Modular engineering and *in vivo* library selection of tissue-targeting protein nanoparticles

Audrey Ellen Olshefsky

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Suzie H. Pun, Co-Chair

Neil P. King, Co-Chair

David Baker

Program Authorized to Offer Degree:

Bioengineering

University of Washington

Abstract

Modular engineering and *in vivo* library selection of tissue-targeting protein nanoparticles

Audrey Ellen Olshefsky

Co-chairs of the Supervisory Committee:
Suzie H. Pun
Department of Bioengineering
Neil P. King
Department of Biochemistry

Targeted delivery remains one of the greatest challenges and opportunities in drug development. In Chapter 1, I review these challenges and opportunities through the lens of the principles of applying self-assembling protein nanomaterials to therapeutic delivery. In Chapter 2, I report an *in vivo* library selection platform based on protein nanoparticles that encapsulate their own genome. Our selection platform offers a unique tool to optimize protein-based therapeutic performance in a living mammal, a complex physiological environment that cannot yet be comprehensively modeled *in silico* or *in vitro*, and to identify targeting ligands of interest. In Chapter 3, I report results towards modularly reprogramming synthetic nucleocapsids for targeted chemotherapeutic delivery by altering both the displayed targeting domains and the encapsulated cargo. I discuss potential future directions in Chapter 4. Together, these results provide new insights and methods for developing self-assembling nanomaterials for *in vivo* targeted delivery.

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my advisors Suzie Pun and Neil King. I could not imagine a more synergistic and dynamic duo to guide me through an enjoyable and rewarding graduate school experience!

I am also very grateful for my committee members (David Baker, Xiaohu Gao, and my GSR Aude Chapuis), Barry Stoddard for leading the training grant program in which I participated in 2018-2020, my incredible collaborators (Halli Benasutti and Jeff Chamberlain here at UW and a team led by Adnan Abu-Yousif at Takeda), and my official and unofficial mentors in the Baker, King, IPD, and Pun Labs (Gabe Butterfield, Marc Lajoie, Heather Gustafson, Meilyn Sylvestre, Ian Cardle, Nataly Kacherovsky, Drew Sellers, Karla Luise-Herpoldt, John Wang, Dan Ellis).

I am thankful for my numerous other lab mates and colleagues at UW, past and present, who've consistently been an important part in a fun and extremely supportive research environment. I am so lucky to have worked with all of you–and there are so many of you that I may have unintentionally left you out! (If so, apologies). I'd like to thank my Pun Lab colleagues (Brynn Olden, Bob Lamm, Gary Liu, David Peeler, Dan Lee, Albert Yen, Emmeline Cheng, Alex Prossnitz, Dan Lee, Jai Raman, Lucy Yang, Yunshi Zhou, Clinton Heinze, Trey Pichon, Kefan Song, Ben Nguyen, Melissa Ling, Abe Wu, Kairui Jiang, Ethan Michelson), IPD colleagues (WiS leaders, Lauren Carter, Brooke Fiala, Anindya Roy, Jake Kraft, Xinru Wang, Erin Yang, Lance Stewart, Sharon Ke, Gargi Kher, Quinton Dowling, Dan Humphrys, Andrew Borst, Alena Khmelinskaia, Aaron Sciore, Stefanie Baker, Yoann Aldon, Masa Somiya, Adam Wargacki, Dane Zambrano, Marti Tooley, Sam Tipps, Marc Langowski, Elias Kinfu, Naveen Jasti, Geoff Hutchinson, Grace Hendricks, Sydney Funk, Helen Eisenach, Annie Dosey, Cameron Criswell, Chloe Adams, Jung-Ho Chun, Cara Chao, Sanela Rankovic, Chelsea Fries, Kate DaPron, Neil Gerstenmaier, Cassie Ogohara, Isaac Sappington, Sam Wrenn, Rashmi Ravichandran, Xinting Li, Stacey Gerbin, Rebecca Cole, Kandise van Wormer, Ratika Krishnamurty, Zari Magness, Courtney Young, Kristina

Herrera), PCR-IO team (Meerit Said, Hugh Haddox, and Ashley Vater and Ryan Caster of the Siegel Lab at UC Davis), Nathaniel Peters from the Keck Institute, my UW 2017 BioE PhD Cohort (Caitlin, Ciana, Elisa, Samantha, Michael, Ian, Saffie, Colleen, Sarah, Qin, and all others), and former UW BioE grad students who helped me navigate through the process of beginning grad school.

I am deeply grateful to Jenny Roberts, Tim Dwyer, Kevin Turner, Penny Chisholm, Andres Cubillos-Ruiz, Allison Coe, Julie Miller, Natalie Kuldell, Carissa Perez Olsen, Robin Dreschler, Sha-Wen Chen, Jennifer Anderson, Cyr de Imus, and Semih Tareen for supporting me at the first stages of my scientific career. I appreciate how impactful positive introductory experiences and superstar mentors are in science, and my gratitude for you all continues to grow.

I'd like to give a shout out to my family, all of my "second families," and my incredible teachers from growing up–I wouldn't be here without all of you (Darren, Sheila, Gma, Gpa, Grandfather, Christine, Jack, and my extended Brogan and Olshefsky family; the Richardson-Grundahl family, the Balises; the MacYoungs; the Franklins; the Eshelmans; the Nielsens/Thorntons; the Roberts; los Marinos, los Fabbri, and many other families; MIT Kappa Alpha Theta; as well as all of my teachers and coaches from FHES, FHMS, FHHS, Dante Alighieri Escuela Bilingua de La Rioja, MIT, and extracurriculars). I feel extremely fortunate to have been supported by such strong communities throughout my life so far.

Last but definitely not least, I'd like to thank my favorite coworker, Christian Richardson, for sharing with me his steady support, whimsical creativity, wise analyses, and quality humor both inside and outside of the lab.

Table of Contents

Dedication

To my dad, who taught me obstacles are things you see when you lose sight of your goals.

To my mom, who taught me balance.

To my grandparents, who have shown our families how to make the most of this life.

And to my sister, who encouraged me to be strong and independent.

I love you all, very much.

Chapter 1. Engineering self-assembling protein nanoparticles for therapeutic delivery *

Audrey Ellen Olshefsky, Christian Richardson, Suzie H. Pun & Neil P. King

Abstract

Despite remarkable advances over the past several decades, many therapeutic nanomaterials fail to overcome major *in vivo* delivery barriers. Controlling immunogenicity, optimizing biodistribution, and engineering environmental responsiveness are key outstanding delivery problems for most nanotherapeutics. However, notable exceptions exist including some lipid and polymeric nanoparticles, some virus-based nanoparticles, and nanoparticle vaccines where immunogenicity is desired. Self-assembling protein nanoparticles offer a powerful blend of modularity and precise designability to the field, and have the potential to solve many of the major barriers to delivery. In this review, we provide a brief overview of key designable features of protein nanoparticles and their implications for therapeutic delivery applications. We anticipate that protein nanoparticles will rapidly grow in their prevalence and impact as clinically-relevant delivery platforms.

* This chapter is reprinted with permission from: Olshefsky, Richardson, Pun, King. Engineering Self-Assembling Protein Nanoparticles for Therapeutic Delivery. Bioconjug Chem. 2022 Nov 16;33(11):2018-2034.

Introduction

Over the past 30 years, tremendous advances in nanoparticle platforms for therapeutic applications have been achieved.[1,2] In 1995, the first cancer-treating nanoparticle was approved by the FDA: Doxil, a liposomal formulation of doxorubicin that significantly decreased cardiomyopathy caused by free doxorubicin.[3,4] Less than three decades later, the first two mRNA vaccines were approved by the FDA: Pfizer-BioNtech's Comirnaty and Moderna's Spikevax, advanced lipid nanoparticle formulations that deliver synthetic mRNA to elicit highly protective immune responses against SARS-CoV-2.[5,6] Despite these and many other advances, it is widely believed that nanotherapeutics have yet to achieve their full potential due to the inconsistent translation from *in vitro* results in cell culture to *in vivo* results in animal models, and preclinical results in animal models to clinical efficacy in humans.[1,7,8] The clinical translation of nanotherapeutics is fundamentally limited by several biological barriers to delivery, manifesting as a much lower than expected clinical trial success rate (**Fig. 1**).[9] Successful therapeutic delivery systems must (1) efficiently localize to the body compartment or organ of interest (e.g., tumor), (2) localize to the cell type of interest (e.g., cancer cells), and (3) engage with the target cells in a therapeutic manner (e.g., kill cancer cells). Overcoming these sequential biological barriers requires precise engineering to endow nanoparticles with functionalities suited to their therapeutic application. Self-assembling protein nanoparticles are an emerging class of delivery vehicles with the potential to satisfy these requirements through directed evolution or rational design.

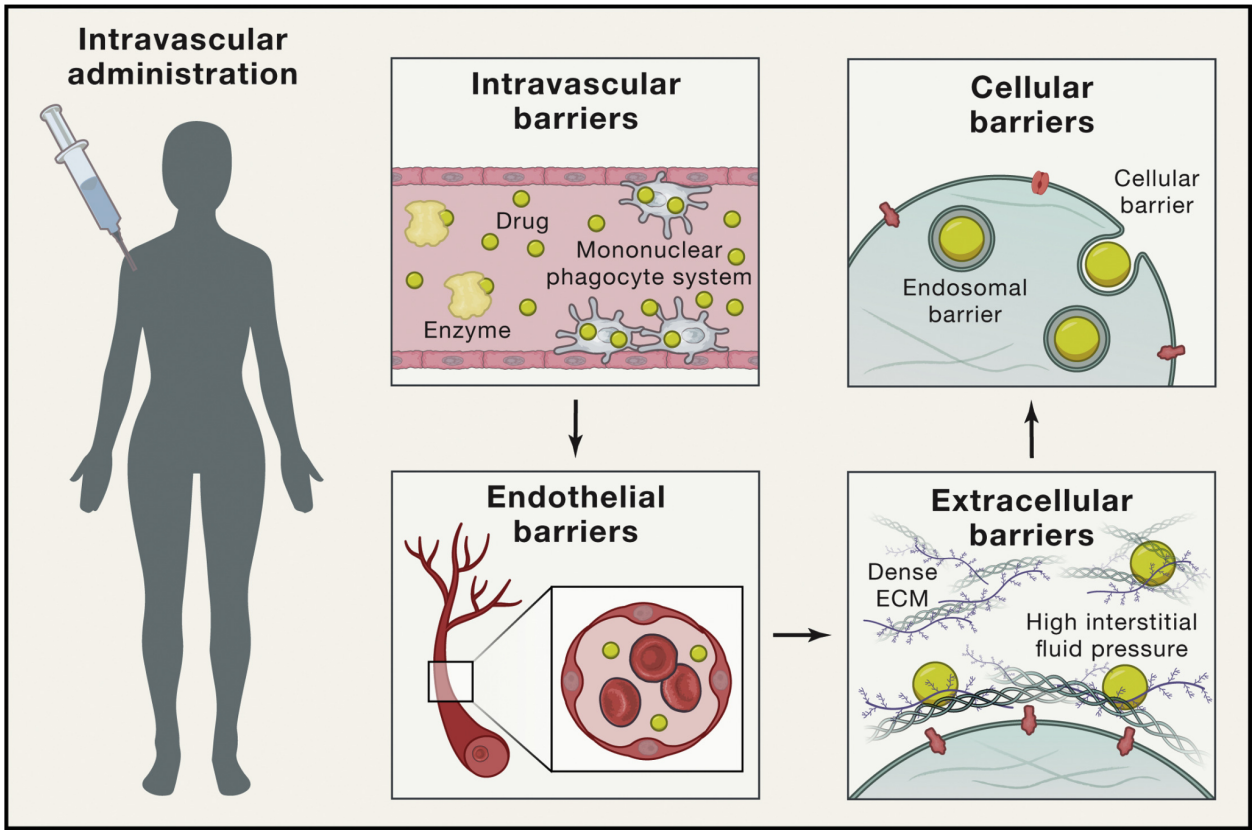


Figure 1. Barriers to *in vivo* nanoparticle delivery following intravascular administration.
For nanoparticles to successfully reach their targets, they must pass several barriers. This example uses intravascular administration as a case study; other routes of administration are reviewed elsewhere.[10–17] Unless immune responses are specifically desired, the nanoparticles must evade recognition by the innate and adaptive immune system during circulation (e.g., the complement system, the mononuclear phagocytic system (MPS), and B cells). Once near the target organ, they must travel past endothelial cells and tight junctions, which line blood vessels. Then the nanoparticles must travel through extracellular matrices within the target tissue, and once they reach the target cells, must access their intended subcellular compartment for therapeutic payload delivery. *Reprinted from Cell, Vol. 181, Issue 1, Zongmin Zhao, Anvay Ukidve, Jayoung Kim, Samir Mitragotri, Targeting Strategies for Tissue-Specific Drug Delivery, Pages 151-167, Copyright 2020, with permission from Elsevier.*[8]

Here, we define self-assembling protein nanoparticles as atomically precise, bounded assemblies composed of symmetric protein oligomers forming three-dimensional shells with hollow interiors. The term "protein nanoparticles" refers to these assemblies unless otherwise noted. As discussed further in "Nanoparticle architecture and geometry," these assemblies have rotational symmetry axes projected into all three dimensions. Examples of such assemblies include computationally designed protein nanoparticles, and natural, non-viral protein nanocontainers like ferritin.[18–21] Although albumin and silk fibroin nanoparticles are also composed of proteins, they do not form atomically precise nanoparticles with defined symmetries.[22,23] Viral vector and virus-derived particle engineering is also outside the scope

of this review due to their distinct complexity. However, we will draw comparisons between many viral and non-viral protein assemblies, as virus-derived delivery systems are a major source of inspiration. Therapeutic applications of these other protein-based nanotherapeutics are thoroughly reviewed elsewhere.[24–30]

Part 1: Designable features of protein nanoparticles

Introduction

Proteins are arguably the most information-rich molecules in the known universe. The information encoded in their amino acid sequences enables proteins to perform a remarkable variety of sophisticated functions that largely constitute the molecular basis of life. The greater the ability to harness this information for precise design, the more sophisticated and diverse the encodable structures can be.[31,32] Computational protein design has become a widely used technique that helps manage the complexity of protein structures and energetics and, in conjunction with complementary approaches, can be used to develop functionalities that can be integrated into protein nanoparticles.[33–36] The main features engineered for delivery purposes are physicochemical properties, environmental responsiveness, scaffold functionalization, and cargo encapsulation.[37] In the following sections, we discuss how these features can be tuned and detail considerations for producing these materials at sufficient scale for clinical application.

Protein nanoparticle scaffold design

To date, three primary approaches to designing and modifying protein nanoparticles have been explored in delivery applications: top-down modification of natural protein nanocontainers, bottom-up design of novel protein nanoparticles, and directed evolution of nanoparticle scaffolds to confer specific functions.

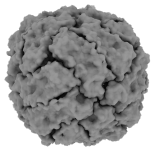
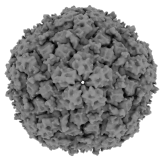
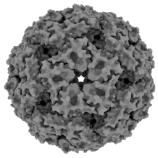
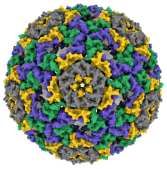
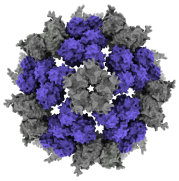
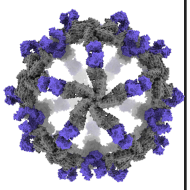
Currently, the most straightforward nanoparticle design method is top-down adaptation of an existing, naturally evolved viral capsid or cellular protein nanocontainer.[38,39] Natural protein nanoparticles like

ferritins, encapsulins, and lumazine synthases evolved over billions of years to serve highly specialized biological functions in specific environments, with many functionalities that would be virtually impossible to design using current methods.(**Table 1**).[33,40–42] Natural protein nanoparticles can also be mutated or modified through the fusion of functional protein domains to confer additional properties. For example, Seo, Yoo, and Kim et al reported a ferritin nanoparticle with a fused fibrinolytic domain to improve tumor accumulation of coadministered anticancer drugs.[33] However, even modified natural protein nanocontainers are largely confined to a design space that is close to their initial states, making it challenging but not impossible to alter their fundamental structural and biochemical properties. Directed evolution, discussed below, has become a particularly powerful tool for modifying some of these underlying properties.

Recent advances in protein engineering have enabled the bottom-up design of purpose-built protein nanoparticles, drastically expanding the protein nanoparticle design space. These methods typically apply principles of symmetry to arrange natural or *de novo* designed cyclic oligomers into protein nanoparticles through interface design or rigid genetic fusion, reviewed in Khmelinskaia, Wargacki, and King (2021).[43–49] Padilla, Colovos, and Yeates et al (2001) reported an early example of a self-assembling protein nanoparticle constructed using the rigid fusion of several oligomeric domains.[48] Recently computational protein design has been applied to interface and rigid fusion methods enabling the design of increasingly complex protein nanoparticles with improving success rates.[19,50,51] King and Baker et al (2012) reported self-assembling protein nanoparticles with a variety of symmetries using computational protein interface design to arrange existing cyclic oligomers into the desired assemblies (**Figure 2B**).[19] Many computational protein design methods currently require specific technical expertise and access to substantial computational resources. However, advances in machine learning are rapidly simplifying the protein design process and reducing the computational burden.[52] This promises to make computational protein design a more widely accessible technique.

Directed evolution is a powerful tool for refining existing protein nanoparticles.[53] This method has been used for a variety of modifications ranging from improving nanoparticle biodistribution, packaging and protection of nucleic acid, and the sequestration of toxic enzymes.[54–56] Notably, Tetter, Terasaka, and Steinauer et al recently reported the design and directed evolution of an icosahedral nanocontainer into a self-mRNA encapsidating nanoparticle with a diameter almost twice as large as the native structure.[56,57] Directed evolution experiments require both a selection assay that will drive designs exclusively towards a desired function and an initial protein nanoparticle that is reasonably evolutionarily close to that function. Choosing a selection assay and initial protein nanoparticle is a nuanced process that requires careful consideration, and these experiments can be labor intensive. On the other hand, a well-designed directed evolution experiment is a powerful tool that can be used to select for poorly understood features that would currently be impossible with rational design alone.[55,56] As demonstrated in many viral vector gene delivery studies, a good selection assay can refine thousands to millions of highly diverse protein variants in a single experiment, making this a powerful complement to computational protein nanoparticle design.[58–61]

Table 1. Selected examples of non-viral protein nanoparticle platforms.

Name	Model[62]	Global Symmetry	Number of Chains and Oligomers in Assembly	Diameter	PDB/ EMDB	Ref.
***H. pylori* ferritin**		Octahedral	24 monomers arranged as 12 dimers	12 nm	1FHA	[63]
***T. maritima* encapsulin**		Icosahedral	60 monomers arranged as 12 pentamers	24 nm	3DKT	[64]
***A. aeolicus* lumazine synthase (AaLS)**		Icosahedral	60 monomers arranged as 12 pentamers	15 nm	1HQK	[65]
Evolved lumazine synthase NC-4		Icosahedral	240 monomers arranged as 12 pseudopentamers and 20 trimers	30 nm	7A4J	[56]
Computationally designed two-component nanoparticle		Icosahedral	120 monomers arranged as 60 trimers and 60 pentamers	27 nm	EMD-0350	[18,55]
Computationally designed antibody nanoparticle		Icosahedral	36 monomers arranged as 12 *de novo* trimers and 12 Fc dimers	25 nm	In Ref.	[21]

Nanoparticle models were made using UCSF ChimeraX.[62]

Nanoparticle architecture and geometry

Nanoparticle geometry describes the arrangement of subunits composing a protein nanoparticle. All protein nanoparticles are constructed from many copies of an asymmetric unit that are arranged in a symmetric manner such that they form a closed, three-dimensional structure (**Fig. 2A-C**). Tetrahedral, octahedral, icosahedral, and some dihedral symmetries are most often used to generate these closed structures.[39,66,67] Cyclic homo-oligomers that can be incorporated into a protein nanoparticle usually need to be on a matched axis of symmetry. For example, a cyclic tetramer must be on the axis of four-fold symmetry in an octahedrally symmetric nanoparticle and is not easily incorporated into tetrahedral and icosahedral nanoparticles. The symmetry group also determines the number of asymmetric units in a complete nanoparticle, and thus the valency of any ligands displayed on the surface of the nanoparticle. Though possible in other symmetry groups, icosahedral architectures are typically more easily designed to have large interior cavities and non-porous shells that can aid in cargo packaging and protection; the myriad naturally existing, icosahedral nanocontainers benefit from these advantages.[42,64,68] While protein nanoparticle subunit size can vary widely, symmetry groups with larger numbers of asymmetric units will yield larger nanoparticles for a subunit of a given size.

The asymmetric unit is composed of one or more protein chains and at least two protein-protein interfaces oriented about axes of cyclic symmetry.[69] An asymmetric unit may be made up of a single chain with two interfaces or have several chains with internal non-symmetric protein-protein interfaces.[18,70] As discussed further in "Manufacturing protein nanoparticles," if an asymmetric unit is composed of two or more chains, the nanoparticle can in principle be assembled *in vitro*. The number of protein chains and their arrangement within the asymmetric unit determines how many unique fusion points (N and C termini) are available on the interior and exterior of the protein nanoparticle for the addition of functional domains.

