc 143, 21630–21636. 10.1021/jacs.1c09967. [PubMed: 34927433]

120. Hamann PR, Hinman LM, Hollander I, Beyer CF, Lindh D, Holcomb R, Hallett W, Tsou H-R, Upeslacis J, Shochat D, et al. (2002). Gemtuzumab Ozogamicin, A Potent and Selective Anti-CD33 Antibody–Calicheamicin Conjugate for Treatment of Acute Myeloid Leukemia. Bioconj. Chem 13, 47–58. 10.1021/bc010021y.

Bigger Picture Statement

In nature, many proteins combine to form architectures that contain exact numbers and oligomeric sequences of different subunits, and such hierarchical structural control influences their biological, catalytic, photophysical, and membrane transport properties. To mimic such structures and potentially surpass their properties, the synthesis of different sequence-encoded protein oligomers is needed.

Herein, a method to organize proteins in such a way using designed nucleic acid sequences is reported. Generalizable bioconjugation reactions are used to install one DNA strand on each protein, and assembly occurs via base-pair recognition of different DNA sequences. Therefore, nearly any protein can be used without requiring mutations or recombinant expression. Using this method, monomers, dimers, and trimers are synthesized that can be further assembled into larger structures, laying the foundation for the study of sequence-encoded protein materials across many areas.

Highlights

- A modular DNA scaffold was designed to organize proteins into defined oligomers
- Oligomers were synthesized with exact numbers and sequences of up to five proteins
- This oligomerization strategy is generalizable to nearly any protein