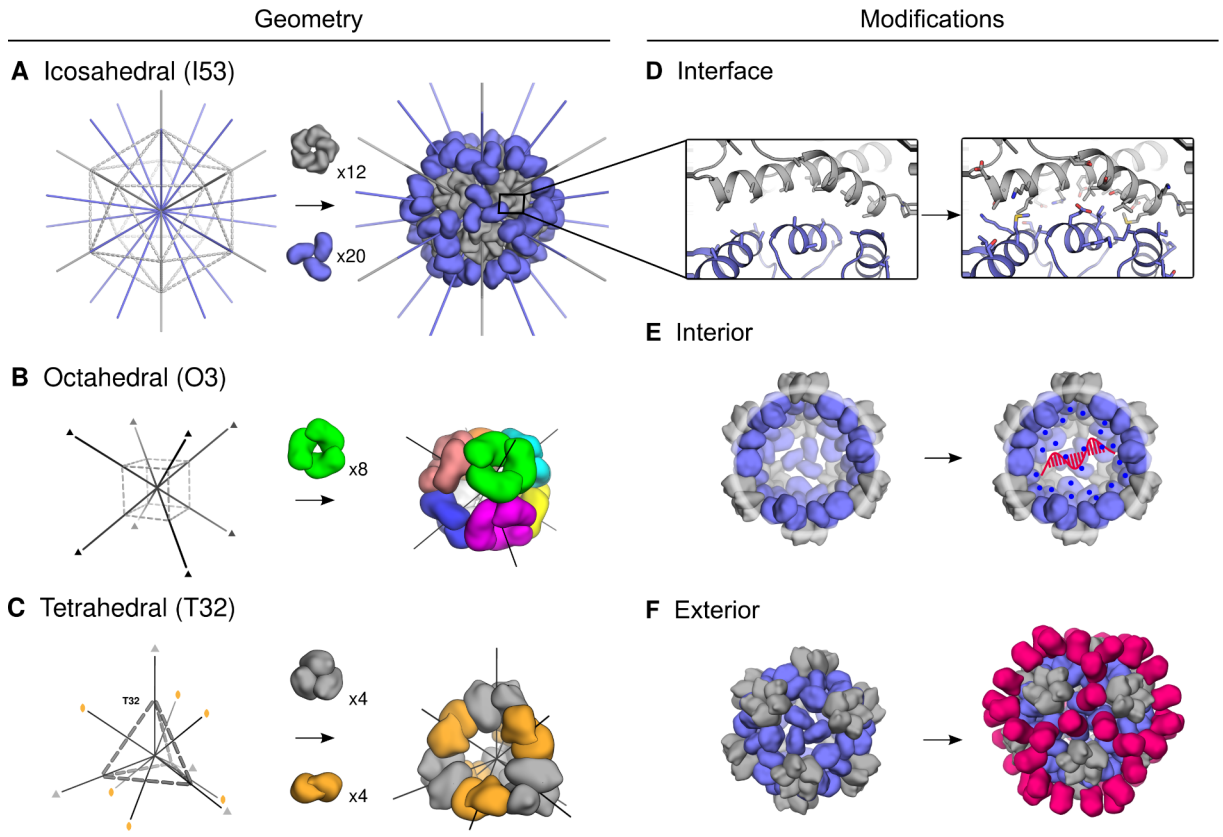


Figure 2. Designable features of nanoparticle scaffolds.
Two main considerations when designing protein nanoparticle delivery systems are scaffold source and scaffold modifications. **(A-C)** Protein nanoparticle geometries commonly used for delivery applications have icosahedral, octahedral, tetrahedral, or dihedral symmetry.[18,19,71] Notably, the octahedral nanoparticle in **(B)** is composed of eight identical trimer subunits, which are colored differently to help visually distinguish individual subunits in the context of the global structure. **(D-F)** Additional functional elements are designed into the nanoparticle through subunit interface, interior, and exterior modifications.[25] **(D)** Interfaces between the trimer subunits (slate) and pentamer subunits (gray) were computationally designed. **(E)** Interior residues were mutated to hold a net positive charge, leading to electrostatic association of mRNA. **(F)** Additional protein domains can be displayed on the surface of existing nanoparticles. **A, D:** *Reprinted with permission pending.*[18] **B:** *Reprinted with permission pending.*[19] **C:** *Reprinted with permission from Springer Nature, Accurate design of co-assembling multi-component protein nanomaterials, Neil P. King et al, Copyright 2014.*[71]

Interface design and modifications

The high degree of cooperativity in protein nanoparticle assembly requires special consideration when designing protein nanoparticle interfaces.[72–74] As with most protein-protein interface design, protein nanoparticle interface design has to date generally relied on hydrophobic packing of interface residues.[43] However, protein nanoparticles assemble with a high degree of cooperativity, which results in efficient assembly of subunits with relatively weak interfaces but can result in kinetic trapping of partial or

off-target assemblies if the inter-subunit interface is too strong.[75] Large hydrophobic interfaces between protein nanoparticle subunits can also interfere with soluble expression. Designing hydrophilic interfaces, as seen in many natural protein nanocontainers, could be one way to overcome this limitation.[65]

Protein nanoparticle interface design can be leveraged to encode environmental responsiveness and nanoparticle functionalization. Protein nanoparticle interfaces can be either scavenged from existing protein-protein interfaces and rigidly fused to a nanoparticle subunit or computationally designed from scratch. This flexibility can enable protein nanoparticles to directly scaffold natural proteins such as antibodies, thus functionalizing the nanoparticle.[21] Additionally, protein nanoparticle interfaces can be selected or designed to allow for controlled assembly and disassembly of the protein nanoparticle under specific environmental conditions. Ionic strength, pH, and metal-dependent protein nanoparticles have been investigated for applications in both packaging and selective drug delivery.[45,76,77]

Interior modifications

In addition to modifying protein subunit interfaces, useful features can be engineered into both the interior and exterior surfaces of nanoparticles through rational design or library selection.[42,55–57] Interior modification approaches are largely inspired by viruses and virus-like particles.[24,78,79] Three main strategies have been used to genetically modify protein nanoparticle interior and exterior surfaces: point mutation, genetic fusion, and loop insertion. Due to the symmetry and repetitive nature of self-assembling protein nanoparticles, small modifications are often unlikely to destabilize protein subunits, but can cause significant functional changes. Genetic modifications often enable chemical conjugations and post-translational modifications, providing opportunities for functionalization that benefit from years of research and development in both academia and industry.[80–82]

The interior nanoparticle surface is commonly modified to encapsulate specific materials (**Fig. 2E**). It has been widely demonstrated that mutating interior-facing surface residues to hold a net charge facilitates

electrostatic-mediated cargo encapsulation (e.g., encapsulation of super-charged fluorescent proteins or nucleic acids) (**Fig. 2E**).[18,55,83] Interior surface residues can also be mutated to include side chains capable of specific chemical reactions like copper-free click chemistry.[84] Alternatively, larger domains like affinity peptides can be seamlessly integrated into the protein nanoparticle interior through genetic fusion. Genetically fusing functional domains to the nanoparticle interior surface can either be bioactive themselves or enable specific cargo encapsulation such as mRNA and siRNA sequences.[55,57,83,85]

Exterior modifications

Interactions between nanoparticles and their surrounding environments are primarily mediated by the exterior surfaces of the nanoparticles (**Fig. 2F**).[86,87] Two common goals of exterior modifications are (1) to alter the physicochemical properties of the scaffold surface and (2) to display a molecular recognition domain (e.g., a pathogen-derived antigen or receptor targeting domain). Similar to interior modification strategies, surface point mutations, chemical conjugations, genetic fusion, and loop insertions are commonly used to modify exterior surfaces.[33,55,88–92] These methods have been used to increase *in vivo* circulation half-life and to incorporate post-translational modifications like glycosylation, which has been reported to increase germinal center delivery for vaccine applications.[86,93] Surface modifications also have important implications for immunogenicity. “Stealth” coatings like poly-ethylene glycol (PEG) are often conjugated to nanoparticle surfaces to increase circulation time and reduce immunogenicity, but this strategy is clinically limited since many humans are sensitized to PEG.[94]

Displaying functional domains on nanoparticle surfaces controls specific biological interactions like targeted delivery to tumor vasculature or immune responses to pathogen-derived antigens in the context of nanoparticle vaccine delivery. As described above, the display of such domains is fundamentally linked to the symmetry and stoichiometry of the nanoparticle platform. Monomeric domains like tumor-targeting DARPins are symmetry-agnostic, while the display of multimeric antigens like influenza hemagglutinin usually requires fusion to a symmetry-matched nanoparticle component and sometimes facilitates antigen

stabilization.[41,95–98] A recent counterexample demonstrated the symmetry-agnostic display of influenza antigens on the surface of a VLP through SpyCatcher-SpyTag conjugation, an alternative to the genetic fusion approach, which has proven to be another robust and versatile technology for modifying nanoparticle exteriors.[99]

Manufacturing protein nanoparticles

Protein production efficiency impacts the cost and market distribution of a therapeutic, thereby determining its commercial viability.[100,101] The main considerations in protein nanoparticle manufacturing are the dose, protein production, nanoparticle assembly, and purification. Vaccines typically require much smaller doses than other protein-based therapeutics such as monoclonal antibodies (micrograms vs. hundreds of milligrams), making it easier to achieve commercially viable production efficiency.[102–104] However, most of the clinical production of protein nanoparticle immunogens has occurred only in the past few years and there is still much to learn (e.g., National Clinical Trials (NTC) 05007951, 04896086, 03186781, 03814720, 04579250, 04784767, 04645147, 05001373, and 03547245 reported at clinicaltrials.gov).

Process development for manufacturing large-scale protein nanoparticle therapeutics must adhere to production efficiency requirements and current good manufacturing process (cGMP) standards.[105,106] Fortunately, the ability to perform *in vitro* nanoparticle assembly offers several advantages for manufacturing multi-component protein nanoparticles. *In vitro* assembly enables two (or more) standard recombinant biologics to be separately purified before nanoparticle assembly, benefiting from decades of research and development in manufacturing recombinant protein biologics like monoclonal antibodies.[106,107] Assembling protein nanoparticles *in vitro* increases sample purity and gives engineers precise control over subunit composition and cargo encapsulation. While simple protein nanoparticle components can be successfully produced and purified in *E. coli*, many components require mammalian post-translational modifications and must be produced in eukaryotic cells.[108,109] *In vitro* assembly also

enables the generation of mosaic nanoparticles simply by mixing different component-antigen fusions together such that different antigens are co-displayed on the same nanoparticle scaffold. Phase I clinical trials are currently underway for mosaic influenza vaccines.[96] Notably, SpyCatcher/SpyTag conjugation of antigens to nanoparticle scaffolds offers an alternative approach to *in vitro* mosaic assembly; mosaics generated using this method are anticipated to reach the clinic soon.[110] While each therapeutic application will likely require unique process development, improving the manufacturing capabilities of protein nanoparticle immunogens will provide useful information to the entire field.

A distinguishing characteristic of protein nanoparticles is their complete genetic encodability. With the recent emergence of clinically validated techniques for delivering genetic information *in vivo,* this feature opens up a unique opportunity to deploy certain nanoparticle therapeutics and vaccines as nucleic acids.[111–115] In concept, genetic delivery would streamline manufacturing and allow rapid prototyping/iteration *in vivo*. However, the inability to control or purify the translated protein(s) *in vivo* emphasizes the importance of optimizing the sequence of the nanoparticle therapeutic during design, as this will determine critical functional features such as expression level, monodispersity, and stability.

Part 1 summary

The key design variables for engineering self-assembling protein nanoparticles as therapeutics are summarized in **Table 2**. While computational design enables control over protein nanoparticle structure, purification and *in vitro* assembly methods enable control over the exact nanoparticle composition. As further discussed in "Therapeutic applications of protein nanoparticles," the nanoparticle characteristics combined with the delivery method critically influence the pharmacokinetics, biodistribution, and efficacy of the protein nanoparticle therapeutic.

Table 2. Key variables influencing the potential outcomes and applications of self-assembling protein nanoparticles.

Property	Variable	Considerations, applications, and outcomes
Scaffold source	● Design strategy	● Modification of a natural scaffold ○ Simple and expedient source of protein nanoparticles ○ Have relatively fixed structural properties ● Rational design ○ High degree of control over features of the nanoparticle ○ Requires a fundamental understanding of the features being designed ○ Can choose or design subunits that are more robust to modification, making the system more modular and alleviating manufacturing challenges ● Evolution ○ Allows selecting for incompletely understood behavior ○ Powerful complement to rational design ○ Labor-intensive ○ Selection strategies can be nuanced
	● Source species	● Recombinant proteins with similarities to the patient's proteins may elicit autoimmune responses
Scaffold architecture	● Nanoparticle symmetry group (geometry)	● Types of cyclic oligomers that can be displayed ● Valency of displayed ligands ● Interior cavity size and porosity
	● Subunits in asymmetric unit	● Number of unique ligands that can be displayed ● Whether or not *in vitro* assembly is possible
Interior and exterior modifications	● Genetically encoded interior or exterior modifications	● Surface amino acid mutations ● Post-translational modifications ● Protein display (e.g., antigen, binding domain, enzyme)
	● Chemical modifications	● Chemical conjugation: Maleimide conjugation to cysteines, NHS ester conjugation to lysines, etc. ● Surface display or cargo loading: Conjugation to polymers, PEG, small molecule drugs, etc.
	● Choice of subunit(s) to modify	● Modular modification, e.g., co-display of two different protein domains (one on each subunit)
Interface Design & Modification	● Strength of interface	● Nanoparticle assembly efficiency and monodispersity

	• Hydrogen bond networks -& Metal Chelation	• Responsiveness to environmental factors (e.g. pH)
Nanoparticle Subunit Choice	• Physicochemical properties	• Hydrophobicity, charge, isoelectric point, and zeta potential affect colloidal stability in buffer solutions and biological fluids • Hydrophobicity and positive charge may increase unwanted cell and tissue accumulation
	• Intrinsic subunit features	• Complex environmental sensing & interaction • Sophisticated functions (e.g., molecular recognition or catalysis)
Manufacturing	• Protein expression host organism	• Expression of components in mammalian cells enables low endotoxin and native post-translational modifications
	• Assembly strategy	• *In vivo* assembly upon expression inside *E. coli* allows for encapsulation of biomolecules like self-mRNA • *In vitro* assembly enables high sample purity and precise control over subunit composition and encapsulated cargo

Part 2: Therapeutic applications of protein nanoparticles

Introduction

In this section, we discuss five common applications of protein nanotherapeutics, the key design criteria that must be considered for each, and their development status. We summarize this information in **Tables 3 and 4**.

While many design criteria are unique to each application, several criteria are shared amongst all *in vivo* therapeutic applications. These criteria generally affect nanoparticle pharmacokinetics and biodistribution, which have been extensively studied using inorganic and polymeric nanoparticle platforms.[10–17,116] The key physicochemical properties influencing these phenomena are shape, size, charge, hydrophobicity, rigidity, and specific molecular interactions. For reference, the physicochemical properties of protein

nanoparticles are extremely diverse but are generally spherical, 10-30 nm in diameter, hold a net negative surface charge at physiological pH, have hydrophilic surfaces, and can be designed to include or exclude specific molecular interactions (e.g., display a binding domain to target a specific receptor). Regardless of delivery route, nanoparticles that are too hydrophobic, too positively charged, too large, or too small face significant delivery barriers such as aggregation, nonspecific cell uptake, liver and spleen accumulation, and kidney clearance (**Fig. 3**). Nanoparticles ~10-150 nm more efficiently travel from the delivery site to the target tissue (e.g., through blood, extracellular matrix, lymphatic system, or mucous membranes). Still, protein corona formation, immunogenicity, and clearance by the mononuclear phagocytic system (MPS) are critical barriers to most delivery systems.[17,117,118] While certain overarching biodistribution phenomena like MPS clearance appear to be consistent between inorganic, polymeric, and protein nanoparticles, the surface chemistries of such nanoparticles are distinct. It is necessary to consider if, and how, findings in one class of materials (e.g., PLGA-based nanoparticles) translate to other classes of materials (e.g., self-assembling protein nanoparticles). Designing successful nanotherapeutic platforms is a balancing act between the design criteria, platform features, and fundamental delivery barriers.

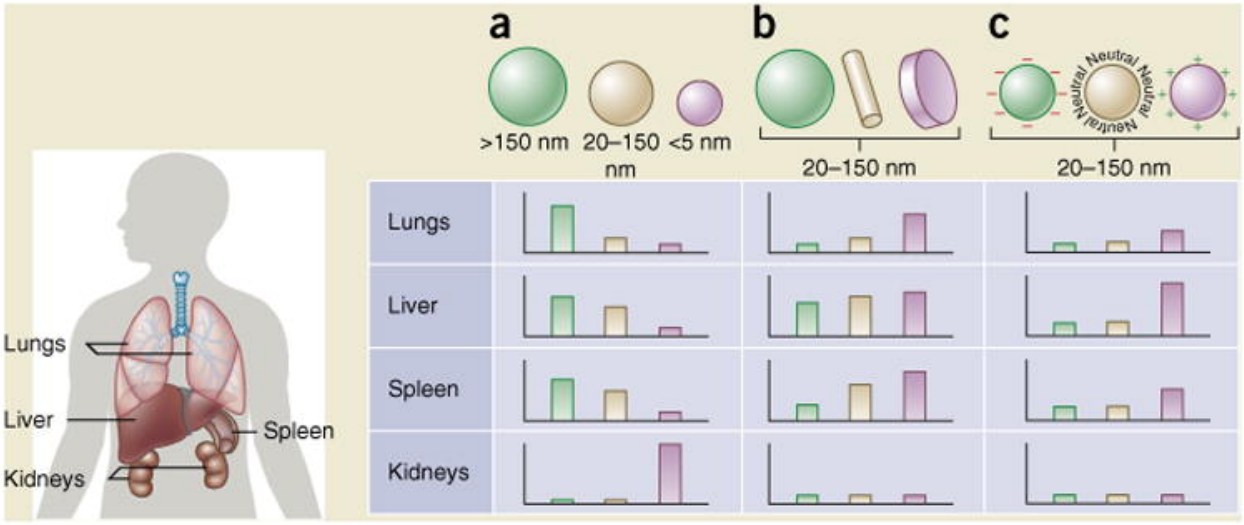


Figure 3. Qualitative impacts of nanoparticle size, shape, and surface charge on biodistribution.
The general effects of surface physicochemical properties on biodistribution were comprehensively reviewed by Blanco, Shen & Ferrari and qualitatively graphed as relative accumulation in major mouse organs.[10] This figure includes data from gold nanoparticles, liposomes, polymer micelles, zwitterionic nanoparticles, hydrogel nanoparticles, and more. (A) Nanoparticles greater than about 150 nm in diameter show increased accumulation in

Figure 3 continued.
the lungs, liver, and spleen, while nanoparticles less than 5 nm in diameter show rapid renal clearance. (B) Spherical nanoparticles tend to have the least uptake by major clearance organs compared to cylindrical and discoidal nanoparticles. (C) Nanoparticles with positively charged surfaces show much higher nonspecific uptake than nanoparticles with negatively charged or neutral surfaces. *Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature Biotechnology, Principles of nanoparticle design for overcoming biological barriers to drug delivery, Elvin Blanco et al, Copyright 2015.*[10]

Vaccine delivery

The goal of vaccines is to safely educate the immune system how to protect against infection or disease upon subsequent encounter with a pathogen. While traditional vaccines derived from whole pathogens provide effective prevention against many diseases, these are inappropriate or have fallen short for several diseases due to factors including safety considerations, insufficient immune stimulation, poor antigen stability, or engineering and manufacturing limitations.[119] Protein nanoparticle immunogens, which deliver antigen in a repetitive array by genetic fusion or chemical or protein-peptide conjugation to nanoparticle scaffolds, aim to address many of these limitations and have been shown to elicit protective immune responses (**Fig. 4**).[97,98,110,120–126] This approach taps into multiple features of the immune system that have evolved to detect "particulate" pathogens featuring repetitive antigenic determinants. First, upon intramuscular injection, protein nanoparticles are the ideal shape and size to be taken up by peripheral or lymph node dendritic cells (spherical, ~20-150 nm in diameter).[11] Additionally, multivalent antigen display facilitates detection of pathogen-associated molecular patterns: protein nanoparticle-mediated B cell receptor cross-linking has been demonstrated to cause stronger immune responses than those elicited by soluble antigens.[117,127,128] Although live-attenuated and inactivated vaccines also benefit from these immunological phenomena, self-assembling protein nanoparticles are widely considered safer since there is no infectious agent. Protein nanoparticles also allow the use of antigens that are engineered in ways that may be incompatible with viral or bacterial growth, such as prefusion-stabilized RSV F (**Fig. 4A**).[129,130] The engineered protein nanoparticle immunogen may possess, for example, increased stability, increasing immune exposure to the antigenic conformation or epitope against which an immune response is desired.[98,131–133] Like VLP-based vaccines, the natural fit between the properties of self-assembling protein nanoparticles and vaccine delivery criteria is likely responsible for vaccines being the only application of

protein nanoparticles that has advanced to the clinic to date – the delivery efficiency of vaccines is less fundamentally limited by immunogenicity compared to other therapeutic applications (**Table 3**).[96,98,122,134–137] However, additional data on the development of antibody responses against nanoparticle scaffolds and their effects is needed.

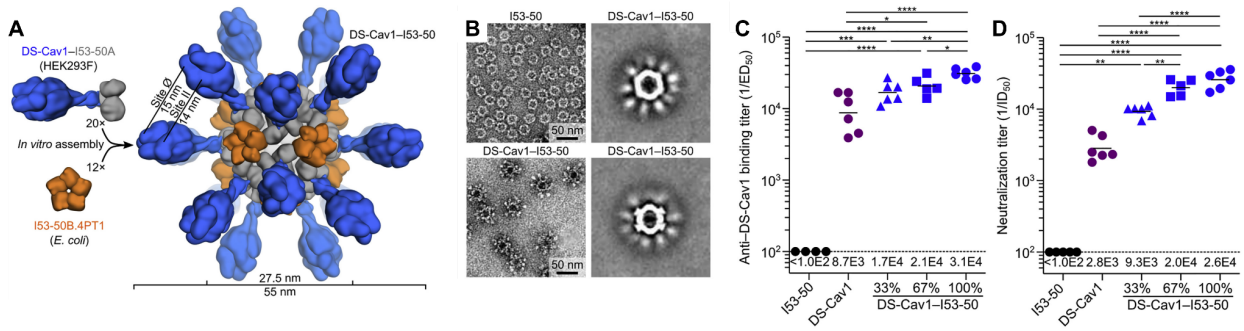


Figure 4. The DS-Cav1-I53-50 nanoparticle immunogen.
(A) A vaccine against respiratory syncytial virus (RSV) was engineered by outwardly displaying a trimeric RSV antigen (DS-Cav1) on the trimeric subunit of a computationally designed nanoparticle (I53-50). **(B)** Negative stain electron microscopy of I53-50 and DS-Cav1-I53-50 nanoparticles. Left: representative micrographs. Right: averages of nanoparticle micrographs. **(C-D)** Antibody binding titers of DS-Cav-1-specific antibodies **(C)** or serum neutralizing antibodies **(D)** from mice immunized with bare nanoparticle (I53-50), free immunogen (DS-Cav1), or nanoparticle immunogens (DS-Cav1-I53-50) at valencies of 33%, 67%, or 100%. Each point represents and individual animal, geometric means are represented by horizontal lines and indicated at the bottom of the plots, and statistical significance is indicated: *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001. See original publication for more details. *Reprinted by permission under the terms of the Creative Commons CC-BY licence, Cell, Vol. 176, Issue 6, Marcandalli J & Fiala B et al, Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus, Pages 1420-1431.e17, Copyright 2019.*[98]

Table 3. Examples of self-assembling protein nanoparticles in the clinic.