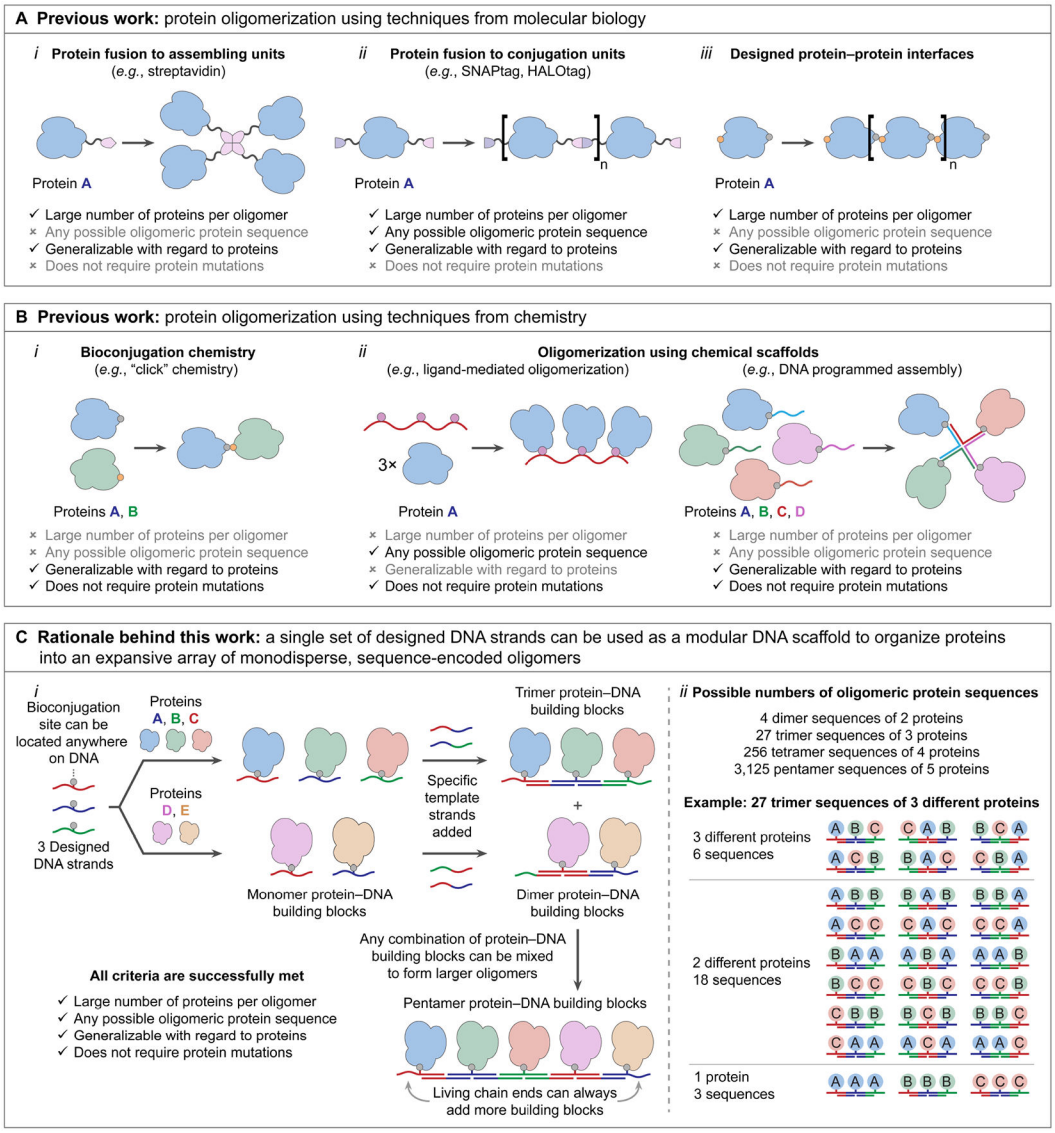


Figure 1. Protein oligomerization techniques and limitations.

(A) Protein oligomers can be synthesized using techniques from molecular biology including (*i*) recombinant expression of linear oligomers, (*ii*) protein fusion to assembling units, and (*iii*) protein fusion to conjugation units. (B) Oligomerization can also be achieved using techniques from chemistry, such as (*i*) bioconjugation chemistry and (*ii*) scaffold-directed oligomerization of proteins. (C) In this work, (*i*) we design a set of six DNA strands that can be used as a modular DNA scaffold to organize proteins into an expansive array of monodisperse and sequence-encoded oligomers. (*ii*) This generalizable method will enable the synthesis of different oligomeric sequences of proteins.

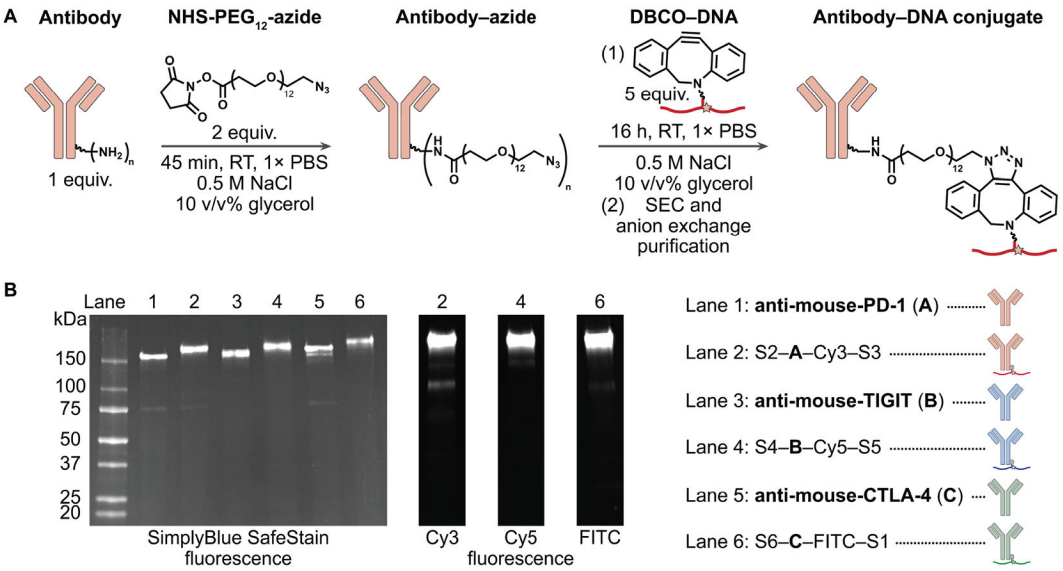


Figure 2. Antibody functionalization with a single DNA strand.

(A) Primary amines on the surface of antibodies were functionalized with azides using an NHS-PEG_{12}-N_3 linker. Next, dibenzocyclooctyne (DBCO)-modified DNA was conjugated to azide-modified antibodies via a strain-promoted azide-alkyne cycloaddition (SPAAC). (B) SDS-PAGE characterization of mouse antibodies (*i.e.*, lane 1: **anti-mouse-PD-1** (**A**), lane 3: **anti-mouse-TIGIT** (**B**), **and lane 5: anti-mouse-CTLA-4** (**C**)) and antibody–DNA conjugates (*i.e.*, lane 2: S2–**A**–Cy3–S3, lane 4: S4–B–Cy5–S5, and lane 6: S6–**C**–FITC–S1). A single gel was imaged for SimplyBlue SafeStain, Cy3, Cy5, and FITC fluorescence.