Scaffold	Example application	Clinical status
I53_dn5 (computationally designed nanoparticle)	Influenza vaccine [96]	Phase 1 clinical trial (NTC04896086)
I53-50 (computationally designed nanoparticle)	SARS-CoV-2 vaccine [122]	Phase 3 clinical trial (NCT05007951)
eOD-GT8 60-mer (engineered envelope glycoprotein (Env))	HIV vaccine [124]	Phase 1 clinical trials (NCT05001373, NCT03547245)
Ferritin (modified natural protein nanoparticle)	Epstein-Barr virus (EBV) vaccine [121]	Phase 1 clinical trial (NCT04645147)

Current vaccine delivery research with protein nanoparticle immunogens is focused on understanding the mechanisms that underlie their interactions with the immune system and developing new technological features that expand their capabilities.[86,93,96,110,128,138–145] The modularity of self-assembling protein nanoparticles and increasingly precise methods for designing them are beginning to allow systematic testing of important features including nanoparticle architecture (e.g., icosahedral vs. tetrahedral scaffolds that present antigen in distinct geometries), antigen identities and combinations (e.g., mosaic nanoparticle immunogens for broadly protective vaccines), post-translational modifications (e.g., glycosylation to increase germinal center accumulation), and surface modifications to tune immunogenicity or focus the immune response to specific sites of interest (e.g., molecular or steric shielding).[51,86,93,96,110,124,146]

Small molecule delivery

The hydrophobic nature of many small molecule drugs often results in poor solubility or nonspecific accumulation throughout the body, reducing therapeutic bioavailability and increasing unwanted toxicities.[147] Loading small molecules onto a carrier such as an antibody is an effective way to improve their biodistribution.[148] However, the hydrophobic small molecule is often chemically conjugated to the surface of non-nanoparticle carriers, limiting drug loading capacity while also requiring careful optimization of drug-carrier ratio to maintain stability and desired biodistribution.[149] Protein nanoparticle carriers possess an interior cavity for encapsulation of hydrophobic small molecules (**Fig. 3E)**. For example, ferritin nanoparticles have been engineered to encapsulate doxorubicin, cisplatin, and other small molecule therapeutics.[20,41,88,150–155] Drug-loaded ferritin nanoparticles have not yet been clinically approved due to complexity in manufacturing and specific tissue targeting, but promising advancements are underway.

Recently, a ferritin nanoparticle was engineered to display a fibrinolytic domain that can dissolve blood clots in tumors. The coadministration of this ferritin nanocage (FNC) and doxorubicin lowered the risk of adverse blood clot-related events and increased doxorubicin tumor penetration in mice (**Fig. 5**).[33] The synergy between these two delivery systems enabled small molecule delivery (by doxorubicin) and extracellular biologics delivery (by fibrinolytic activity), highlighting the potential of combination therapies.

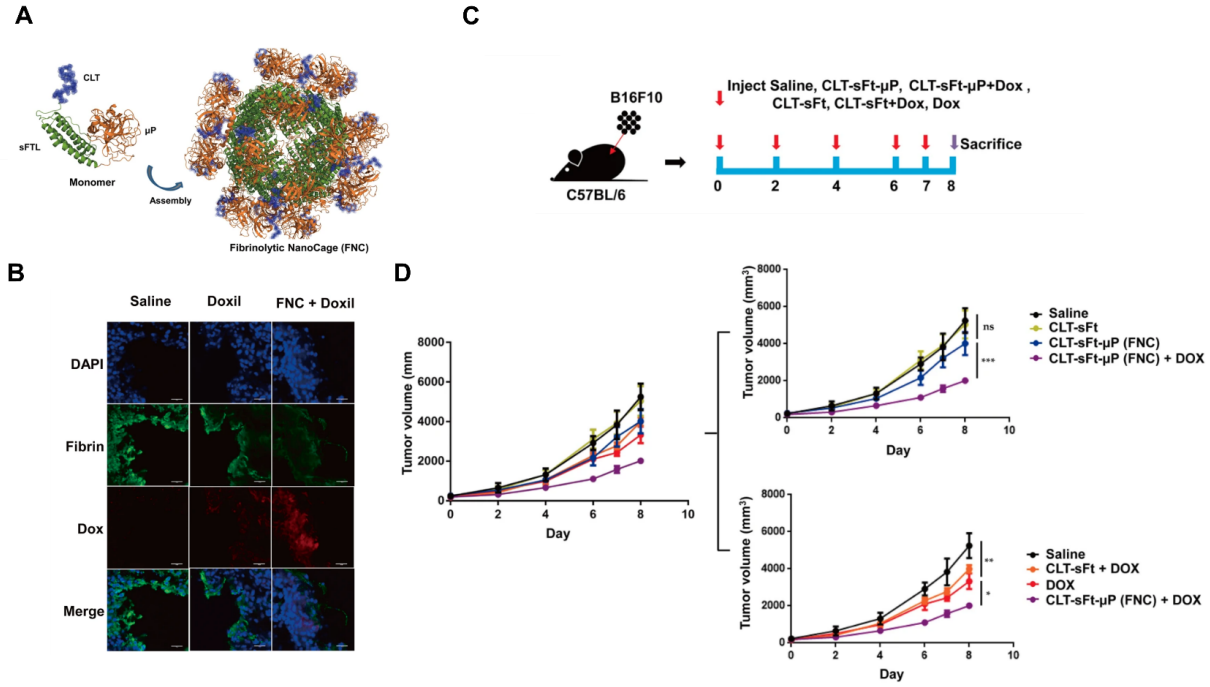


Figure 5. Coadministration of doxorubicin with ferritin co-displaying clot-targeting and fibrinolytic domains. **(A)** The ferritin monomer was engineered to display a clot-targeting peptide (CLT) and fibrinolytic domain (μP), resulting in surface display on ferritin nanoparticles. **(B)** Fibrinolytic nanocages (FNC) coadministered with Doxil show increased tumor cell accumulation and decreased fibrin signal compared to saline or Doxil controls. **(C-D)** When delivered to tumor-bearing mice at the indicated dosing schedule (C), Doxil-FNC coadministration show increased tumor volume reduction (purple) compared to mice treated with Doxil alone (red) and other controls (D). *Reprinted by permission under the terms of the Creative Commons Attribution 4.0 International License, Experimental & Molecular Medicine, Volume 53, Pages 1592–1601, Seo J et al, Fibrinolytic nanocages dissolve clots in the tumor microenvironment, improving the distribution and therapeutic efficacy of anticancer drugs, Copyright 2021.*[33]

Extracellular biologics delivery

Targeted stimulation or inhibition of cell signaling receptors is the key mechanism of many therapeutics ranging from receptor tyrosine kinase inhibitors to bispecific antibodies. Clinical translation of many such

therapeutics is hindered by challenges with ligand stability, pharmacokinetics, biodistribution, and hepatotoxicity.[156–158] Displaying receptor agonists or antagonists on protein nanoparticles offers several opportunities to address these challenges: the (ant)agonist can be stabilized by scaffolding on a solid foundation, the signaling magnitude can be tuned through altering display valency and nanoparticle architecture, and the biodistribution profile can be improved through optimizing the physicochemical features of the nanoparticle.[21,38,159–161] Inspired by receptor signaling applications based on VLPs, ferritin and computationally designed nanoparticles have recently been engineered to display a variety of therapeutic receptor signaling domains. Examples include interleukin-4 receptor (IL-4R)-targeting peptides to ameliorate asthma symptoms, anti-mesenchymal epithelial transition (MET) peptide pharmacophores to stimulate hepatocyte growth factor receptor, death receptor ligands (TRAIL) to kill tumor cells, and angiopoietin-agonizing antibodies to promote angiogenesis (**Fig. 6**).[21,38,159,162,163] These modular platforms offer exciting opportunities for research and development through systematically varying the nanoparticle features to interrogate and optimize biological responses.

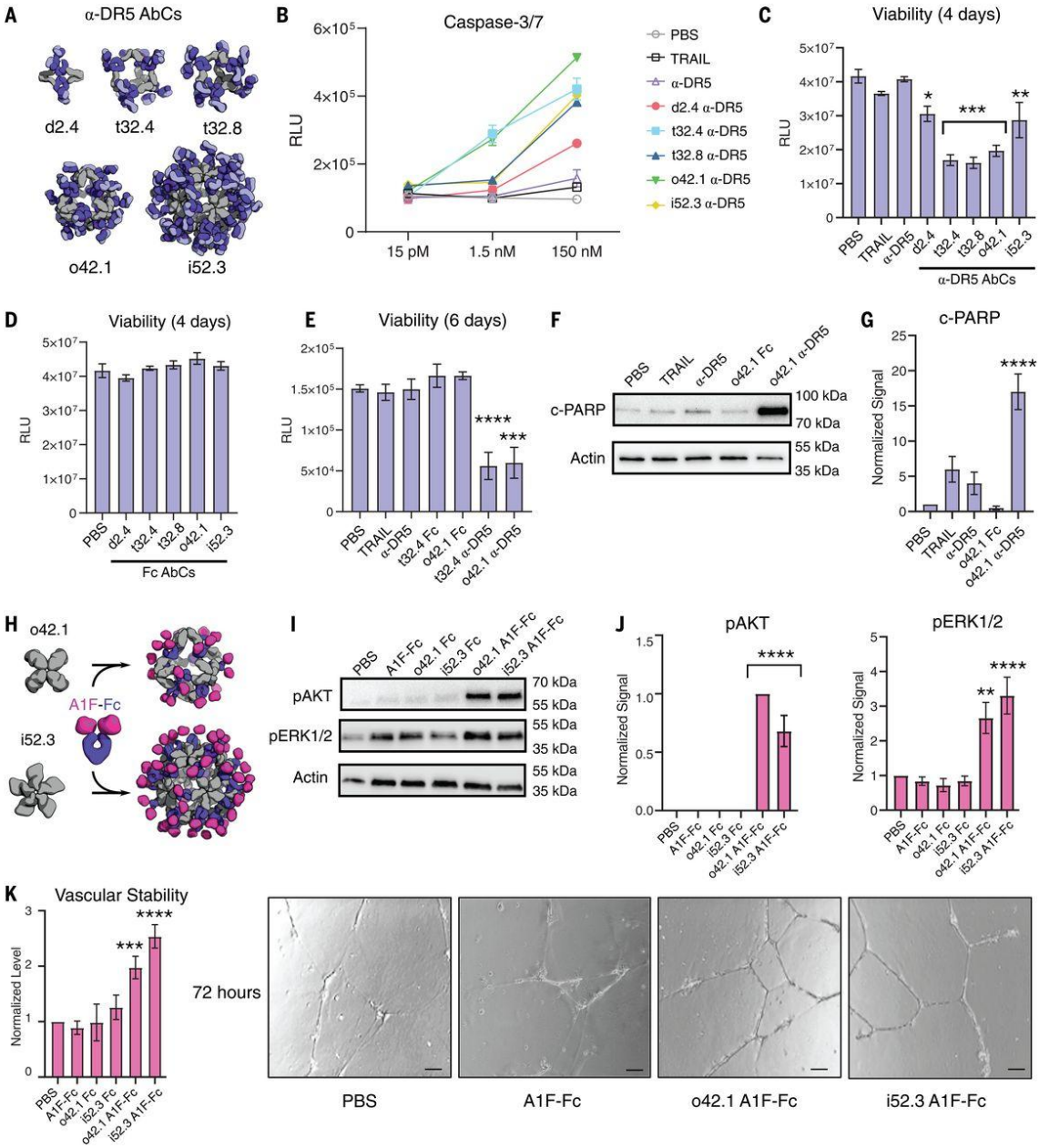


Figure 6. Computationally designed antibody cages (AbCs) activate apoptosis and angiogenic signaling pathways.

(A and B) Caspase-3/7 is activated by AbCs formed with α-DR5 antibody (A), but not the free antibody, in RCC4 renal cancer cells (B). (C and D) α-DR5 AbCs (C), but not Fc AbC controls (D), reduce cell viability 4 days after treatment. (E) α-DR5 AbCs reduce viability 6 days after treatment. (F and G) o42.1 α-DR5 AbCs enhance PARP cleavage, a marker of apoptotic signaling; (G) is a quantification of (F) relative to PBS control. (H) The F-domain from angiopoietin-1 was fused to Fc (A1F-Fc) and assembled into octahedral (o42.1) and icosahedral (i52.3) AbCs. (I) Representative Western blots show that A1F-Fc AbCs, but not controls, increase pAKT and pERK1/2 signals. (J) Quantification of (I): pAKT quantification is normalized to o42.1 A1F-Fc signaling (no pAKT signal in the PBS control); pERK1/2 is normalized to PBS. (K) A1F-Fc AbCs increase vascular stability after 72 hours. (Left) Quantification of vascular stability compared with PBS. (Right) Representative images; scale bars, 100 μm. All error bars represent means ± SEM; means were compared using analysis of variance and Dunnett post-hoc tests (tables S8 and S9). *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001. *Reprinted by permission from American Association for the Advancement of Science, Science, Volume 372, Issue 6537, eabd9994, Divine R et al, Designed proteins assemble antibodies into modular nanocages, Copyright 2021.*[21]

Intracellular biologics delivery

The goal of intracellular biologics delivery is to package macromolecular drugs, target specific cell types, and deliver these membrane-impermeable molecules to specific subcellular locations for therapeutic effect. Viral vectors, virus-like particles, and lipid nanoparticles have been extensively developed for these applications, seeing clinical use as genetic vaccines, *ex vivo* gene editing for CAR-T therapies, and *in vivo* gene therapies.[111–115,164,165] However, expansion of these technologies to new clinical applications is challenging due to MPS clearance and neutralizing antibody responses.[17,117,118]

Designed protein nanoparticles are a comparatively nascent technology but have the potential to greatly expand the intracellular biologics delivery design space. Protein nanoparticles that deliver biologics such as siRNA to cultured cells through the endolysosomal pathway have been reported (**Fig. 7**).[166] Protein nanoparticles have also been reported with modifications that improve circulation half-life and target the nanoparticles to specific cells.[55,167] For systemic, *in vivo* delivery, protein nanoparticles need to incorporate design features that provide stability during circulation with the ability to recognize, enter, and traffic within target cells. Viruses, VLPs, and some designed delivery vehicles achieve this by responding to the endosomal environment and switching states, satisfying the opposing needs of extracellular and endosomal activity.[111,168–170] An exciting new space to explore with protein nanoparticles is engineering environmental responsive mechanisms, inspired by virus-derived particles.

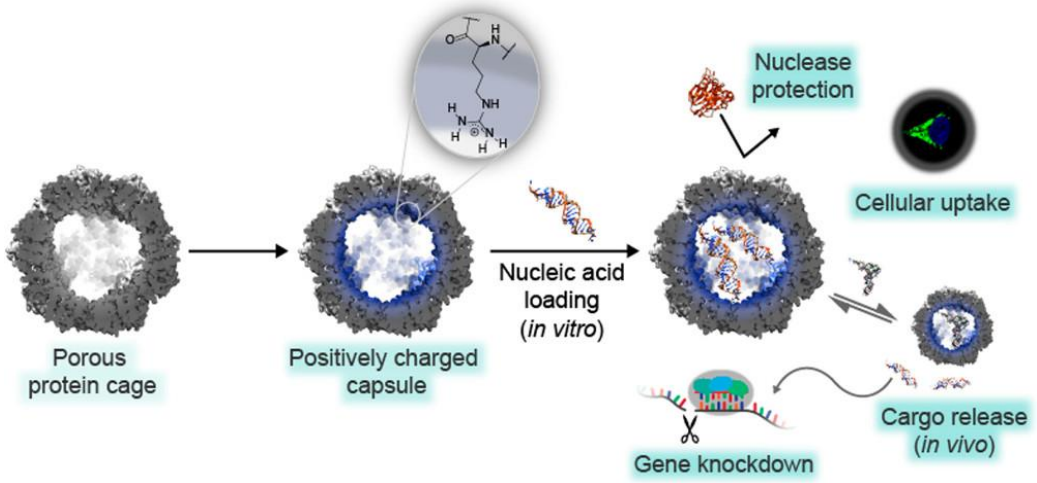


Figure 7. Computationally designed nanoparticle O3-33 was redesigned for siRNA delivery.
The computationally designed octahedral protein nanoparticle O3-33 (porous protein cage) containing 144 hexahistidine tags on its surface was redesigned to have a positively charged interior to electrostatically associate with nucleic acid *in vitro* (positively charged capsule loaded with nucleic acid).[19,166] The authors reported nuclease protection, uptake of the nanoparticles by HeLa cells, and subsequent cargo release leading to knockdown of intracellular GFP mRNA. The authors attributed successful endosomal escape to the hexahistidine tags. *Reprinted by permission from the American Chemical Society, Journal of the American Chemical Society, 140(33):10439–10442, Edwardson TGW et al, Rational Engineering of a Designed Protein Cage for siRNA Delivery, Copyright 2018.*[166]

Theranostics

Combining minimally invasive diagnostic techniques with therapeutic delivery, "theranostics" conveniently synergizes the strengths of two medical procedures.[26,83,171–173] Non-invasive imaging techniques like magnetic resonance imaging (MRI) are vital for diagnosing and treating disease, yet have low resolution at the cellular and molecular level. Protein nanoparticles can provide alternative optical properties and could be engineered to inform researchers or physicians about real-time cellular and molecular processes through environmental responsiveness.[174,175] However, knowledge on protein nanoparticle pharmacokinetics and biodistribution is often lacking.[176] The ability to switch between "imaging mode" (e.g., encapsulating gadolinium) and "therapeutic mode" (e.g., encapsulating a small molecule drug), or to simultaneously operate in both modes, offers an opportunity to more deeply probe the *in vivo* behavior of protein nanoparticles. Ferritin was recently used in this manner to deliver a cytotoxic peptide to tumors while simultaneously displaying green fluorescent protein (GFP) (**Fig. 8A-C**).[83] The option to operate in imaging mode, therapeutic mode, or both also offers a potential strategy

for screening patients for therapeutic response before treatment.[171] Theranostics could be especially impactful in the field of solid tumor delivery, where the heterogeneity within and between tumor types remains a medical challenge.[177–179] However, like most other delivery applications, clinical translation of protein nanoparticle theranostics is still fundamentally limited by scaffold immunogenicity.

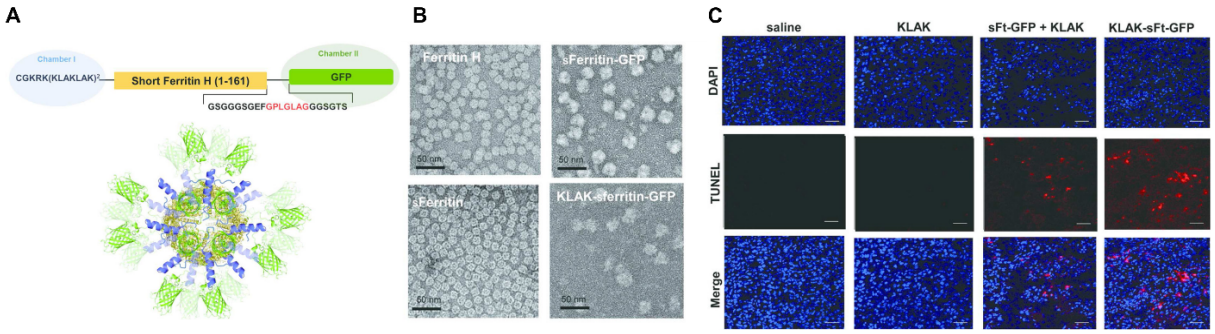


Figure 8. Example of ferritin engineered as a theranostic by displaying a fluorescent protein and encapsulating a cytotoxic peptide.
(A) Schematic for genetically fusing a cytotoxic peptide (chamber 1) and a fluorescent protein (chamber 2) to ferritin. **(B)** Electron micrographs of ferritin, ferritin displaying GFP, and ferritin fused to both GFP and a cytotoxic peptide ("KLAK" repeats). **(C)** Tumor cryosections from mice treated with a saline control, cytotoxic peptide (KLAK), ferritin-GFP and cytotoxic peptide (sFt-GFP + KLAK), or ferritin fused to both cytotoxic peptide and GFP (KLAK-sFt-GFP). Scale bars are 40 μm. *Reprinted by permission from American Chemical Society, Biomacromolecules, 17(1):12-9, Kim S et al, Double-Chambered Ferritin Platform: Dual-Function Payloads of Cytotoxic Peptides and Fluorescent Protein, Copyright 2016.*[83]

Table 4. Platform requirements for therapeutic applications of protein nanoparticles.

Application	Translational status	Platform requirements
Vaccine delivery	Clinical trials	• Stable display of antigens • Drains to lymph nodes • Internalized by APCs • Compatible with adjuvant • Does not elicit autoimmune responses against patients
Receptor signaling	Preclinical	• Stable display of receptor agonist or antagonist • Localizes to target receptors

Small molecule delivery	Preclinical	• Stable encapsulation of small molecules
		• Accumulates in target organ or cell type
		• Mechanism exists for small molecule release
		  ○ Carrier degradation
		  ○ Controlled release

Biologics delivery	Preclinical	• Capacity to encapsulate biologic of interest • Accumulates in target organ or cell type • Endocytosis by target cells • Endosomal escape

		• Controlled release of biologic
Theranostics	Preclinical	• Stably encapsulates contrast agent • Accumulates in tissue(s) of interest • Contrast agent signal is detectable in specific environment of interest • Performs desired function in specific environment (e.g., drug release)

Remaining challenges and opportunities for protein-based nanotherapeutics

Adaptive and innate immune responses are perhaps the largest challenge for the broad application of protein nanoparticle therapeutics.[180,181] Preclinical and clinical experience with therapeutics and vaccines based on viral vectors has clearly established vector-neutralizing antibodies as a significant limitation. Antibody responses against the nanoparticle surface and MPS clearance would likely reduce therapeutic efficacy by preventing cargo delivery to the target tissues or cells.[17,117,118] Altering physicochemical surface properties like lipid envelopes or glycosylation could significantly reduce such antibody responses. Although preliminary studies suggest that structural features such as aspect ratio and display of phagocytosis-preventing signaling molecules could reduce MPS clearance, new concepts and strategies will be needed to minimize the impact of immune responses against protein nanoparticle therapeutics.[13–15,182] The growing body of data on protein nanoparticle vaccines—where immune responses are desirable instead of problematic—provides a valuable opportunity to learn how protein nanoparticles are perceived by the immune system.[183,184] Lessons learned from these studies could be used to develop techniques for engineering immune evasion into protein nanoparticles.

Challenges in design and manufacturing hinder the development of efficacious protein nanotherapeutics. Although computational advances are broadening the accessibility of protein design, repurposing naturally existing protein nanocontainers and designing novel nanomaterials still require significant expertise and effort. Successfully designed protein nanotherapeutics must be manufactured from nucleic

acid templates in biological systems, posing a challenge to meeting the demands of clinical production efficiency but offering opportunities for genetic delivery. Fortunately, molecular analysis and quality control of protein nanoparticles benefits greatly from the abundance of nanoscale analytical techniques. The monodisperse and atomically precise nature of protein nanoparticles enables straightforward and reliable use of analytical techniques, allowing for rapid design-build-test cycles.

There are many opportunities for meaningful advancement in the near future of protein nanoparticle engineering. Naturally occurring proteins and protein nanoparticles present innumerable examples of the sensitive and precise environmental responsiveness proteins can achieve; learning how to harness and engineer these kinds of responsiveness into protein nanoparticle therapeutics is a major opportunity for the field.[111,170] Engineers are now able to selectively design and evolve nanoparticles with features such as larger interior cavities, non-spherical scaffolds, and specific responses to environmental cues of interest.[33,46,56] Engineering and clinically translating self-assembling protein nanoparticles promises to be an exciting and abundant area of research over the next several years.

Declaration of interests

A.O., C.R., and S.H.P. declare no conflicts of interest. N.P.K. is a co-founder, shareholder, paid consultant, and chair of the scientific advisory board of Icosavax, Inc. The King lab has received an unrelated sponsored research agreement from Pfizer.

Acknowledgments

We would like to thank Chelsea Fries for comments on the manuscript. This work was supported by the National Institutes of Health (P50 AI150464 to N.P.K., 1R21CA232430-01 to S.H.P. and N.P.K.., and 1R01CA257563 to S.H.P.), the National Science Foundation NSF (CHE 1629214 to N.P.K. and DGE-2140004 to C.R.), and the Audacious Project. Figures 2, 4, 5, and 8 were made using Inkscape.[185] The TOC graphic was created with Biorender.com.

The following journal content was used: **Figure 1:** Reprinted by permission from Cell, Vol. 181, Issue 1, Zongmin Zhao, Anvay Ukidve, Jayoung Kim, Samir Mitragotri, Targeting Strategies for Tissue-Specific Drug Delivery, Pages 151-167, Copyright 2020, with permission from Elsevier.[8] **Figure 2: A, D:** Reprinted with permission pending.[18] **B:** Reprinted with permission pending.[19] **C:** Reprinted with permission from Springer Nature, Accurate design of co-assembling multi-component protein nanomaterials, Neil P. King et al, Copyright 2014.[71] **Figure 3:** Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature Biotechnology, Principles of nanoparticle design for overcoming biological barriers to drug delivery, Elvin Blanco et al, Copyright 2015.[10] **Figure 4:** Reprinted by permission under the terms of the Creative Commons CC-BY licence, Cell, Vol. 176, Issue 6, Marcandalli J & Fiala B et al, Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus, Pages 1420-1431.e17, Copyright 2019.[98] **Figure 5:** Reprinted by permission under the terms of the Creative Commons Attribution 4.0 International License, Experimental & Molecular Medicine, Volume 53, Pages 1592–1601, Seo J et al, Fibrinolytic nanocages dissolve clots in the tumor microenvironment, improving the distribution and therapeutic efficacy of anticancer drugs, Copyright 2021.[33] **Figure 6:** Reprinted by permission from American Association for the Advancement of Science, Science, Volume 372, Issue 6537, eabd9994, Divine R et al, Designed proteins assemble antibodies into modular nanocages, Copyright 2021.[21] **Figure 7:** Reprinted by permission from the American Chemical Society, Journal of the American Chemical Society, 140(33):10439–10442, Edwardson TGW et al, Rational Engineering of a Designed Protein Cage for siRNA Delivery, Copyright 2018.[166] **Figure 8:** Reprinted by permission from American Chemical Society, Biomacromolecules, 17(1):12-9, Kim S et al, Double-Chambered Ferritin Platform: Dual-Function Payloads of Cytotoxic Peptides and Fluorescent Protein, Copyright 2016.[83]

References

(1) Mitchell, M. J.; Billingsley, M. M.; Haley, R. M.; Wechsler, M. E.; Peppas, N. A.; Langer, R. Engineering Precision Nanoparticles for Drug Delivery. *Nat. Rev. Drug Discov.* **2021**, *20* (2), 101–124.

(2) Davis, M. E.; Chen, Z. G.; Shin, D. M. Nanoparticle Therapeutics: An Emerging Treatment Modality for Cancer. *Nat. Rev. Drug Discov.* **2008**, *7* (9), 771–782.

(3) Waterhouse, D. N.; Tardi, P. G.; Mayer, L. D.; Bally, M. B. A Comparison of Liposomal Formulations of Doxorubicin with Drug Administered in Free Form: Changing Toxicity Profiles. *Drug Saf.* **2001**, *24* (12), 903–920.

(4) Barenholz, Y. Doxil®--the First FDA-Approved Nano-Drug: Lessons Learned. *J. Control. Release* **2012**, *160* (2), 117–134.

(5) Office of the Commissioner. FDA approves first COVID-19 vaccine https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine (accessed 2021 -11 -02).

(6) Office of the Commissioner. Coronavirus (COVID-19) update: FDA takes key action by approving second COVID-19 vaccine https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-takes-key-action-approving-second-covid-19-vaccine (accessed 2022 -03 -01).

(7) Paunovska, K.; Sago, C. D.; Monaco, C. M.; Hudson, W. H.; Castro, M. G.; Rudoltz, T. G.; Kalathoor, S.; Vanover, D. A.; Santangelo, P. J.; Ahmed et al. A Direct Comparison of in Vitro and in Vivo Nucleic Acid Delivery Mediated by Hundreds of Nanoparticles Reveals a Weak Correlation. *Nano Lett.* **2018**, *18* (3), 2148–2157.

(8) Zhao, Z.; Ukidve, A.; Kim, J.; Mitragotri, S. Targeting Strategies for Tissue-Specific Drug Delivery. *Cell* **2020**, *181* (1), 151–167.

(9) Bae, Y. H.; Park, K. Advanced Drug Delivery 2020 and beyond: Perspectives on the Future. *Adv. Drug Deliv. Rev.* **2020**, *158*, 4–16.

(10) Blanco, E.; Shen, H.; Ferrari, M. Principles of Nanoparticle Design for Overcoming Biological Barriers to Drug Delivery. *Nat. Biotechnol.* **2015**, *33* (9), 941–951.

(11) Reddy, S. T.; Swartz, M. A.; Hubbell, J. A. Targeting Dendritic Cells with Biomaterials: Developing the next Generation of Vaccines. *Trends Immunol.* **2006**, *27* (12), 573–579.

(12) Qi, R.; Wang, Y.; Bruno, P. M.; Xiao, H.; Yu, Y.; Li, T.; Lauffer, S.; Wei, W.; Chen, Q.; Kang, X. et al. Nanoparticle Conjugates of a Highly Potent Toxin Enhance Safety and Circumvent Platinum Resistance in Ovarian Cancer. *Nat. Commun.* **2017**, *8* (1), 2166.

(13) Zhao, Y.; Wang, Y.; Ran, F.; Cui, Y.; Liu, C.; Zhao, Q.; Gao, Y.; Wang, D.; Wang, S. A Comparison between Sphere and Rod Nanoparticles Regarding Their in Vivo Biological Behavior and Pharmacokinetics. *Sci. Rep.* **2017**, *7* (1), 4131.

(14) Gratton, S. E. A.; Ropp, P. A.; Pohlhaus, P. D.; Luft, J. C.; Madden, V. J.; Napier, M. E.; DeSimone, J. M. The Effect of Particle Design on Cellular Internalization Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (33), 11613–11618.

(15) Jiang, W.; Kim, B. Y. S.; Rutka, J. T.; Chan, W. C. W. Nanoparticle-Mediated Cellular Response Is Size-Dependent. *Nat. Nanotechnol.* **2008**, *3* (3), 145–150.

(16) Madathiparambil Visalakshan, R.; González García, L. E.; Benzigar, M. R.; Ghazaryan, A.; Simon, J.; Mierczynska-Vasilev, A.; Michl, T. D.; Vinu, A.; Mailänder, V.; Morsbach, S.; Landfester, K.; Vasilev, K. The Influence of Nanoparticle Shape on Protein Corona Formation. *Small* **2020**, *16* (25), e2000285.

(17) Lazarovits, J.; Chen, Y. Y.; Sykes, E. A.; Chan, W. C. W. Nanoparticle-Blood Interactions: The Implications on Solid Tumour Targeting. *Chem. Commun.* **2015**, *51* (14), 2756–2767.

(18) Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C.; Cascio, D.; Yeates, T. O.; Gonen, T.; King, N. P. et al. Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes. *Science* **2016**, *353* (6297), 389–394.

(19) King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **2012**, *336* (6085), 1171–1174.

(20) Sun, X.; Hong, Y.; Gong, Y.; Zheng, S.; Xie, D. Bioengineered Ferritin Nanocarriers for Cancer Therapy. *Int. J. Mol. Sci.* **2021**, *22* (13). https://doi.org/10.3390/ijms22137023.

(21) Divine, R.; Dang, H. V.; Ueda, G.; Fallas, J. A.; Vulovic, I.; Sheffler, W.; Saini, S.; Zhao, Y. T.; Raj, I. X.; Morawski, P. A. et al. Designed Proteins Assemble Antibodies into Modular Nanocages. *bioRxiv* **2020**. https://doi.org/10.1101/2020.12.01.406611.

(22) Rowinsky, E. K.; Donehower, R. C. Paclitaxel (taxol). *N. Engl. J. Med.* **1995**, *332* (15), 1004–1014.

(23) Pham, D. T.; Tiyaboonchai, W. Fibroin Nanoparticles: A Promising Drug Delivery System. *Drug Deliv.* **2020**, *27* (1), 431–448.

(24) Lee, E. J.; Guenther, C. M.; Suh, J. Adeno-Associated Virus (AAV) Vectors: Rational Design Strategies for Capsid Engineering. *Curr Opin Biomed Eng* **2018**, *7*, 58–63.

(25) Douglas, T.; Young, M. Viruses: Making Friends with Old Foes. *Science* **2006**, *312* (5775), 873–875.

(26) Schwarz, B.; Douglas, T. Development of Virus-like Particles for Diagnostic and Prophylactic Biomedical Applications. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2015**, *7* (5), 722–735.

(27) Steinmetz, N. F.; Lim, S.; Sainsbury, F. Protein Cages and Virus-like Particles: From Fundamental Insight to Biomimetic Therapeutics. *Biomater Sci* **2020**, *8* (10), 2771–2777.

(28) Miao, Y.; Yang, T.; Yang, S.; Yang, M.; Mao, C. Protein Nanoparticles Directed Cancer Imaging and Therapy. *Nano Converg* **2022**, *9* (1), 2.

(29) Habibi, N.; Quevedo, D. F.; Gregory, J. V.; Lahann, J. Emerging Methods in Therapeutics Using Multifunctional Nanoparticles. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2020**, *12* (4), e1625.

(30) Wagner, H. J.; Weber, W.; Fussenegger, M. Synthetic Biology: Emerging Concepts to Design and Advance Adeno-Associated Viral Vectors for Gene Therapy. *Adv. Sci.* **2021**, *8* (9), 2004018.

(31) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537* (7620), 320–327.

(32) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A. et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589.

(33) Seo, J.; Do Yoo, J.; Kim, M.; Shim, G.; Oh, Y.-K.; Park, R.-W.; Lee, B.; Kim, I.-S.; Kim, S. Fibrinolytic Nanocages Dissolve Clots in the Tumor Microenvironment, Improving the Distribution and Therapeutic Efficacy of Anticancer Drugs. *Exp. Mol. Med.* **2021**. https://doi.org/10.1038/s12276-021-00688-7.

(34) Langan, R. A.; Boyken, S. E.; Ng, A. H.; Samson, J. A.; Dods, G.; Westbrook, A. M.; Nguyen, T. H.; Lajoie, M. J.; Chen, Z.; Berger, S. et al. De Novo Design of Bioactive Protein Switches. *Nature* **2019**, *572* (7768), 205–210.

(35) Lajoie, M. J.; Boyken, S. E.; Salter, A. I.; Bruffey, J.; Rajan, A.; Langan, R. A.; Olshefsky, A.; Muhunthan, V.; Bick, M. J.; Gewe, M. et al. Designed Protein Logic to Target Cells with Precise Combinations of Surface Antigens. *Science* **2020**, *369* (6511), 1637–1643.

(36) Boyken, S. E.; Benhaim, M. A.; Busch, F.; Jia, M.; Bick, M. J.; Choi, H.; Klima, J. C.; Chen, Z.; Walkey, C.; Mileant, A. et al. De Novo Design of Tunable, pH-Driven Conformational Changes. *Science* **2019**, *364* (6441), 658–664.

(37) Melnyk, T.; Đorđević, S.; Conejos-Sánchez, I.; Vicent, M. J. Therapeutic Potential of Polypeptide-Based Conjugates: Rational Design and Analytical Tools That Can Boost Clinical Translation. *Adv. Drug Deliv. Rev.* **2020**, *160*, 136–169.

(38) Yoo, J. D.; Bae, S. M.; Seo, J.; Jeon, I. S.; Vadevoo, S. M. P.; Kim, S.-Y.; Kim, I.-S.; Lee, B.; Kim, S. Designed Ferritin Nanocages Displaying Trimeric TRAIL and Tumor-Targeting Peptides Confer Superior Anti-Tumor Efficacy. *Sci. Rep.* **2020**, *10* (1), 19997.

(39) Rome, L. H.; Kickhoefer, V. A. Development of the Vault Particle as a Platform Technology. *ACS Nano* **2013**, *7* (2), 889–902.

(40) Yeates, T. O.; Liu, Y.; Laniado, J. The Design of Symmetric Protein Nanomaterials Comes of Age in Theory and Practice. *Curr. Opin. Struct. Biol.* **2016**, *39*, 134–143.

(41) Van de Steen, A.; Khalife, R.; Colant, N.; Mustafa Khan, H.; Deveikis, M.; Charalambous, S.; Robinson, C. M.; Dabas, R.; Esteban Serna, S.; Catana, D. A. et al. Bioengineering Bacterial Encapsulin Nanocompartments as Targeted Drug Delivery System. *Synth Syst Biotechnol* **2021**, *6* (3), 231–241.

(42) Azuma, Y.; Edwardson, T. G. W.; Hilvert, D. Tailoring Lumazine Synthase Assemblies for Bionanotechnology. *Chem. Soc. Rev.* **2018**, *47* (10), 3543–3557.

(43) Khmelinskaia, A.; Wargacki, A.; King, N. P. Structure-Based Design of Novel Polyhedral Protein Nanomaterials. *Curr. Opin. Microbiol.* **2021**, *61*, 51–57.

(44) Malay, A. D.; Miyazaki, N.; Biela, A.; Chakraborti, S.; Majsterkiewicz, K.; Stupka, I.; Kaplan, C. S.; Kowalczyk, A.; Piette, B. M. A. G.; Hochberg, G. K. A. et al. An Ultra-Stable Gold-Coordinated Protein Cage Displaying Reversible Assembly. *Nature* **2019**, *569* (7756), 438–442.

(45) Churchfield, L. A.; Tezcan, F. A. Design and Construction of Functional Supramolecular Metalloprotein Assemblies. *Acc. Chem. Res.* **2019**, *52* (2), 345–355.

(46) Ballister, E. R.; Lai, A. H.; Zuckermann, R. N.; Cheng, Y.; Mougous, J. D. In Vitro Self-Assembly of

Tailorable Nanotubes from a Simple Protein Building Block. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (10), 3733–3738.

(47) Laniado, J.; Meador, K.; Yeates, T. O. A Fragment-Based Protein Interface Design Algorithm for Symmetric Assemblies. *Protein Eng. Des. Sel.* **2021**, *34*. https://doi.org/10.1093/protein/gzab008.

(48) Padilla, J. E.; Colovos, C.; Yeates, T. O. Nanohedra: Using Symmetry to Design Self Assembling Protein Cages, Layers, Crystals, and Filaments. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (5), 2217–2221.

(49) Cannon, K. A.; Nguyen, V. N.; Morgan, C.; Yeates, T. O. Design and Characterization of an Icosahedral Protein Cage Formed by a Double-Fusion Protein Containing Three Distinct Symmetry Elements. *ACS Synth. Biol.* **2020**, *9* (3), 517–524.

(50) Vulovic, I.; Yao, Q.; Park, Y.-J.; Courbet, A.; Norris, A.; Busch, F.; Sahasrabuddhe, A.; Merten, H.; Sahtoe, D. D.; Ueda, G. et al. Generation of Ordered Protein Assemblies Using Rigid Three-Body Fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (23). https://doi.org/10.1073/pnas.2015037118.

(51) Ueda, G.; Antanasijevic, A.; Fallas, J. A.; Sheffler, W.; Copps, J.; Ellis, D.; Hutchinson, G. B.; Moyer, A.; Yasmeen, A.; Tsybovsky, Y. et al. Tailored Design of Protein Nanoparticle Scaffolds for Multivalent Presentation of Viral Glycoprotein Antigens. *Elife* **2020**, *9*. https://doi.org/10.7554/eLife.57659.

(52) Anishchenko, I.; Pellock, S. J.; Chidyausiku, T. M.; Ramelot, T. A.; Ovchinnikov, S.; Hao, J.; Bafna, K.; Norn, C.; Kang, A.; Bera, A. K. et al. De Novo Protein Design by Deep Network Hallucination. *Nature* **2021**, *600* (7889), 547–552.

(53) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat. Rev. Genet.* **2015**, *16* (7), 379–394.

(54) Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D. Directed Evolution of a Protein Container. *Science* **2011**, *331* (6017), 589–592.

(55) Butterfield, G. L.; Lajoie, M. J.; Gustafson, H. H.; Sellers, D. L.; Nattermann, U.; Ellis, D.; Bale, J. B.; Ke, S.; Lenz, G. H.; Yehdego, A. et al. Evolution of a Designed Protein Assembly Encapsulating Its Own RNA Genome. *Nature* **2017**, *552* (7685), 415–420.

(56) Tetter, S.; Terasaka, N.; Steinauer, A.; Bingham, R. J.; Clark, S.; Scott, A. J. P.; Patel, N.; Leibundgut, M.; Wroblewski, E.; Ban, N. et al. Evolution of a Virus-like Architecture and Packaging Mechanism in a Repurposed Bacterial Protein. *Science* **2021**, *372* (6547), 1220–1224.

(57) Terasaka, N.; Azuma, Y.; Hilvert, D. Laboratory Evolution of Virus-like Nucleocapsids from Nonviral Protein Cages. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (21), 5432–5437.

(58) Ojala, D. S.; Sun, S.; Santiago-Ortiz, J. L.; Shapiro, M. G.; Romero, P. A.; Schaffer, D. V. In Vivo Selection of a Computationally Designed SCHEMA AAV Library Yields a Novel Variant for Infection of Adult Neural Stem Cells in the SVZ. *Mol. Ther.* **2018**, *26* (1), 304–319.

(59) Deverman, B. E.; Pravdo, P. L.; Simpson, B. P.; Kumar, S. R.; Chan, K. Y.; Banerjee, A.; Wu, W.-L.; Yang, B.; Huber, N.; Pasca, S. P.; Gradinaru, V. Cre-Dependent Selection Yields AAV Variants for Widespread Gene Transfer to the Adult Brain. *Nat. Biotechnol.* **2016**, *34* (2), 204–209.

(60) Gustafson, H. H.; Olshefsky, A.; Sylvestre, M.; Sellers, D. L.; Pun, S. H. Current State of in Vivo Panning Technologies: Designing Specificity and Affinity into the Future of Drug Targeting. *Adv.*

Drug Deliv. Rev. **2018**, *130*, 39–49.

(61) Tabebordbar, M.; Lagerborg, K. A.; Stanton, A.; King, E. M.; Ye, S.; Tellez, L.; Krunnfusz, A.; Tavakoli, S.; Widrick, J. J.; Messemer, K. A. et al. Directed Evolution of a Family of AAV Capsid Variants Enabling Potent Muscle-Directed Gene Delivery across Species. *Cell* **2021**, *184* (19), 4919–4938.e22.

(62) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* **2021**, *30* (1), 70–82.

(63) Lawson, D. M.; Artymiuk, P. J.; Yewdall, S. J.; Smith, J. M.; Livingstone, J. C.; Treffry, A.; Luzzago, A.; Levi, S.; Arosio, P.; Cesareni, G. Solving the Structure of Human H Ferritin by Genetically Engineering Intermolecular Crystal Contacts. *Nature* **1991**, *349* (6309), 541–544.

(64) Sutter, M.; Boehringer, D.; Gutmann, S.; Günther, S.; Prangishvili, D.; Loessner, M. J.; Stetter, K. O.; Weber-Ban, E.; Ban, N. Structural Basis of Enzyme Encapsulation into a Bacterial Nanocompartment. *Nat. Struct. Mol. Biol.* **2008**, *15* (9), 939–947.

(65) Zhang, X.; Meining, W.; Fischer, M.; Bacher, A.; Ladenstein, R. X-Ray Structure Analysis and Crystallographic Refinement of Lumazine Synthase from the Hyperthermophile Aquifex Aeolicus at 1.6 A Resolution: Determinants of Thermostability Revealed from Structural Comparisons. *J. Mol. Biol.* **2001**, *306* (5), 1099–1114.

(66) Yeates, T. O. Geometric Principles for Designing Highly Symmetric Self-Assembling Protein Nanomaterials. *Annu. Rev. Biophys.* **2017**, *46*, 23–42.

(67) Laniado, J.; Yeates, T. O. A Complete Rule Set for Designing Symmetry Combination Materials from Protein Molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (50), 31817–31823.