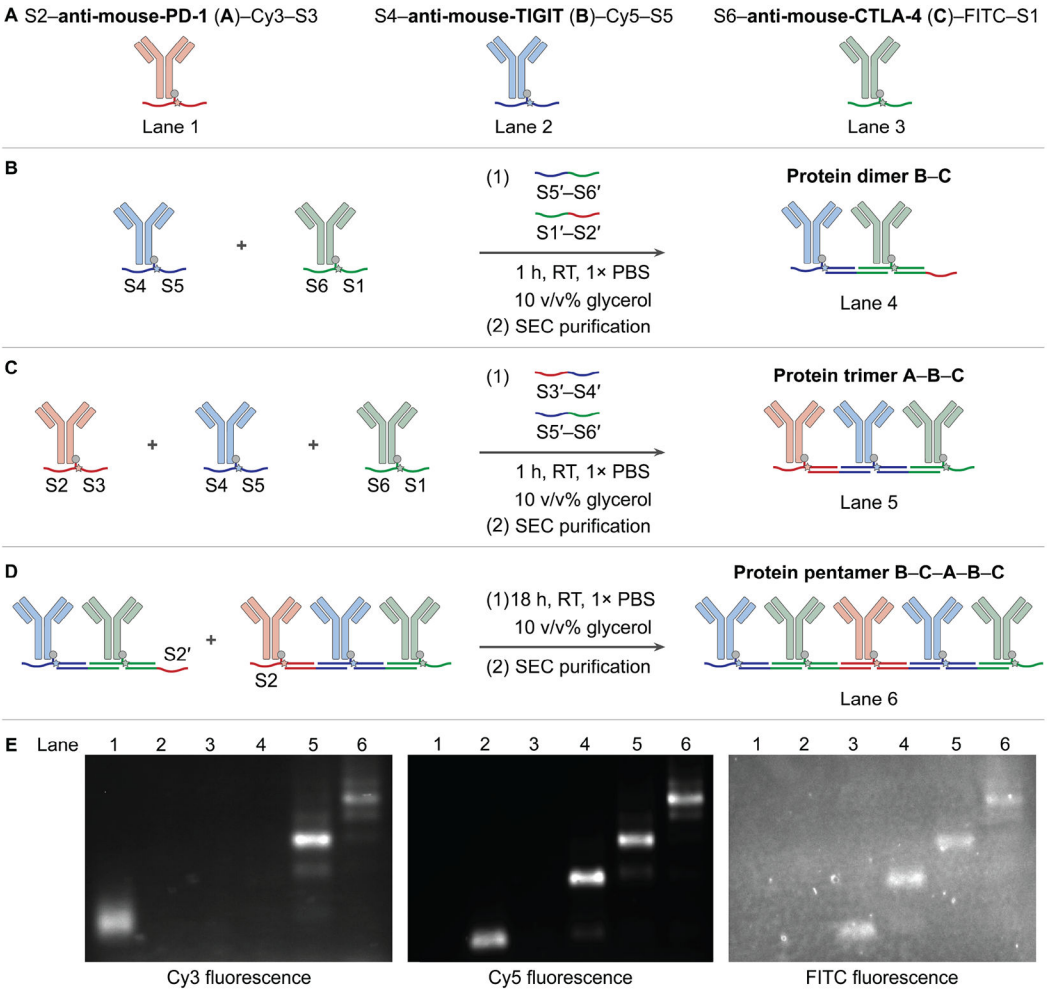


Figure 3. Antibody oligomerization into encoded sequences using DNA–DNA interactions.

(A) Antibody–DNA conjugates and template DNA strands were assembled using DNA–DNA interactions into sequence-encoded protein (B) dimers and (C) trimers. Dimers and trimers were subsequently assembled using DNA–DNA interactions into sequence-encoded protein (D) pentamers. (E) Agarose gel characterization of antibody–DNA conjugates (*i.e.*, lane 1: S2–**anti-mouse-PD-1** (**A**)–Cy3–S3, lane 2: S4–**anti-mouse-TIGIT** (**B**)–Cy5–S5, and lane 3: S6–**anti-mouse-CTLA-4** (**C**)–FITC–S1) along with sequence-encoded antibody dimers (*i.e.*, lane 4: S4–**B**–**C**–S2'), trimers (*i.e.*, lane 5: S2–**A**–**B**–**C**–S1), and pentamers (*i.e.*, lane 6: S4–**B**–**C**–**A**–**B**–**C**–S1). A single gel was imaged for Cy3, Cy5, and FITC fluorescence and these images are merged into one composite image in Figure S9.