(68) Ladenstein, R.; Meyer, B.; Huber, R.; Labischinski, H.; Bartels, K.; Bartunik, H. D.; Bachmann, L.; Ludwig, H. C.; Bacher, A. Heavy Riboflavin Synthase from Bacillus Subtilis. Particle Dimensions, Crystal Packing and Molecular Symmetry. *J. Mol. Biol.* **1986**, *187* (1), 87–100.

(69) Wukovitz, S. W.; Yeates, T. O. Why Protein Crystals Favour Some Space-Groups over Others. *Nat. Struct. Biol.* **1995**, *2* (12), 1062–1067.

(70) Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C.; Huang, P.-S.; Ravichandran, R. et al. Corrigendum: Design of a Hyperstable 60-Subunit Protein Icosahedron. *Nature* **2016**, *540* (7631), 150.

(71) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D. Accurate Design of Co-Assembling Multi-Component Protein Nanomaterials. *Nature* **2014**, *510* (7503), 103–108.

(72) Wargacki, A. J.; Wörner, T. P.; van de Waterbeemd, M.; Ellis, D.; Heck, A. J. R.; King, N. P. Complete and Cooperative in Vitro Assembly of Computationally Designed Self-Assembling Protein Nanomaterials. *Nat. Commun.* **2021**, *12* (1), 883.

(73) Zlotnick, A. Are Weak Protein-Protein Interactions the General Rule in Capsid Assembly? *Virology* **2003**, *315* (2), 269–274.

(74) Deeds, E. J.; Bachman, J. A.; Fontana, W. Optimizing Ring Assembly Reveals the Strength of Weak Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (7), 2348–2353.

(75) Hagan, M. F.; Chandler, D. Dynamic Pathways for Viral Capsid Assembly. *Biophys. J.* **2006**, *91* (1), 42–54.

(76) Shanbhag, B. K.; Liu, C.; Haritos, V. S.; He, L. Understanding the Interplay between Self-Assembling Peptides and Solution Ions for Tunable Protein Nanoparticle Formation. *ACS Nano* **2018**, *12* (7), 6956–6967.

(77) Kim, M.; Rho, Y.; Jin, K. S.; Ahn, B.; Jung, S.; Kim, H.; Ree, M. pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly. *Biomacromolecules* **2011**, *12* (5), 1629–1640.

(78) Sharma, J.; Uchida, M.; Miettinen, H. M.; Douglas, T. Modular Interior Loading and Exterior Decoration of a Virus-like Particle. *Nanoscale* **2017**, *9* (29), 10420–10430.

(79) McCoy, K.; Selivanovitch, E.; Luque, D.; Lee, B.; Edwards, E.; Castón, J. R.; Douglas, T. Cargo Retention inside P22 Virus-Like Particles. *Biomacromolecules* **2018**, *19* (9), 3738–3746.

(80) Hermanson, G. T. *Bioconjugate Techniques*; Academic Press, 2013.

(81) Drachman, J. G.; Senter, P. D. Antibody-Drug Conjugates: The Chemistry behind Empowering Antibodies to Fight Cancer. *Hematology Am. Soc. Hematol. Educ. Program* **2013**, *2013*, 306–310.

(82) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. Copper-Free Click Chemistry for Dynamic in Vivo Imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (43), 16793–16797.

(83) Kim, S.; Kim, G. S.; Seo, J.; Gowri Rangaswamy, G.; So, I.-S.; Park, R.-W.; Lee, B.-H.; Kim, I.-S. Double-Chambered Ferritin Platform: Dual-Function Payloads of Cytotoxic Peptides and Fluorescent Protein. *Biomacromolecules* **2016**, *17* (1), 12–19.

(84) Wang, Y.-H.; Jian, M.-L.; Chen, P.-J.; Tsou, J.-C.; Truong, L. P.; Wang, Y.-S. Ferritin Conjugates With Multiple Clickable Amino Acids Encoded by C-Terminal Engineered Pyrrolysyl-tRNA Synthetase. *Front Chem* **2021**, *9*, 779976.

(85) Choi, K.-M.; Choi, S.-H.; Jeon, H.; Kim, I.-S.; Ahn, H. J. Chimeric Capsid Protein as a Nanocarrier for siRNA Delivery: Stability and Cellular Uptake of Encapsulated siRNA. *ACS Nano* **2011**, *5* (11), 8690–8699.

(86) Tokatlian, T.; Read, B. J.; Jones, C. A.; Kulp, D. W.; Menis, S.; Chang, J. Y. H.; Steichen, J. M.; Kumari, S.; Allen, J. D.; Dane, E. L. et al. Innate Immune Recognition of Glycans Targets HIV Nanoparticle Immunogens to Germinal Centers. *Science* **2019**, *363* (6427), 649–654.

(87) Spicer, C. D.; Jumeaux, C.; Gupta, B.; Stevens, M. M. Peptide and Protein Nanoparticle Conjugates: Versatile Platforms for Biomedical Applications. *Chem. Soc. Rev.* **2018**, *47* (10), 3574–3620.

(88) Moon, H.; Lee, J.; Min, J.; Kang, S. Developing Genetically Engineered Encapsulin Protein Cage Nanoparticles as a Targeted Delivery Nanoplatform. *Biomacromolecules* **2014**, *15* (10), 3794–3801.

(89) Choi, H.; Eom, S.; Kim, H.-U.; Bae, Y.; Jung, H. S.; Kang, S. Load and Display: Engineering Encapsulin as a Modular Nanoplatform for Protein-Cargo Encapsulation and Protein-Ligand Decoration Using Split Intein and SpyTag/SpyCatcher. *Biomacromolecules* **2021**, *22* (7), 3028–3039.

(90) Wang, L.; Xing, D.; Le Van, A.; Jerse, A. E.; Wang, S. Structure-Based Design of Ferritin Nanoparticle Immunogens Displaying Antigenic Loops of Neisseria Gonorrhoeae. *FEBS Open Bio*

2017, *7* (8), 1196–1207.

(91) Moon, H.; Lee, J.; Kim, H.; Heo, S.; Min, J.; Kang, S. Genetically Engineering Encapsulin Protein Cage Nanoparticle as a SCC-7 Cell Targeting Optical Nanoprobe. *Biomater Res* **2014**, *18*, 21.

(92) Michel-Souzy, S.; Hamelmann, N. M.; Zarzuela-Pura, S.; Paulusse, J. M. J.; Cornelissen, J. J. L. M. Introduction of Surface Loops as a Tool for Encapsulin Functionalization. *Biomacromolecules* **2021**, *22* (12), 5234–5242.

(93) Read, B. J.; Won, L.; Kraft, J. C.; Sappington, I.; Aung, A.; Wu, S.; Bals, J.; Chen, C.; Lee, K. K.; Lingwood, D. et al. Mannose-Binding Lectin and Complement Mediate Follicular Localization and Enhanced Immunogenicity of Diverse Protein Nanoparticle Immunogens. *Cell Rep.* **2022**, *38* (2), 110217.

(94) Suk, J. S.; Xu, Q.; Kim, N.; Hanes, J.; Ensign, L. M. PEGylation as a Strategy for Improving Nanoparticle-Based Drug and Gene Delivery. *Adv. Drug Deliv. Rev.* **2016**, *99* (Pt A), 28–51.

(95) Zahnd, C.; Wyler, E.; Schwenk, J. M.; Steiner, D.; Lawrence, M. C.; McKern, N. M.; Pecorari, F.; Ward, C. W.; Joos, T. O.; Plückthun, A. A Designed Ankyrin Repeat Protein Evolved to Picomolar Affinity to Her2. *J. Mol. Biol.* **2007**, *369* (4), 1015–1028.

(96) Boyoglu-Barnum, S.; Ellis, D.; Gillespie, R. A.; Hutchinson, G. B.; Park, Y.-J.; Moin, S. M.; Acton, O. J.; Ravichandran, R.; Murphy, M.; Pettie, D. et al. Quadrivalent Influenza Nanoparticle Vaccines Induce Broad Protection. *Nature* **2021**, *592* (7855), 623–628.

(97) Kanekiyo, M.; Wei, C.-J.; Yassine, H. M.; McTamney, P. M.; Boyington, J. C.; Whittle, J. R. R.; Rao, S. S.; Kong, W.-P.; Wang, L.; Nabel, G. J. Self-Assembling Influenza Nanoparticle Vaccines Elicit Broadly Neutralizing H1N1 Antibodies. *Nature* **2013**, *499* (7456), 102–106.

(98) Marcandalli, J.; Fiala, B.; Ols, S.; Perotti, M.; de van der Schueren, W.; Snijder, J.; Hodge, E.; Benhaim, M.; Ravichandran, R.; Carter, L. et al. Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **2019**, *176* (6), 1420–1431.e17.

(99) Rahikainen, R.; Rijal, P.; Tan, T. K.; Wu, H.-J.; Andersson, A.-M. C.; Barrett, J. R.; Bowden, T. A.; Draper, S. J.; Townsend, A. R.; Howarth, M. Overcoming Symmetry Mismatch in Vaccine Nanoassembly through Spontaneous Amidation. *Angew. Chem. Int. Ed Engl.* **2021**, *60* (1), 321–330.

(100) Palmer, E. UPDATED: J&J's Doxil shortage to last until at least end of 2014 https://www.fiercepharma.com/m-a/updated-j-j-s-doxil-shortage-to-last-until-at-least-end-of-2014 (accessed 2022 -01 -19).

(101) Sethuraman, N.; Stadheim, T. A. Challenges in Therapeutic Glycoprotein Production. *Curr. Opin. Biotechnol.* **2006**, *17* (4), 341–346.

(102) Kelley, B. Industrialization of mAb Production Technology: The Bioprocessing Industry at a Crossroads. *MAbs* **2009**, *1* (5), 443–452.

(103) Kushnir, N.; Streatfield, S. J.; Yusibov, V. Virus-like Particles as a Highly Efficient Vaccine Platform: Diversity of Targets and Production Systems and Advances in Clinical Development. *Vaccine* **2012**, *31* (1), 58–83.

(104) Gogtay, N.; Thatte, U.; Kshirsagar, N.; Leav, B.; Molrine, D.; Cheslock, P.; Kapre, S. V.; Kulkarni, P. S.; SII RMab author group. Safety and Pharmacokinetics of a Human Monoclonal

Antibody to Rabies Virus: A Randomized, Dose-Escalation Phase 1 Study in Adults. *Vaccine* **2012**, *30* (50), 7315–7320.

(105) Wu, L.; Su, S.; Liu, F.; Xu, T.; Wang, X.; Huang, Y.; Sun, X.; Ge, X.; Chen, T.; Liu, H. et al. Removal of the Tag from His-Tagged ILYd4, a Human CD59 Inhibitor, Significantly Improves Its Physical Properties and Its Activity. *Curr. Pharm. Des.* **2012**, *18* (27), 4187–4196.

(106) Puetz, J.; Wurm, F. M. Recombinant Proteins for Industrial versus Pharmaceutical Purposes: A Review of Process and Pricing. *Processes* **2019**, *7* (8), 476.

(107) Jiang, H.; Horwitz, A. A.; Wright, C.; Tai, A.; Znameroski, E. A.; Tsegaye, Y.; Warbington, H.; Bower, B. S.; Alves, C.; Co, C. et al. Challenging the Workhorse: Comparative Analysis of Eukaryotic Micro-Organisms for Expressing Monoclonal Antibodies. *Biotechnol. Bioeng.* **2019**, *116* (6), 1449–1462.

(108) Brady, J. R.; Love, J. C. Alternative Hosts as the Missing Link for Equitable Therapeutic Protein Production. *Nat. Biotechnol.* **2021**, *39* (4), 404–407.

(109) Matthews, C. B.; Kuo, A.; Love, K. R.; Love, J. C. Development of a General Defined Medium for Pichia Pastoris. *Biotechnol. Bioeng.* **2018**, *115* (1), 103–113.

(110) Cohen, A. A.; Gnanapragasam, P. N. P.; Lee, Y. E.; Hoffman, P. R.; Ou, S.; Kakutani, L. M.; Keeffe, J. R.; Wu, H.-J.; Howarth, M.; West, A. P. et al. Mosaic Nanoparticles Elicit Cross-Reactive Immune Responses to Zoonotic Coronaviruses in Mice. *Science* **2021**, *371* (6530), 735–741.

(111) Wang, D.; Tai, P. W. L.; Gao, G. Adeno-Associated Virus Vector as a Platform for Gene Therapy Delivery. *Nat. Rev. Drug Discov.* **2019**, *18* (5), 358–378.

(112) Bulcha, J. T.; Wang, Y.; Ma, H.; Tai, P. W. L.; Gao, G. Viral Vector Platforms within the Gene Therapy Landscape. *Signal Transduct Target Ther* **2021**, *6* (1), 53.

(113) Pardi, N.; Hogan, M. J.; Weissman, D. Recent Advances in mRNA Vaccine Technology. *Curr. Opin. Immunol.* **2020**, *65*, 14–20.

(114) Xu, Z.; Chokkalingam, N.; Tello-Ruiz, E.; Wise, M. C.; Bah, M. A.; Walker, S.; Tursi, N. J.; Fisher, P. D.; Schultheis, K.; Broderick, K. E. et al. A DNA-Launched Nanoparticle Vaccine Elicits CD8+ T-Cell Immunity to Promote In Vivo Tumor Control. *Cancer Immunol Res* **2020**, *8* (11), 1354–1364.

(115) Melo, M.; Porter, E.; Zhang, Y.; Silva, M.; Li, N.; Dobosh, B.; Liguori, A.; Skog, P.; Landais, E.; Menis, S. et al. Immunogenicity of RNA Replicons Encoding HIV Env Immunogens Designed for Self-Assembly into Nanoparticles. *Mol. Ther.* **2019**, *27* (12), 2080–2090.

(116) Hoshyar, N.; Gray, S.; Han, H.; Bao, G. The Effect of Nanoparticle Size on in Vivo Pharmacokinetics and Cellular Interaction. *Nanomedicine* **2016**, *11* (6), 673–692.

(117) De Groot, A. S.; Scott, D. W. Immunogenicity of Protein Therapeutics. *Trends Immunol.* **2007**, *28* (11), 482–490.

(118) Gustafson, H. H.; Holt-Casper, D.; Grainger, D. W.; Ghandehari, H. Nanoparticle Uptake: The Phagocyte Problem. *Nano Today* **2015**, *10* (4), 487–510.

(119) Fries, C. N.; Curvino, E. J.; Chen, J.-L.; Permar, S. R.; Fouda, G. G.; Collier, J. H. Advances in Nanomaterial Vaccine Strategies to Address Infectious Diseases Impacting Global Health. *Nat.*

Nanotechnol. **2021**, *16* (4), 1–14.

(120) Rodrigues, M. Q.; Alves, P. M.; Roldão, A. Functionalizing Ferritin Nanoparticles for Vaccine Development. *Pharmaceutics* **2021**, *13* (10). https://doi.org/10.3390/pharmaceutics13101621.

(121) Kanekiyo, M.; Bu, W.; Joyce, M. G.; Meng, G.; Whittle, J. R. R.; Baxa, U.; Yamamoto, T.; Narpala, S.; Todd, J.-P.; Rao, S. S. et al. Rational Design of an Epstein-Barr Virus Vaccine Targeting the Receptor-Binding Site. *Cell* **2015**, *162* (5), 1090–1100.

(122) Walls, A. C.; Fiala, B.; Schäfer, A.; Wrenn, S.; Pham, M. N.; Murphy, M.; Tse, L. V.; Shehata, L.; O'Connor, M. A.; Chen, C. et al. Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell* **2020**, *183* (5), 1367–1382.e17.

(123) Antanasijevic, A.; Ueda, G.; Brouwer, P. J. M.; Copps, J.; Huang, D.; Allen, J. D.; Cottrell, C. A.; Yasmeen, A.; Sewall, L. M.; Bontjer, I. et al. Structural and Functional Evaluation of de Novo-Designed, Two-Component Nanoparticle Carriers for HIV Env Trimer Immunogens. *PLoS Pathog.* **2020**, *16* (8), e1008665.

(124) Duan, H.; Chen, X.; Boyington, J. C.; Cheng, C.; Zhang, Y.; Jafari, A. J.; Stephens, T.; Tsybovsky, Y.; Kalyuzhniy, O.; Zhao, P. et al. Glycan Masking Focuses Immune Responses to the HIV-1 CD4-Binding Site and Enhances Elicitation of VRC01-Class Precursor Antibodies. *Immunity* **2018**, *49* (2), 301–311.e5.

(125) Lagoutte, P.; Mignon, C.; Stadthagen, G.; Potisopon, S.; Donnat, S.; Mast, J.; Lugari, A.; Werle, B. Simultaneous Surface Display and Cargo Loading of Encapsulin Nanocompartments and Their Use for Rational Vaccine Design. *Vaccine* **2018**, *36* (25), 3622–3628.

(126) Choi, B.; Moon, H.; Hong, S. J.; Shin, C.; Do, Y.; Ryu, S.; Kang, S. Effective Delivery of Antigen-Encapsulin Nanoparticle Fusions to Dendritic Cells Leads to Antigen-Specific Cytotoxic T Cell Activation and Tumor Rejection. *ACS Nano* **2016**, *10* (8), 7339–7350.

(127) Bachmann, M. F.; Jennings, G. T. Vaccine Delivery: A Matter of Size, Geometry, Kinetics and Molecular Patterns. *Nat. Rev. Immunol.* **2010**, *10* (11), 787–796.

(128) Kato, Y.; Abbott, R. K.; Freeman, B. L.; Haupt, S.; Groschel, B.; Silva, M.; Menis, S.; Irvine, D. J.; Schief, W. R.; Crotty, S. Multifaceted Effects of Antigen Valency on B Cell Response Composition and Differentiation In Vivo. *Immunity* **2020**, *53* (3), 548–563.e8.

(129) McLellan, J. S.; Chen, M.; Joyce, M. G.; Sastry, M.; Stewart-Jones, G. B. E.; Yang, Y.; Zhang, B.; Chen, L.; Srivatsan, S.; Zheng, A. et al. Structure-Based Design of a Fusion Glycoprotein Vaccine for Respiratory Syncytial Virus. *Science* **2013**, *342* (6158), 592–598.

(130) Mukhamedova, M.; Wrapp, D.; Shen, C.-H.; Gilman, M. S. A.; Ruckwardt, T. J.; Schramm, C. A.; Ault, L.; Chang, L.; Derrien-Colemyn, A.; Lucas, S. A. M. et al. Vaccination with Prefusion-Stabilized Respiratory Syncytial Virus Fusion Protein Induces Genetically and Antigenically Diverse Antibody Responses. *Immunity* **2021**, *54* (4), 769–780.e6.

(131) Pauthner, M.; Havenar-Daughton, C.; Sok, D.; Nkolola, J. P.; Bastidas, R.; Boopathy, A. V.; Carnathan, D. G.; Chandrashekar, A.; Cirelli, K. M.; Cottrell, C. A. et al. Elicitation of Robust Tier 2 Neutralizing Antibody Responses in Nonhuman Primates by HIV Envelope Trimer Immunization Using Optimized Approaches. *Immunity* **2017**, *46* (6), 1073–1088.e6.

(132) Cirelli, K. M.; Carnathan, D. G.; Nogal, B.; Martin, J. T.; Rodriguez, O. L.; Upadhyay, A. A.;

Enemuo, C. A.; Gebru, E. H.; Choe, Y.; Viviano, F. et al. Slow Delivery Immunization Enhances HIV Neutralizing Antibody and Germinal Center Responses via Modulation of Immunodominance. *Cell* **2020**, *180* (1), 206.

(133) Sesterhenn, F.; Yang, C.; Bonet, J.; Cramer, J. T.; Wen, X.; Wang, Y.; Chiang, C.-I.; Abriata, L. A.; Kucharska, I.; Castoro, G. et al. De Novo Protein Design Enables the Precise Induction of RSV-Neutralizing Antibodies. *Science* **2020**, *368* (6492). https://doi.org/10.1126/science.aay5051.

(134) Zhao, Q.; Potter, C. S.; Carragher, B.; Lander, G.; Sworen, J.; Towne, V.; Abraham, D.; Duncan, P.; Washabaugh, M. W.; Sitrin, R. D. Characterization of Virus-like Particles in GARDASIL® by Cryo Transmission Electron Microscopy. *Hum. Vaccin. Immunother.* **2014**, *10* (3), 734–739.

(135) Joyce, M. G.; Wheatley, A. K.; Thomas, P. V.; Chuang, G.-Y.; Soto, C.; Bailer, R. T.; Druz, A.; Georgiev, I. S.; Gillespie, R. A.; Kanekiyo, M. et al. Vaccine-Induced Antibodies That Neutralize Group 1 and Group 2 Influenza A Viruses. *Cell* **2016**, *166* (3), 609–623.

(136) Roldão, A.; Mellado, M. C. M.; Castilho, L. R.; Carrondo, M. J. T.; Alves, P. M. Virus-like Particles in Vaccine Development. *Expert Rev. Vaccines* **2010**, *9* (10), 1149–1176.

(137) Jeong, H.; Seong, B. L. Exploiting Virus-like Particles as Innovative Vaccines against Emerging Viral Infections. *J. Microbiol.* **2017**, *55* (3), 220–230.

(138) Sirleaf, E. J.; Clark, H. Report of the Independent Panel for Pandemic Preparedness and Response: Making COVID-19 the Last Pandemic. *Lancet* **2021**, *398* (10295), 101–103.

(139) Abbott, R. K.; Lee, J. H.; Menis, S.; Skog, P.; Rossi, M.; Ota, T.; Kulp, D. W.; Bhullar, D.; Kalyuzhniy, O.; Havenar-Daughton, C. et al. Precursor Frequency and Affinity Determine B Cell Competitive Fitness in Germinal Centers, Tested with Germline-Targeting HIV Vaccine Immunogens. *Immunity* **2018**, *48* (1), 133–146.e6.

(140) Jardine, J.; Julien, J.-P.; Menis, S.; Ota, T.; Kalyuzhniy, O.; McGuire, A.; Sok, D.; Huang, P.-S.; MacPherson, S.; Jones, M. et al. Rational HIV Immunogen Design to Target Specific Germline B Cell Receptors. *Science* **2013**, *340* (6133), 711–716.

(141) Jardine, J. G.; Kulp, D. W.; Havenar-Daughton, C.; Sarkar, A.; Briney, B.; Sok, D.; Sesterhenn, F.; Ereño-Orbea, J.; Kalyuzhniy, O.; Deresa, I. et al. HIV-1 Broadly Neutralizing Antibody Precursor B Cells Revealed by Germline-Targeting Immunogen. *Science* **2016**, *351* (6280), 1458–1463.

(142) Mu, Z.; Wiehe, K.; Saunders, K. O.; Henderson, R.; Cain, D. W.; Parks, R.; Martik, D.; Mansouri, K.; Edwards, R. J.; Newman, A. et al. Ability of Nucleoside-Modified mRNA to Encode HIV-1 Envelope Trimer Nanoparticles. *bioRxiv* **2021**. https://doi.org/10.1101/2021.08.09.455714.

(143) Wise, M. C.; Xu, Z.; Tello-Ruiz, E.; Beck, C.; Trautz, A.; Patel, A.; Elliott, S. T.; Chokkalingam, N.; Kim, S.; Kerkau, M. G. et al. In Vivo Delivery of Synthetic DNA-Encoded Antibodies Induces Broad HIV-1-Neutralizing Activity. *J. Clin. Invest.* **2020**, *130* (2), 827–837.

(144) Kanekiyo, M.; Joyce, M. G.; Gillespie, R. A.; Gallagher, J. R.; Andrews, S. F.; Yassine, H. M.; Wheatley, A. K.; Fisher, B. E.; Ambrozak, D. R.; Creanga, A. et al. Mosaic Nanoparticle Display of Diverse Influenza Virus Hemagglutinins Elicits Broad B Cell Responses. *Nat. Immunol.* **2019**, *20* (3), 362–372.

(145) Walls, A. C.; Miranda, M. C.; Schäfer, A.; Pham, M. N.; Greaney, A.; Arunachalam, P. S.; Navarro, M.-J.; Tortorici, M. A.; Rogers, K.; O'Connor, M. A. et al. Elicitation of Broadly Protective

Sarbecovirus Immunity by Receptor-Binding Domain Nanoparticle Vaccines. *Cell* **2021**, *184* (21), 5432–5447.e16.

(146) Sliepen, K.; van Montfort, T.; Melchers, M.; Isik, G.; Sanders, R. W. Immunosilencing a Highly Immunogenic Protein Trimerization Domain. *J. Biol. Chem.* **2015**, *290* (12), 7436–7442.

(147) Singla, A. K.; Garg, A.; Aggarwal, D. Paclitaxel and Its Formulations. *Int. J. Pharm.* **2002**, *235* (1-2), 179–192.

(148) Horwitz, S.; O'Connor, O. A.; Pro, B.; Illidge, T.; Fanale, M.; Advani, R.; Bartlett, N. L.; Christensen, J. H.; Morschhauser, F.; Domingo-Domenech, E. et al. Brentuximab Vedotin with Chemotherapy for CD30-Positive Peripheral T-Cell Lymphoma (ECHELON-2): A Global, Double-Blind, Randomised, Phase 3 Trial. *Lancet* **2019**, *393* (10168), 229–240.

(149) Hamblett, K. J.; Senter, P. D.; Chace, D. F.; Sun, M. M. C.; Lenox, J.; Cerveny, C. G.; Kissler, K. M.; Bernhardt, S. X.; Kopcha, A. K.; Zabinski, R. F et al. Effects of Drug Loading on the Antitumor Activity of a Monoclonal Antibody Drug Conjugate. *Clin. Cancer Res.* **2004**, *10* (20), 7063–7070.

(150) Khoshnejad, M.; Greineder, C. F.; Pulsipher, K. W.; Villa, C. H.; Altun, B.; Pan, D. C.; Tsourkas, A.; Dmochowski, I. J.; Muzykantov, V. R. Ferritin Nanocages with Biologically Orthogonal Conjugation for Vascular Targeting and Imaging. *Bioconjug. Chem.* **2018**, *29* (4), 1209–1218.

(151) Liang, M.; Fan, K.; Zhou, M.; Duan, D.; Zheng, J.; Yang, D.; Feng, J.; Yan, X. H-Ferritin-Nanocaged Doxorubicin Nanoparticles Specifically Target and Kill Tumors with a Single-Dose Injection. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (41), 14900–14905.

(152) Caldeira, J. C.; Perrine, M.; Pericle, F.; Cavallo, F. Virus-Like Particles as an Immunogenic Platform for Cancer Vaccines. *Viruses* **2020**, *12* (5). https://doi.org/10.3390/v12050488.

(153) Schoonen, L.; van Hest, J. C. M. Functionalization of Protein-Based Nanocages for Drug Delivery Applications. *Nanoscale* **2014**, *6* (13), 7124–7141.

(154) Ghosh, S.; Banerjee, M. A Smart Viral Vector for Targeted Delivery of Hydrophobic Drugs. *Sci. Rep.* **2021**, *11* (1), 7030.

(155) Edwardson, T. G. W.; Tetter, S.; Hilvert, D. Two-Tier Supramolecular Encapsulation of Small Molecules in a Protein Cage. *Nat. Commun.* **2020**, *11* (1), 5410.

(156) van Erp, N. P.; Gelderblom, H.; Guchelaar, H.-J. Clinical Pharmacokinetics of Tyrosine Kinase Inhibitors. *Cancer Treat. Rev.* **2009**, *35* (8), 692–706.

(157) Thakur, A.; Huang, M.; Lum, L. G. Bispecific Antibody Based Therapeutics: Strengths and Challenges. *Blood Rev.* **2018**, *32* (4), 339–347.

(158) Lemke, J.; von Karstedt, S.; Zinngrebe, J.; Walczak, H. Getting TRAIL Back on Track for Cancer Therapy. *Cell Death Differ.* **2014**, *21* (9), 1350–1364.

(159) Jeon, J. O.; Kim, S.; Choi, E.; Shin, K.; Cha, K.; So, I.-S.; Kim, S.-J.; Jun, E.; Kim, D.; Ahn, H. J. et al. Designed Nanocage Displaying Ligand-Specific Peptide Bunches for High Affinity and Biological Activity. *ACS Nano* **2013**, *7* (9), 7462–7471.

(160) Rujas, E.; Kucharska, I.; Tan, Y. Z.; Benlekbir, S.; Cui, H.; Zhao, T.; Wasney, G. A.; Budylowski, P.; Guvenc, F.; Newton, J. C. et al. Multivalency Transforms SARS-CoV-2 Antibodies into Ultrapotent Neutralizers. *Nat. Commun.* **2021**, *12* (1), 3661.

(161) Zhao, Y. T.; Fallas, J. A.; Saini, S.; Ueda, G.; Somasundaram, L.; Zhou, Z.; Xavier Raj, I.; Xu, C.; Carter, L.; Wrenn, S. et al. F-Domain Valency Determines Outcome of Signaling through the Angiopoietin Pathway. *EMBO Rep.* **2021**, *22* (12), e53471.

(162) Kaczmarczyk, S. J.; Sitaraman, K.; Young, H. A.; Hughes, S. H.; Chatterjee, D. K. Protein Delivery Using Engineered Virus-like Particles. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (41), 16998–17003.

(163) Komatsu, Y.; Terasaka, N.; Sakai, K.; Mihara, E.; Wakabayashi, R.; Matsumoto, K.; Hilvert, D.; Takagi, J.; Suga, H. De Novo Peptide Grafting to a Self-Assembling Nanocapsule Yields a Hepatocyte Growth Factor Receptor Agonist. *iScience* **2021**, *24* (11), 103302.

(164) Jensen, M. C.; Riddell, S. R. Designing Chimeric Antigen Receptors to Effectively and Safely Target Tumors. *Curr. Opin. Immunol.* **2015**, *33*, 9–15.

(165) Eyquem, J.; Mansilla-Soto, J.; Giavridis, T.; van der Stegen, S. J. C.; Hamieh, M.; Cunanan, K. M.; Odak, A.; Gönen, M.; Sadelain, M. Targeting a CAR to the TRAC Locus with CRISPR/Cas9 Enhances Tumour Rejection. *Nature* **2017**, *543* (7643), 113–117.

(166) Edwardson, T. G. W.; Mori, T.; Hilvert, D. Rational Engineering of a Designed Protein Cage for siRNA Delivery. *J. Am. Chem. Soc.* **2018**, *140* (33), 10439–10442.

(167) Levasseur, M. D.; Mantri, S.; Hayashi, T.; Reichenbach, M.; Hehn, S.; Waeckerle-Men, Y.; Johansen, P.; Hilvert, D. Cell-Specific Delivery Using an Engineered Protein Nanocage. *ACS Chem. Biol.* **2021**, *16* (5), 838–843.

(168) Cheng, Y.; Yumul, R. C.; Pun, S. H. Virus-Inspired Polymer for Efficient In Vitro and In Vivo Gene Delivery. *Angew. Chem. Int. Ed Engl.* **2016**, *55* (39), 12013–12017.

(169) Yuan, Y.-Y.; Mao, C.-Q.; Du, X.-J.; Du, J.-Z.; Wang, F.; Wang, J. Surface Charge Switchable Nanoparticles Based on Zwitterionic Polymer for Enhanced Drug Delivery to Tumor. *Adv. Mater.* **2012**, *24* (40), 5476–5480.

(170) Banskota, S.; Raguram, A.; Suh, S.; Du, S. W.; Davis, J. R.; Choi, E. H.; Wang, X.; Nielsen, S. C.; Newby, G. A.; Randolph, P. B. et al. Engineered Virus-like Particles for Efficient in Vivo Delivery of Therapeutic Proteins. *Cell* **2022**. https://doi.org/10.1016/j.cell.2021.12.021.

(171) Man, F.; Lammers, T.; T M de Rosales, R. Imaging Nanomedicine-Based Drug Delivery: A Review of Clinical Studies. *Mol. Imaging Biol.* **2018**, *20* (5), 683–695.

(172) Diaz, D.; Vidal, X.; Sunna, A.; Care, A. Bioengineering a Light-Responsive Encapsulin Nanoreactor: A Potential Tool for In Vitro Photodynamic Therapy. *ACS Appl. Mater. Interfaces* **2021**, *13* (7), 7977–7986.

(173) Putri, R. M.; Fredy, J. W.; Cornelissen, J. J. L. M.; Koay, M. S. T.; Katsonis, N. Labelling Bacterial Nanocages with Photo-Switchable Fluorophores. *Chemphyschem* **2016**, *17* (12), 1815–1818.

(174) Farhadi, A.; Ho, G.; Kunth, M.; Ling, B.; Lakshmanan, A.; Lu, G.; Bourdeau, R. W.; Schröder, L.; Shapiro, M. G. Recombinantly Expressed Gas Vesicles as Nanoscale Contrast Agents for Ultrasound and Hyperpolarized MRI. *AIChE J.* **2018**, *64* (8), 2927–2933.

(175) Bennett, K. M.; Shapiro, E. M.; Sotak, C. H.; Koretsky, A. P. Controlled Aggregation of Ferritin to Modulate MRI Relaxivity. *Biophys. J.* **2008**, *95* (1), 342–351.

(176) Kaku, T. S.; Lim, S. Protein Nanoparticles in Molecular, Cellular, and Tissue Imaging. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2021**, *13* (5), e1714.

(177) Geninatti Crich, S.; Bussolati, B.; Tei, L.; Grange, C.; Esposito, G.; Lanzardo, S.; Camussi, G.; Aime, S. Magnetic Resonance Visualization of Tumor Angiogenesis by Targeting Neural Cell Adhesion Molecules with the Highly Sensitive Gadolinium-Loaded Apoferritin Probe. *Cancer Res.* **2006**, *66* (18), 9196–9201.

(178) Elzoghby, A. O.; Hemasa, A. L.; Freag, M. S. Hybrid Protein-Inorganic Nanoparticles: From Tumor-Targeted Drug Delivery to Cancer Imaging. *J. Control. Release* **2016**, *243*, 303–322.

(179) Hwang, M. P.; Lee, J.-W.; Lee, K. E.; Lee, K. H. Think Modular: A Simple Apoferritin-Based Platform for the Multifaceted Detection of Pancreatic Cancer. *ACS Nano* **2013**, *7* (9), 8167–8174.

(180) Mingozzi, F.; High, K. A. Immune Responses to AAV Vectors: Overcoming Barriers to Successful Gene Therapy. *Blood* **2013**, *122* (1), 23–36.

(181) Martino, A. T.; Markusic, D. M. Immune Response Mechanisms against AAV Vectors in Animal Models. *Mol Ther Methods Clin Dev* **2020**, *17*, 198–208.

(182) Rodriguez, P. L.; Harada, T.; Christian, D. A.; Pantano, D. A.; Tsai, R. K.; Discher, D. E. Minimal "Self" Peptides That Inhibit Phagocytic Clearance and Enhance Delivery of Nanoparticles. *Science* **2013**, *339* (6122), 971–975.

(183) Nguyen, B.; Tolia, N. H. Protein-Based Antigen Presentation Platforms for Nanoparticle Vaccines. *NPJ Vaccines* **2021**, *6* (1), 70.

(184) Laurens, M. B. RTS,S/AS01 Vaccine (Mosquirix™): An Overview. *Hum. Vaccin. Immunother.* **2020**, *16* (3), 480–489.

(185) Inkscape Project, 2020. Inkscape, Available at: https://inkscape.org.

Chapter 2. *In vivo* selection of synthetic nucleocapsids for tissue targeting

Audrey Ellen Olshefsky, Halli Benasutti, Meilyn Sylvestre, Gabriel L. Butterfield, Gabriel J. Rocklin, Christian Richardson, Derrick R. Hicks, Marc J. Lajoie, Elizabeth Leaf, Catherine Treichel, Sharon Ke, Gargi Kher, Lauren Carter, Jeffrey S. Chamberlain, David Baker, Neil P. King* and Suzie H. Pun*

*Co-senior authors

Abstract

Controlling the biodistribution of protein- and nanoparticle-based therapeutic formulations remains challenging. *In vivo* library selection is an effective method for identifying constructs that exhibit desired distribution behavior; library variants can be selected based on their ability to localize to the tissue or compartment of interest despite complex physiological challenges. Here, we describe further development of an *in vivo* library selection platform based on self-assembling protein nanoparticles encapsulating their own mRNA genomes (synthetic nucleocapsids or synNCs). We show that libraries composed of mutations on the surface of synNCs are likely insufficient to identify therapeutically relevant biodistribution shifts beyond increased circulation half-life in blood, while libraries of synNCs displaying miniproteins with binder-like properties can successfully be utilized to identify variants that shift accumulation towards lungs, muscles, or tumors in just two rounds of *in vivo* selection. Our approach should contribute to achieving specific tissue homing patterns and identifying targeting ligands for diseases of interest.

Introduction

The emerging drug class of biologics with intracellular targets (e.g., mRNA, siRNA, ribonucleoproteins) is currently limited in clinical translation by a lack of effective systemic delivery systems (*1*, *2*). These biologics often require packaging into delivery systems such as viral vectors or lipid nanoparticles (LNPs) that improve drug stability and mediate cellular entry. However, the default biodistribution of most nanoparticles when delivered systemically skews toward uptake by the mononuclear phagocytic system (MPS) and accumulation in liver and spleen (*3*). *In vivo* library selection is a powerful method to identify delivery vehicles possessing both physicochemical characteristics and molecular ligands that enable organ targeting. For example, *in vivo* library selection has been successfully employed to alter the vascular targeting profile of lipid nanoparticles (LNPs) by identifying specific surface chemistries that enable therapeutically-relevant mRNA delivery to bone marrow, lungs, spleen, liver, and tumor (*4–8*). *In vivo* library selection is also amenable to the discovery of targeting ligand moieties, wherein millions of potential ligands are simultaneously screened for desired biodistribution behaviors (*9*). Unlike *in vitro* library selection, *in vivo* library selection accounts for both known and unknown physiological complexity and can utilize established animal models that recapitulate disease physiology. Although several *in vivo* library selection platforms exist, every platform has limitations. Current state-of-the-art library selection platforms are limited by large particle sizes that prevent tissue penetration (e.g., M13 phage display), low stability (e.g., ribosome display and one-bead one-compound display), incompatibility with *in vivo* selection (one-bead one-compound display), and sensitivity to surface display modifications (e.g., adeno-associated virus display) (*9–11*). Additional library selection platforms that better represent nanoparticle delivery systems and are amenable to *in vivo* selection will further enable the discovery of novel targeting ligands for specific cell and tissue types.

In prior work, our group engineered a virus-sized, self-assembling protein nanoparticle that encapsulates its own mRNA genome, I53-50-v4 (*12*). I53-50-v4 was derived from I53-50, a computationally designed,

two-component self-assembling protein nanoparticle with icosahedral symmetry (*13*). This mRNA-encapsulating version of I53-50, termed a "synthetic nucleocapsid" (synNC), links genotype and phenotype to enable the evolution of new and improved functional properties through diversification and selection. After four stages of evolution, synNC stability in blood was increased by almost 20-fold and *in vivo* circulation half-life was increased from less than 5 minutes to about 4.5 hours. We named the resulting synNC “I53-50-v4” or simply “v4,” for the fourth evolved version of I53-50. However, we did not observe specific or preferential association of I53-50-v4 with any major organs. Samples collected either 5 minutes or 4 hours after I53-50-v4 administration showed the majority of I53-50-v4 mRNA remained in the blood, while we detected an order of magnitude less I53-50-v4 mRNA in the kidneys, lungs, heart, brain, spleen, or liver compared to the blood. This led us to hypothesize that a modified selection strategy that enriches library variants in specific tissues would yield novel targeting moieties. We sought to address two main questions: First, can the same library with which circulation half-life was evolved (*12*) also be used to enrich tissue targeting behavior of the synthetic nucleocapsid itself? Second, can the synthetic nucleocapsids be used in a similar manner as phage display, in which binding domains are displayed on the surface of the particles and tested for accumulation in specific tissues? Our approach is similar to the state-of-the-art M13 phage display platform (*14*). However, since synNCs are an order of magnitude smaller than M13 phage, they might be better able to access different physiological compartments such as tumor microenvironments. Additionally, computationally designed protein nanoparticles can be engineered modularly to display or encapsulate a wide variety of cargoes (*13*, *15*, *16*); and can be manufactured at scale (*17*), so any promising library variant could be directly formulated for applications in therapeutic delivery.

Here, we utilize murine models to evolve two separate synNC libraries towards altered tissue biodistribution. We tested two different libraries in BALB/c mice: a surface library in which point mutations are made on the surface of the synNCs (*12*), and a miniprotein display library in which the synNCs display a semi-random library of designed miniproteins on their surface. We show that specific

tissue targeting is not likely to be achieved by simple synNC surface modifications, while miniproteins can shift synNC biodistribution patterns towards selected tissues, e.g., lung, muscle, or tumor.

Results

Synthetic nucleocapsid libraries can be constructed in a variety of ways. Here, we constructed surface libraries where surface residues were mutated (**Fig. 1A**, top, yellow point mutations), and miniprotein display libraries where designed miniproteins were displayed on the surface of the nucleocapsids via genetic fusion (**Fig. 1A,** bottom, orange miniproteins). Following the delivery of library plasmids into *E. coli*, the transgene is transcribed and translated into synthetic nucleocapsid (synNC) protein subunits. Since transgene expression is driven by the strong T7 promoter, there is an abundance of self-mRNA inside the *E. coli* cells and synNCs can encapsulate their own mRNA genomes through electrostatic interactions (*12*). For both the surface library and miniprotein library, we injected the libraries into healthy mice or 4T1 tumor-bearing mice, allowed the libraries to circulate for the desired amount of time (e.g., 30 minutes), and extracted RNA from the desired tissues (**Fig. 1B**). After specifically reverse-transcribing and amplifying synNC mRNA by RT-qPCR, we sequenced the samples and cloned a new library for the next round of selection. We visualized enrichments of specific sequences in the collected organs using hierarchically clustered heatmaps, from which we identified variants of interest (**Fig. 1C**).

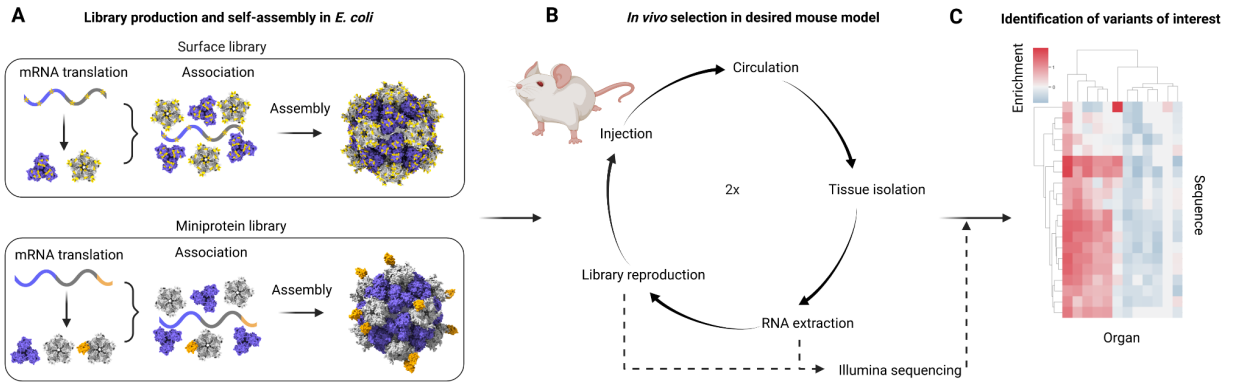


Figure 1. Experimental process for *in vivo* library selection based on synthetic nucleocapsids.
(A) Library production and self-assembly inside *E. coli.* Top: surface library with yellow point mutations. Bottom: Miniprotein library with orange miniproteins. (**B**) *In vivo* selection scheme. (**C**) Example heatmap format used to identify variants of interest.

In vivo library selection of I53-50-v4 surface mutant library

In an attempt to modulate the biodistribution of I53-50-v4, we first used the same surface library that was used to evolve longer circulation half-life in our previous work (**Fig. 1, Fig. 2A**) (*12*). This library contained roughly 10,000 unique, single-residue amino acid mutations which altered the charge or hydrophobicity of surface residues spanning both subunits of I53-50 (trimer positions 223, 226, 243, 246, 249, 251 relative to the published I53-50-v4 design model; pentamer positions 1, 4, 15, 34, 38, 39, 65, 100 relative to the published I53-50-v4 design model) (**Fig. 2B, Tab. S1**) (*12*). Instead of selecting for variants with longer circulation half-lives in blood, we enriched synNC RNA recovered from major organs (brain, heart, kidneys, liver, lungs, and spleen) over two rounds of selection. We primarily detected synNC library mRNA in the blood, liver, and spleen (**Fig. S1, Fig. 2C-D**). After sequencing the synNC mRNA from each collected tissue and calculating $\log_{10}$(enrichment) scores for each sequence compared to v4, we identified two variants of interest (**Eq. 1** in Materials and Methods, **Fig. 2E**). The selected variants were representative of two different biodistribution patterns. First, v4_D100K, a negative-to-positive mutation on the pentamer subunit, exhibited positive enrichment in major organs and negative enrichment in blood (**Fig. 2B** (left structure), **Fig. 2E** (left graph)). Second, v4_E243D, a negative-to-negative mutation on the trimer subunit, exhibited positive enrichment in the blood and major organs, albeit low in magnitude (**Fig. 2B** (right structure), **Fig. 2E** (right graph)).

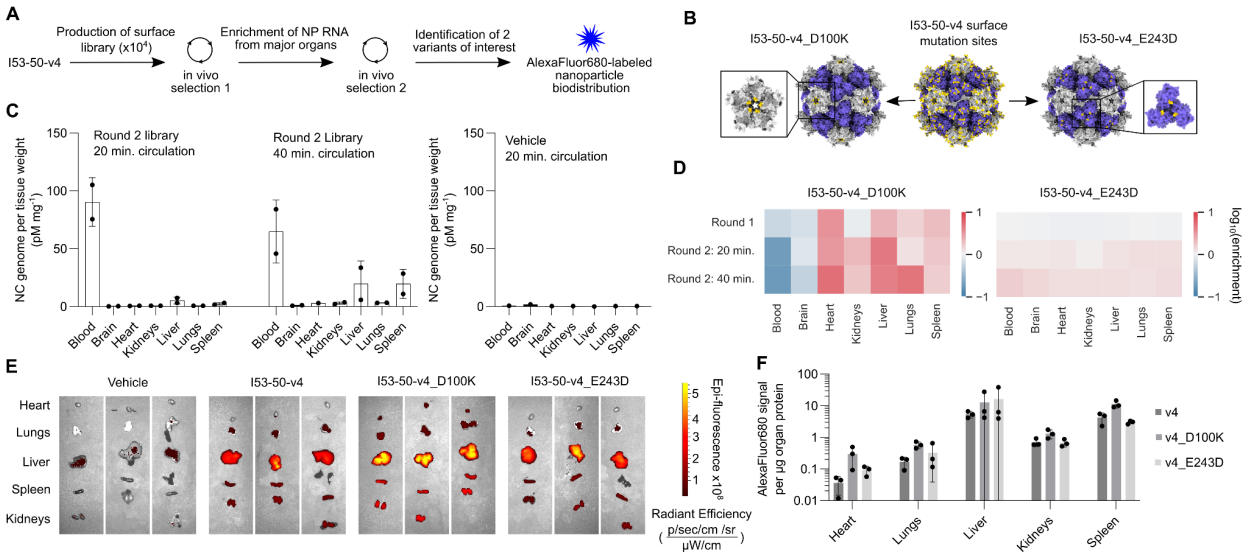


Figure 2. *In vivo* library selection and biodistribution of I53-50-v4 surface library.
(A) Experimental trajectory of surface library selection and fluorescence biodistribution studies. SynNC library mRNA was obtained from organs (brain, heart, kidneys, liver, lungs, spleen) and amplified for *in vivo* selection round 2. SynNC mRNA-seq was used to identify two variants of interest (v4_D100K and v4_E243D) that were produced as individual synNCs and fluorescently labeled for biodistribution studies. **(B)** Cartoon of locations of variants of interest (left and right) selected from the I53-50-v4 surface library (middle). The D100K mutation resides near the small central pore of the pentamer and the E243D mutation resides near the three-fold axis of the trimer. **(C)** Bulk biodistribution of the i.v.-injected surface library from *in vivo* selection round 2 after 20 or 40 minutes of circulation (left) and the vehicle control (right), measured by RT-qPCR. The majority of library mRNA was detected in the blood at 20 minutes, and in the blood, liver, and spleen at 40 minutes. **(D)** Enrichment of the two variants of interest compared to v4 (**Eq. 1**). v4_D100K shows positive enrichment in all tissues except blood and brain. v4_E243D shows slight positive enrichment in all tissues. **(E)** SynNC-Alexa Fluor 680 signal in the whole organs of mice injected with library variants of interest or vehicle control shows statistically insignificant increase in organ accumulation by v4_D100K compared to v4 and v4_E243D, and similar tissue accumulation of v4 and v4_E243D. Radiant efficiency color scale: min = 2.8×10^7, max = 5.5×10^8. **(F)** BCA-fluorescence assay on homogenized organs from (E) to measure the AF680 fluorescence per µg organ protein in major organs of interest. B: n=2 mice per group; error bars depict standard error of the mean. C: n=1 mouse per group. G: n=3 mice per group; error bars depict standard deviations. No statistically significant differences were observed in (F) by one-way ANOVA in each different organ.

We next evaluated the biodistribution of fluorescently labeled v4_D100K and v4_E243D as an orthogonal method to mRNA sequencing (**Fig. 2B, E-F**). Instead of administering self-mRNA-encapsulating synNCs to mice, we administered synNCs in which the exterior residues were conserved but the interior residues were mutated to both ablate nucleic acid encapsulation and introduce a unique cysteine handle for conjugation to Alexa Fluor 680 C2-maleimide (AF680) (**Fig. S2**). These modified synNCs are herein referred to as protein nanoparticles, as they do not encapsulate their mRNA genomes, and denoted

"v4v0-Cys-AF680," which stands for v4 exterior (version 4: surface evolved for increased circulation half-life), v0 interior (version 0: non-mRNA-encapsulating), and a unique internal cysteine (Cys) on the trimer component to which Alexa Fluor 680-C2-Maleimide was conjugated (AF680)). After injecting 1 nanomole (nmole) of asymmetric unit of each sample into mice (the asymmetric unit comprises one subunit from the pentamer and one subunit from the trimer, with 60 asymmetric units per nanoparticle, organized into 12 pentamers and 20 trimers), including a non-targeted v4v0-Cys-AF680 control and a vehicle control, we performed a PBS perfusion, imaged the whole organs for qualitative analysis, and performed a BCA-fluorescence assay on homogenized organs to quantitatively examine the AF680 fluorescence per microgram (μg) protein in each organ sample (**Fig. 2E-F**). We observed similar overall biodistribution patterns in the original v4v0-Cys-AF680 synNC and v4v0-Cys-AF680_E243D and increased but nonspecific tissue accumulation by v4v0-Cys-AF680_D100K. We did not observe any statistically significant differences in organ accumulation but note the congruence between enrichment values (**Fig. 2D**) and fluorescence signal per μg protein (**Fig. 2F**). While we did not identify any variants that shifted biodistribution in a therapeutically relevant manner, this experiment validated our platform for organ enrichment and laid the groundwork for our miniprotein display library.

***In vivo* selection of an I53-50-v4 miniprotein display library**

After observing only minor enrichment of the I53-50-v4 surface library variants in major organs, we hypothesized that displaying miniproteins on the surface of the synNC could better enable specific organ accumulation. The Baker group had previously designed millions of miniproteins intended to fold into stable structures and bind specific target proteins. For our purposes, we viewed these miniproteins as simply small, stable domains with surface patches that have physicochemical properties typical of protein-protein interfaces (*18*). To make a large miniprotein display library based on synNCs, we genetically fused a random subset of designed miniprotein libraries to either the N terminus of the v4 trimer subunit at 100% display valency or the C terminus of the v4 pentamer subunit at about 20-40% valency by using a programmed ribosomal frameshifting B sequence (prfB), herein denoted v4-trimer

display library or v4-pentamer display library, respectively (**Figure 3A-C**). The prfB sequence shifts the open reading frame +1 to avoid a stop codon and instead introduce a flexible polypeptide linker leading into the subsequent encoded miniprotein (*19*, *20*). The v4-trimer display library resulted in very low synNC yield and significant aggregation by dynamic light scattering (DLS), while the v4-pentamer display library exhibited high yield, mRNA encapsulation and protection, and monodisperse nanoparticle sizes by DLS and negative stain transmission electron microscopy (nsEM) (**Fig. 3D-G**). We noted that the exact molecular weights of the library display subunits were difficult to determine by SDS-PAGE (**Fig. 3D**), likely due to the range of sizes of miniproteins being displayed, as observed by Illumina sequencing analysis (~10-120 amino acids (aa) long with a mode of 44 aa). DLS confirmed a larger hydrodynamic radius of the v4-pentamer display library compared to v4 alone (**Fig. 3F**), and purification of individual library clones confirmed the presence of miniproteins fused to the pentamers at reduced valency (**Fig. S3**). We also observed increased susceptibility to RNase degradation of encapsulated mRNA by the v4-pentamer library compared to the original I53-50-v4 synthetic nucleocapsid ("v4"), but encapsulation and protection were efficient enough to sequence the mRNA (**Fig. 3E, S2**). The resulting synthetic nucleocapsid (synNC) library, the v4-pentamer display library, comprised ~1.7×10^7 different synNC variants that each displayed a different miniprotein fused to the surface via the pentamer (**Fig. 3C**).

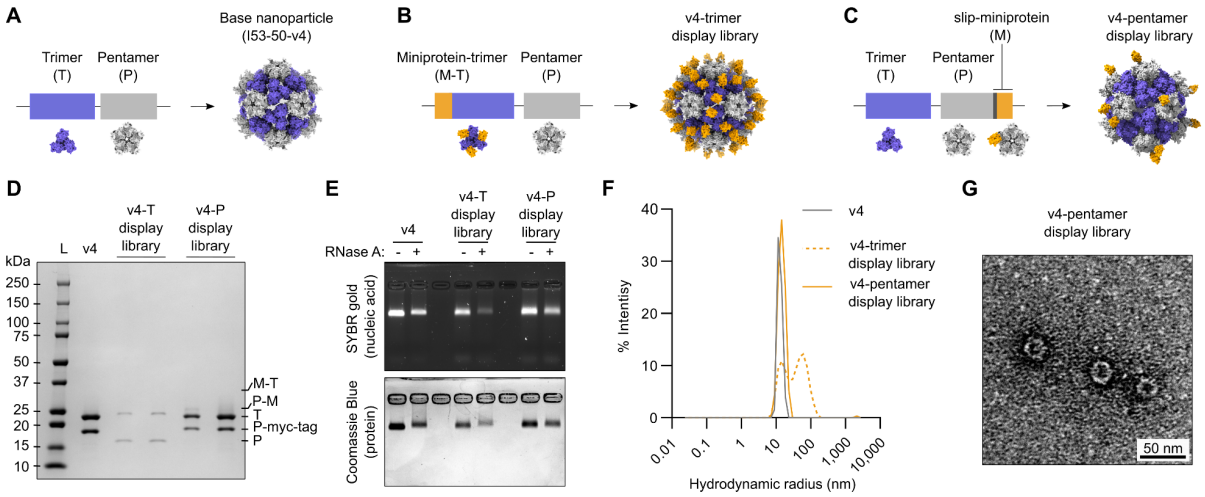


Figure 3. Biochemical characterization of synthetic nucleocapsid miniprotein display library.
(A-C) Cartoon of genetic constructs of base nanoparticle (A) and the nanoparticle displaying miniproteins on the N terminus of the trimer (B, v4-trimer display library) or C terminus of the pentamer (C, v4-pentamer display library). **(D)** Reducing SDS-PAGE of libraries compared to unmodified synNC (v4) shows lack of trimer-miniprotein fusions ("M-T") but presence of pentamer-miniprotein fusions ("P-M") at a variety of sizes as demonstrated by the smeared P-M band. The reduced valency of fused miniproteins is evident by the ratio of P-myc-tag band intensity to P-M smear intensity. These data agree with the individual library variant characterization in **Fig. S3**. **(E)** Evaluation of nuclease resistance. SynNC libraries were incubated in the presence or absence of RNase H and then analyzed for nucleic acid and protein migration on the same native agarose gel (SYBR gold nucleic acid stain, top; Coomassie Blue protein stain, bottom). All three proteins co-migrate with mRNA in the presence and absence of RNase A, suggesting mRNA encapsulation and protection. However, the v4-trimer display library shows qualitatively lighter bands after RNase A treatment, suggesting instability. **(F)** The hydrodynamic radii of the synthetic nucleocapsids were analyzed by dynamic light scattering (DLS) (I53-50-v4: gray, v4-trimer display library: dotted orange line, v4-pentamer display library: solid orange line). The v4-pentamer display library exhibited a slightly larger hydrodynamic radius than v4 alone, as expected. The v4-trimer display library exhibited aggregation, further pointing to instability of this library format. **(G)** Representative negative stain transmission electron micrograph of v4-pentamer display library shows nanoparticles of expected sizes and shapes. Scale bar: 50 nm.

In a nearly identical manner as described in the surface library section, we tested the miniprotein display library in BALB/c mice (**Fig. 1**). However, this time we utilized both healthy (wild-type) mice, and 4T1 tumor-bearing mice (*21*, *22*). 4T1 tumors are a murine model for triple-negative breast cancer. Here, we were especially interested in miniprotein enrichment in tumors, lungs, and muscles, as these tissues are important targets for cancer therapeutics, respiratory illness therapeutics, and muscular dystrophy gene therapies, respectively. We performed three stages of library selection (rounds 1A, 1B, and 2), reducing the library size at each step from 1.7×10^7 unique variants to 2.4×10^5 to 1×10^4, ultimately selecting 8 variants of interest with unique organ enrichment patterns (**Fig. 4A**). Previous work in our group demonstrated that synNCs displaying miniproteins that target specific receptors are rapidly endocytosed

and most synNC mRNA is degraded within 1-2 hours (data not shown). To ensure the integrity of encapsulated synNC mRNA, we elected to collect mouse organs after allowing the library to circulate for either 5 or 30 minutes. We observed similar synNC mRNA yields at both time points, and thus used 30 minute circulation times in all following *in vivo* studies (**Fig. S4A-C**). We performed RT-qPCR on the bulk synNC library genomes collected from major tissues at each round of selection, and detected lower blood accumulation and higher liver and spleen accumulation compared to the surface library (**Fig. 4B-C, Fig. S4A-F**). This was expected, as the display of targeting domains on the surface of synNCs shields part of the protein nanoparticle surface that was previously evolved for a longer circulation half-life (*12*). Additionally, we estimated a shorter circulation half-life for the bulk miniprotein display library compared to I53-50-v4 by RT-qPCR (about 44 minutes compared to 4.5 hours, respectively) (**Fig. S4D**).

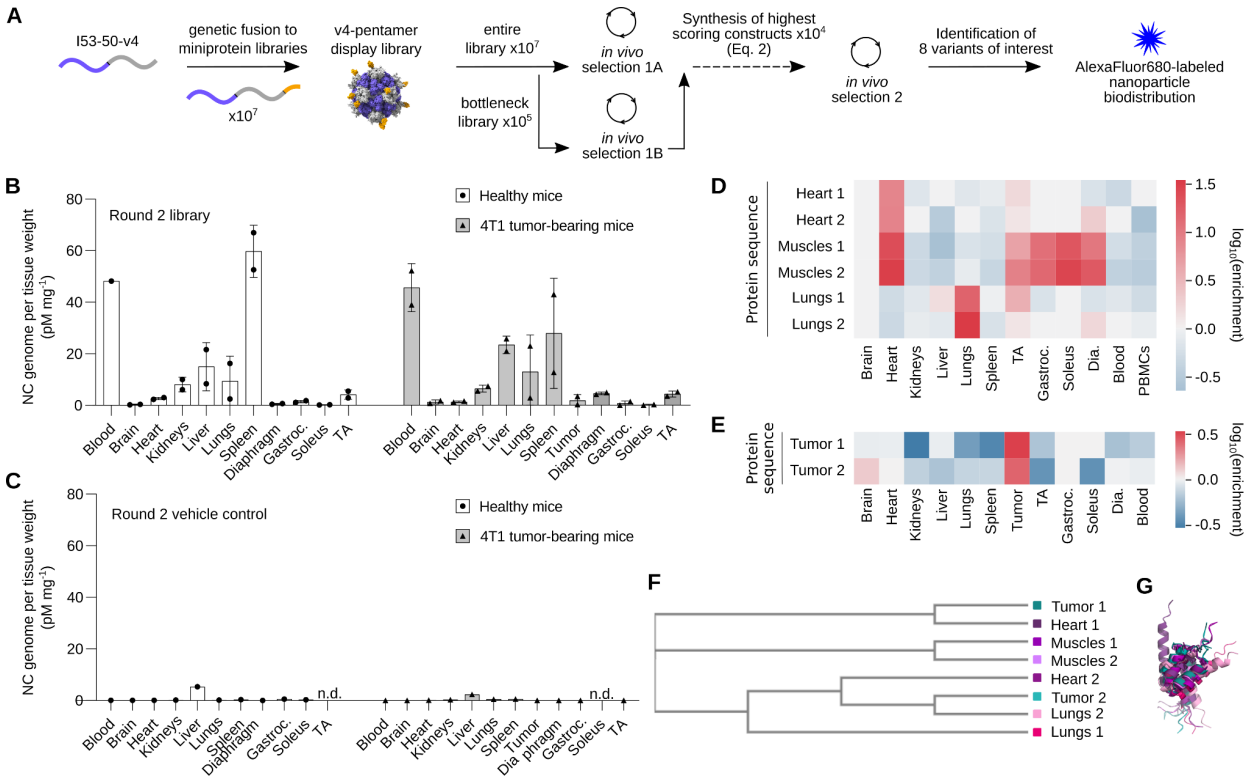


Figure 4. Biodistribution and sequencing analysis of miniprotein display library.
(A) Experimental trajectory of library selection and fluorescence biodistribution studies. SynNCs produced in *E. coli* were used in two rounds of *in vivo* library selection. In round 1A, the entire library was used. In round 1B, a smaller ("bottleneck") library was injected to achieve even higher sequencing coverage. After rounds 1A and 1B, enriched miniprotein sequences (n = 10,000) in the heart or tumor were selected and synthesized as a new library solely containing these sequences. After *in vivo* selection round 2, eight variants of interest with preferential enrichment in

Figure 4 continued.
the lungs, heart, heart and muscles, or tumor were selected for individual biodistribution assessment. **(B-C)** Biodistribution of bulk synNC mRNA recovered from major organs by RT-qPCR for the library (B) and vehicle control (C), reported as picomolar (pM) nucleocapsid (NC) genome per mg of tissue. **(D-E)** Enrichment of sequences of interest in tissues of interest (heart only, heart and all other muscles tested (Muscles), lungs, tumor) in healthy mice (D) and tumor-bearing mice (E). **(F)** Phylogenetic tree from MUltiple Sequence Comparison by Log-Expectation (MUSCLE) of miniproteins identified in D, E. **(G)** Superimposition of miniprotein structures predicted by ColabFold shows distinct conformations, despite similar protein sequences. The color scale matches the phylogenetic tree labels in (F). B: n=2 mice per group; error bars depict standard error of the mean. C: n=1 mouse per group. "n.d.": no data.

We next sequenced the synNC RNA recovered from the blood, heart, lungs, liver, kidneys, and spleen from all healthy mice and tumors isolated from tumor-bearing mice. We calculated the $\log_{10}$ enrichment values of sequences identified from these organs compared to the input library that was administered to the mice, and selected 10,000 sequences that were preferentially enriched in heart or tumor (**Eq. 2** in Materials and Methods).

We re-synthesized a library containing these 10,000 variants for the third round of *in vivo* selection, after which we performed mRNA-Seq on all aforementioned organs in addition to major muscle groups (diaphragm, gastrocnemius, soleus, and tibialis anterior) and selected two variants enriched in each target organ or organ group for further characterization: heart only, heart and major muscle groups, lungs, or 4T1 tumors from tumor-bearing mice (**Fig. 4D-E**, **Eq. 1**). Although some of the eight selected miniproteins have homologous sequences, AlphaFold2 predictions revealed unique 3D conformations, surface hydrophobic networks, and surface charge distributions (**Fig. 4F-G**, **Fig. 5**, **Tab. S2**) (*23*, *24*). We were encouraged to see that each of the eight variants, with the exception of the second tumor-enriched miniprotein (**Fig. 5I**), featured hydrophobic surface patches that are characteristic of high-affinity minibinders (*25*).

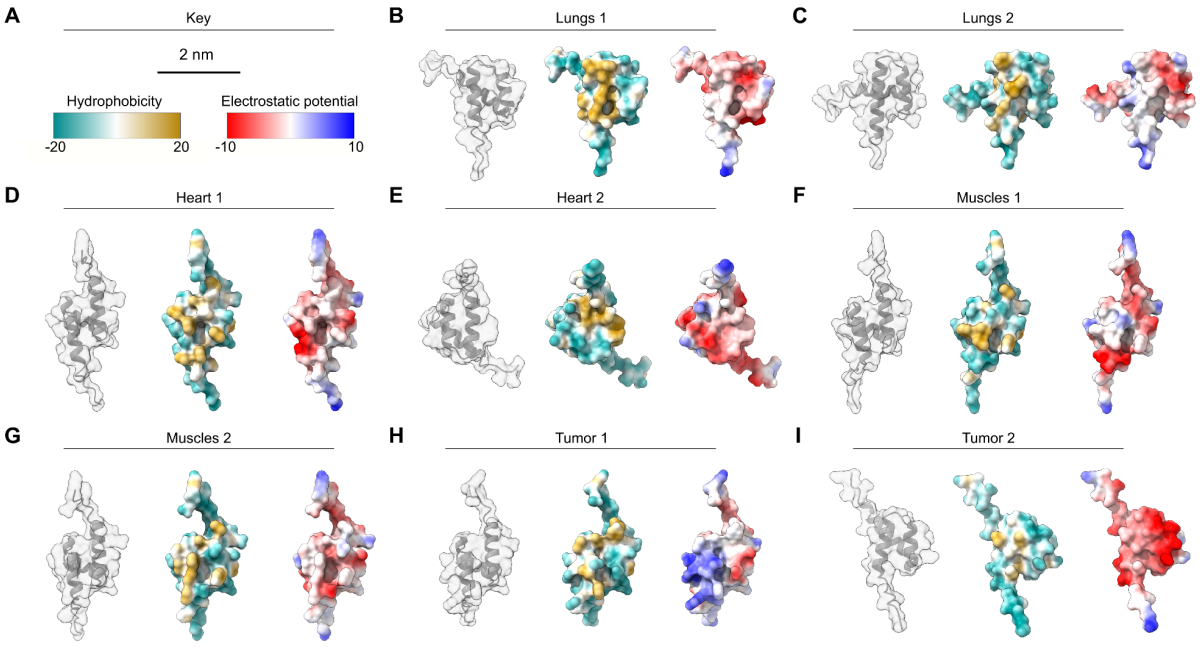


Figure 5. Enriched miniproteins have distinct topologies, surface hydrophobic networks and surface charge distributions.
(A) Key. **(B-I)** AlphaFold2 predictions (*23*) of eight selected miniprotein structures enriched in lungs, heart, muscles, or tumor. The secondary structure, surface hydrophobicity, and surface electrostatic potentials of each miniprotein were visualized in UCSF ChimeraX (*24*).

We then tested the biodistribution of fluorescently labeled nanoparticles displaying each of the eight selected miniproteins. Using the same method as for the surface library variants of interest (**Fig. 1F-G**), we individually produced and purified the protein nanoparticles, as well as a bare nanoparticle (i.e., non-targeted) control, each with AF680 covalently conjugated to the interior surface of the nanoparticle (**Fig. S2, Fig. S5**). We also produced two fluorescent nanoparticles displaying miniprotein sequences that were enriched in the spleen, but found that the observed spleen enrichment was likely the result of free mRNA escaping unstable synNC variants (**Fig. S5N-O, Fig. S5S-T, Fig. S6**). As with the surface library variants, we examined fluorescence biodistribution by both qualitative whole organ IVIS imaging (**Fig. S7**) and a quantitative BCA-fluorescence assay on homogenized organs (**Fig. 6** and **Fig. S8**). We measured the AF680 fluorescence per µg organ protein in each sample and normalized to the summed values across all analyzed tissues in each mouse to enable relative comparisons between samples. The

same data are plotted as raw values in **Fig. S8**. In the non-muscle organs, we observed 5.7- and 4.8-fold higher relative normalized accumulation of the Lungs 1 and Lungs 2 variants in the lungs compared to the non-targeted control, respectively ($p \leq 0.01$) (**Fig. 6C**); 5.7-fold and 3.0-fold higher relative accumulation of the Muscles 1 and Muscles 2 variants in heart tissue compared to the non-targeted control, respectively ($p \leq 0.01$) (**Fig. 6E**); and did not observe statistically significant relative accumulation of any other variant. We suspected the minor increased relative accumulation of Tumor 1 in 4T1 tumors could have been due to nonspecific accumulation enabled by the large, positively charged patch on the surface of the miniprotein (**Fig. 6H**). Additionally, since *in vivo* phage display often yields targeting domains that bind tumor vasculature (*26*, *27*), we suspected Tumor 1 could have bound tumor vasculature instead of the 4T1 tumor cells. However, we did not observe binding of Tumor 1 to 4T1 tumor cells or to murine brain endothelial (BEND3) cells *in vitro* (**Fig. S9**). To examine distribution- and elimination-related effects on organ accumulation, we measured the serum circulation half-lives of non-targeted NP-AF680, L1-AF680, M1-AF680, and T1-AF680 (**Fig. S10**, **Tab. S3**). Based on a two-compartment pharmacokinetic model, all samples exhibited distribution half-lives less than 30 minutes except M1-AF680, which exhibited a distribution half-life of 1.26 hours. We also noted increased relative target organ accumulation by fluorescent nanoparticles displaying organ-matched miniproteins compared to nanoparticles displaying non-organ-matched miniproteins. For example, the Lungs 1 and Lungs 2 samples were the only samples to exhibit significantly increased accumulation in the lungs, and Muscles 1 and Muscles 2 samples were the only samples to exhibit significantly increased relative accumulation in heart tissue (**Fig. 6**). Lastly, we observed a slight positive correlation between enrichment score and fold-change of both non-relative and relative organ accumulation compared to non-targeted nanoparticles (**Fig. S11**).

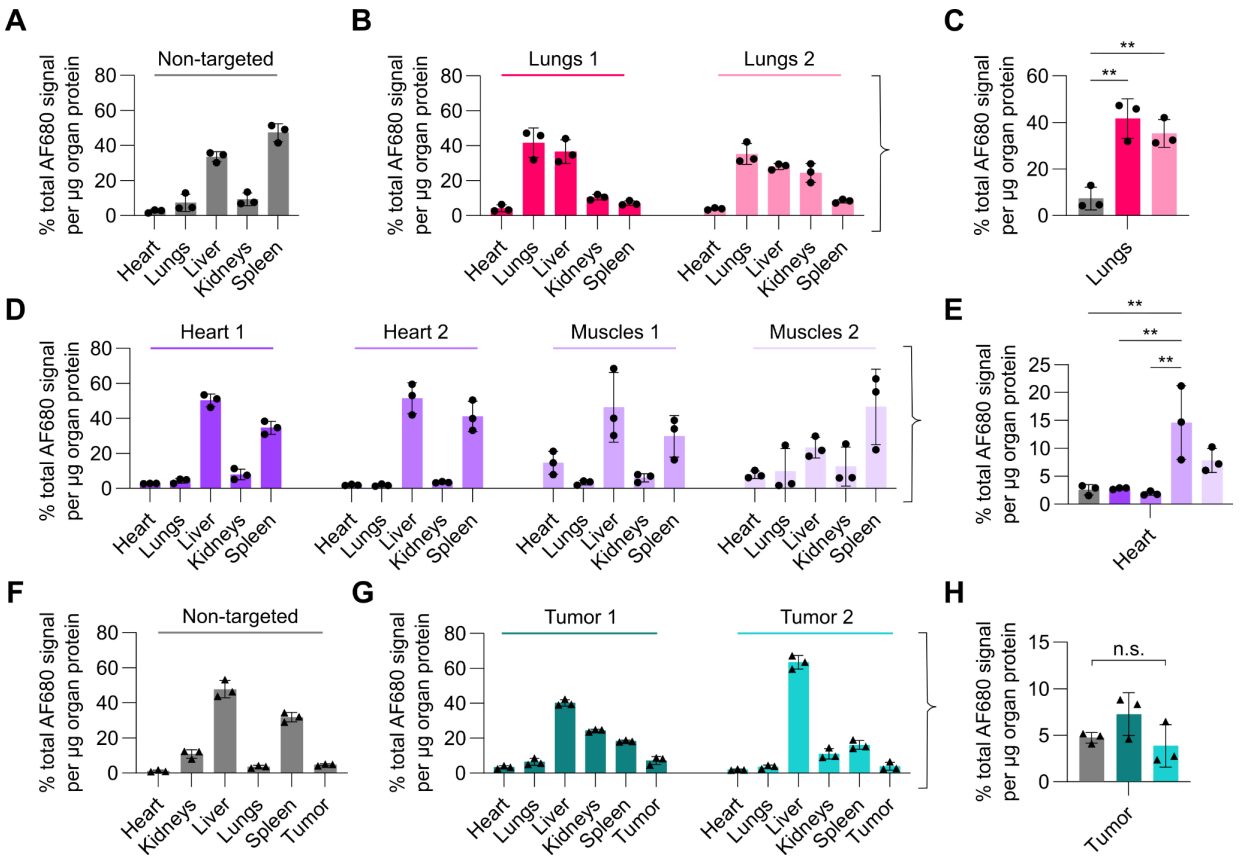


Figure 6. Normalized fluorescence biodistribution of variants of interest.
Each graph shows the normalized fluorescence signal per µg homogenized organ protein from mice injected with AF680-labeled nanoparticles (% total fluorescence per µg organ protein). In other words, each organ's proportion of the total signal from each mouse is plotted. **(A)** Non-targeted nanoparticle in healthy mice; **(B-C)** Lungs 1 and 2 library variants in all organs (D) and compared to non-targeted nanoparticle in healthy mice in the lungs only (E); **(D-E)** Heart 1, Heart 2, Muscles 1, and Muscles 2 library variants in all organs (D) and compared to non-targeted nanoparticle in healthy mice in the heart only (E); **(F)** Non-targeted nanoparticle-AF680 in 4T1 tumor-bearing mice; **(G-H)** Tumor 1 and 2 library variants in all organs (G) and compared to non-targeted nanoparticle in the tumor only (H). Circles depict individual values from healthy mice. Triangles depict individual values from 4T1 tumor-bearing mice. n=3 mice per group. Error bars depict standard deviations. Statistics: Ordinary one-way ANOVA with Tukey's correction for multiple comparisons. **: $p \leq 0.01$; ns: not statistically significant. Only statistically significant comparisons are indicated in C, E.

Discussion

Here, we employ *in vivo* library selection of synthetic nucleocapsids to identify variants with altered tissue biodistribution and to select for specific tissue accumulation. *In vivo* library selection is a powerful method for targeting domain identification because millions of candidates can be simultaneously injected and evaluated *in vivo*, and those exhibiting the desired behavior can be identified by linking genotype to

phenotype (e.g., biodistribution). Unlike *in vitro* library selection approaches, which are typically based on a subset of the functional requirements required to achieve a desired outcome (e.g., receptor binding), *in vivo* library selection subjects the libraries to the complete set of known and unknown functional requirements and physiological barriers. We are still unable to comprehensively recapitulate and account for these physiological barriers during *in silico* design or *in vitro* testing.

The stringency of *in vivo* selection together with the semi-rational library design (i.e., randomly selected miniproteins that were computationally designed to fold into binder-like structures) enabled the identification of well-performing variants in just two rounds of selection. Other *in vivo* library selection techniques like peptide phage display have also been shown to yield promising results in just 2-3 rounds of *in vivo* selection (*28*, *29*). In comparison, *in vitro* selection techniques like SELEX, in which random oligonucleotide libraries are screened against a target, typically require 5-10 rounds of selection to identify aptamers that bind specific cell populations or 10-15 rounds of selection to identify aptamers that bind specific proteins (*30*–*32*). Notably, incorporating more rational design into both library design and *in vivo* selection strategy could further decrease the number of rounds of selection needed to achieve a desired result and further improve the quality of the molecules identified. For example, designing high-quality libraries of miniprotein binders against specific targets is now possible using cutting-edge machine learning-based design methods, and would enable the *in vivo* library selection strategy to be tailored to those targets (*33*, *34*).

The ability to sequence library mRNA from multiple tissues enabled simultaneous positive selection for accumulation in tissues of interest and negative selection for accumulation in off-target tissues. However, this selection scheme is not immune to the pervasive hurdle of distinguishing false positives from true positives that recognize the target through a specific molecular interaction. Four of the eight selected variants exhibited statistically significant shifts in biodistribution and target organ accumulation compared to non-targeted nanoparticles (Lungs 1, Lungs 2, Muscles 1, Muscles 2), and one library variant showed a

minor, non-statistically significant increase in relative 4T1 tumor accumulation compared to non-targeted nanoparticles (Tumor 1). Consistent with previous results (*12*), none of the variants with $\log_{10}$(enrichment) scores below 1.0 exhibited a statistically significant increase in fluorescent nanoparticle accumulation in the target tissue(s), while three of the four variants with $\log_{10}$(enrichment) scores greater than 1.0 did (**Fig. S9**). However, we have yet to demonstrate whether these shifts in biodistribution are caused by specific receptor-ligand interactions; other undesired or nonspecific phenomena could be responsible. Additionally, during sequencing analysis and structure prediction, we observed several miniprotein sequences with unpaired, surface exposed cysteines enriched in the tumor (data not shown). It has previously been demonstrated that the incorporation of free cysteines in peptides increases tumor accumulation by forming disulfide bonds to serum albumin rather than a specific receptor-ligand interaction; therefore, we did not select these sequences for individual testing (*35*). Further experiments such as organ cryosectioning and immunofluorescence imaging are needed to identify where in the target tissues our selected nanoparticles are accumulating and to what molecular targets the miniproteins might be binding.

Importantly, the ability to detect congruence between separate readouts may have helped avoid the selection of false positives in some cases. Both of the variants that were selected for preferential enrichment in all five of the muscles—heart, diaphragm, gastrocnemius, soleus, and tibialis anterior—were mutually corroborating (Muscles 1 and Muscles 2, **Fig. 4D**). Indeed, both of these variants successfully showed increased accumulation in the target tissues by fluorescence biodistribution (**Fig. 6D-E, Fig. S8D-I**). Additionally, the ability to rapidly predict the 3D conformation of miniproteins of interest with reasonably high confidence, confirming they were likely folding into tertiary protein structures with binder-like properties, aided our selection of variants to further examine (**Fig. 5**) (*23*, *36*).

Our results support the growing body of research highlighting the power of *in vivo* library selection for identifying nanoparticles with desired physiological and functional properties (*4*, *5*, *12*). Not only could

miniprotein binders of interest from this study be modularly applied to other delivery platforms to potentially confer altered tissue tropism (e.g., the Muscles 2 miniprotein could be displayed on lentivirus or AAV via SpyCatcher-SpyTag (*37*, *38*)), synNCs displaying miniproteins of interest could also be formulated as delivery vehicles themselves by encapsulating therapeutic molecules instead of self-mRNA or fluorophores (*39*). Alternatively, miniproteins could be optimized for desired *in vivo* behaviors using this platform, then extracted and used on their own in a manner similar to antibody-based therapeutics (*40*). These methods could also be applied to virus-like particles or protein nanoparticles of different sizes, shapes, and surface chemistries that could be better-suited to specific delivery applications (*34*, *41–46*). In conclusion, the synthetic nucleocapsid library selection platform described in this work can be implemented in various ways to advance the discovery and optimization of protein-based therapeutics.

Materials and Methods

Equations

Equation 1. Log_{10}(enrichment) score of synNC surface library variant sequence *s* in tissue *t* compared to v4 sequence *v* from tissue *t* containing up to *x* unique sequences.

$$\log_{10}(\text{enrichment})_{s,o} = [\ (\text{representation of } s \text{ in } t) / (\text{representation of } s \text{ in } i)\ ]$$

$$= [\ (\sum_{s}^{t} \text{reads} / \sum_{0-x}^{t} \text{reads}\ ) / (\sum_{v}^{t} \text{reads} / \sum_{0-x}^{t} \text{reads}\ )\ ]$$

$$= [\ (\sum_{s}^{t} \text{reads}\ ) / (\sum_{v}^{t} \text{reads})\ ]$$

Equation 2. Log_{10}(enrichment) score of miniprotein sequence *m* in tissue *t* containing up to *x* unique sequences compared to input library dose *i* containing up to *y* unique sequences.

$$\log_{10}(\text{enrichment})_{m,t} = [\ (\text{representation of } m \text{ in } t) / (\text{representation of } m \text{ in } i)\ ]$$

$$= [\ (\sum_{m}^{t} \text{reads} / \sum_{0-x}^{t} \text{reads}\ ) / (\sum_{m}^{i} \text{reads} / \sum_{0-y}^{i} \text{reads}\ )\ ]$$

Equation 3. AF680 signal per µg tissue protein normalized to the sum of each organ for each individual mouse.

% total AF680 signal per µg tissue protein t from mouse m with x number of tissues =

$$[\ (\text{AF680 RFU in } x_t) / (\mu\text{g protein in } x_t)\ ] / [\ \sum_{0-x}^{m} (\text{AF680 RFU}) / (\mu\text{g protein})$$

Cloning synthetic nucleocapsid libraries.

The I53-50-v4 synthetic nucleocapsid (synNC) genome was PCR-amplified as previously described (*12*). Randomly chosen (double-blind) libraries of miniprotein designs were obtained from the Baker Lab oligonucleotide chip stocks, synthesized by Twist, IDT, CustomArray, or Agilent. Miniprotein sequences were qPCR-amplified by primers that included overlap sites with the synNC genome (**Tab. S1**). Assembly PCR was performed to genetically fuse the miniproteins to the synthetic nucleocapsids. Assemblies were cloned into pet29b(+) protein production vectors via Gibson Assembly (NEB E2611L), electroporated into DH10beta cells (NEB C3019H), and grown at 37°C on agar-kanamycin plates overnight. Plates were scraped, plasmid DNA was isolated, plasmid DNA was electroporated into BL21 Star (DE3) (ThermoFisher C601003) or Lemo21(DE3) (NEB C2528J) producer cells, and cells were grown at 37°C on agar-kanamycin plates overnight. Fresh chemically competent cells were expanded and prepared for electroporation per manufacturer's instructions in Milli-Q water. Bacterial lawns were collected from the plates and used to inoculate large producer cultures for protein production. At both plating steps, bacteria were serially diluted and spotted onto separate agar plates to calculate the approximate library size and coverage. Library coverages of 10- to 100-fold were obtained in all selection steps except miniprotein library selection round 1A.

Protein production and fluorescent labeling.

Proteins were produced and purified as previously described (*12*). Briefly, proteins were expressed in *E. coli* via IPTG induction or autoinduction (*47*), cell pellets were homogenized and microfluidized, and the

supernatant was subjected to immobilized metal affinity chromatography (IMAC) at pH 8.0. Triton X-114 (Sigma Aldrich X114-500ML, CAS 9036-19-5) was used to remove lipopolysaccharides (endotoxin). Samples were incubated with 10 µg/mL RNase H (ThermoFisher 18021071) for 10 minutes at room temperature to remove free RNA. Immediately following RNase treatment, size exclusion chromatography (SEC) was performed on a Superose 6 Increase 10/300 GL column (Cytiva 29-0915-96) on a GE AKTA Pure with PBS or HEPES buffer at pH 7.4. Samples were stored at 4°C and sterile filtered before use. For fluorescent labeling, samples were labeled with 5- to 10-fold molar excess of Alexa Fluor 488 C5 maleimide (ThermoFisher A10254) or Alexa Fluor 680 C2 maleimide (ThermoFisher A20344) at 4°C overnight or at room temperature for 2 hours prior to dialysis to remove free fluorophore, quenching with 10-fold molar excess dithiothreitol (DTT, ThermoFisher P2325), purification on a PD-10 desalting column (Cytiva 17085101), and purification by SEC.

Biochemical characterization.

Proteins were characterized as previously described (*12*). Briefly, molecular weight was analyzed by SDS-PAGE, nucleic acid and protein co-migration and RNase resistance was analyzed by native agarose gel electrophoresis, and polydispersity and size were measured by dynamic light scattering (DLS) and negative stain transmission electron microscopy (nsEM). Protein concentrations were measured by either Qubit (Invitrogen Q33212), UV-Vis spectroscopy, or Bradford assay (Thermo Scientific 23200). Molecular protein masses were measured by intact mass spectra obtained via reverse-phase LC/MS on an Agilent G6230B TOF on an AdvanceBio RP-Desalting column, and deconvoluted with Bioconfirm using a total entropy algorithm. Endotoxin was measured (Charles River Endosafe LAL Endotoxin Cartridges, PTS201F) (**Tab. S4**).

In vivo library circulation, fluorescence biodistribution, and organ recovery.

BALB/c mice (BALB/cAnNCrl, Inbred, strain code 028) were purchased from Charles River Laboratories at 6-8 weeks of age. 9-10-week old tumor-bearing mice groups were injected with 1-2 million 4T1 cells 1-2 weeks before the study and monitored per an approved IACUC protocol. Mice were intravenously (retro-orbitally) injected with 150 µL of sample (0.5 - 4 nanomoles protein nanoparticle asymmetric unit). After the designated circulation time, mice were euthanized with Avertin overdose and cardiac puncture. Blood was collected from the vena cava into EDTA-lined tubes. PBS perfusions through the left ventricle were performed. Organs were collected and immediately snap-frozen. For fluorescent imaging, organs were imaged with Xenogen IVIS before snap-freezing.

RNA isolation from frozen tissues was performed within one week of the *in vivo* study. Frozen tissues were ground in liquid nitrogen with a mortar and pestle over dry ice to remain frozen. A subset of the ground tissue was weighed, dissolved in TRIzol, sonicated with a Sonic Dismembrator 60 (Fisher Scientific), and stored at -80°C until RNA extraction. We experienced great difficulty homogenizing the muscles with the hand-held homogenizer. Snap-freezing and grinding the frozen muscles with a mortar and pestle, as we did to recover RNA from muscles and organs during the selection steps, could be a more reliable method to measure accumulation in muscles moving forward.

To measure AF680 fluorescence per µg tissue protein, tissues were thawed and suspended in RIPA buffer (Thermo 89901) supplemented with DNase (Thermo 90083) before homogenizing with Omni Tissue Homogenizer (TH). Homogenized tissues were spun down and lysate relative fluorescence units at 679 nm were read on a plate reader (infinite M200PRO plate reader, Tecan). In parallel, Pierce BCA Protein Assay Kits were utilized according to the manufacturer's instructions to quantify the amount of tissue protein per sample (ThermoFisher 23225). AF680 signal per µg organ (or tissue) protein was calculated by dividing the AF680 RFU by the µg organ protein calculated by the BCA assay. These values were then normalized (**Eq. 3** in Materials and Methods).

To measure serum circulation half-lives, 0.5 nmoles of AF680-labeled nanoparticle ASU was I.V. R.O.I.-injected into healthy BALB/c mice. Each sample was administered to two groups of mice (n = 3 mice per group). Blood was collected from group A at 5 minutes, 2 hours, and 24 hours. Blood was collected from group B at 30 minutes, 6 hours, and 24 hours. Blood was immediately centrifuged, serum was aspirated, and samples were frozen on dry ice. Samples were stored at -80°C for up to 4 days prior to measuring fluorescence with an infinite M200PRO plate reader (Tecan). Two-phase exponential decay was performed in GraphPad Prism 9.

All animal experiments were following approved IACUC protocol methods meeting OAW standards.

RNA extraction, reverse transcription, quantitative PCR, and Illumina sequencing.

Samples were stored in TRIzol (Invitrogen 15596026) at -80°C. RNA was purified from samples using RNeasy Kits (Qiagen 74104 or 74106) according to the manufacturer's instructions. RNA and DNA concentrations were measured with NanoDrop or Qubit (Invitrogen Q32852, Q32854). Reverse transcription (PrimeScript 1st strand cDNA Synthesis Kit, Takara 6110A), quantitative PCR (KAPA HiFi HotStart Ready Mix, Roche; SYBR Green, Invitrogen S7563), and sequencing on the Illumina MiSeq (Illumina MiSeq Reagent Kit v3, MS-102-3003; 5% PhiX Control v3, FC-110-3001) were performed as previously described and/or according to manufacturer's instructions unless otherwise noted (*12*).

The skpp-132-R reverse primer was used for reverse transcription instead of the primers and 6mers provided by the Takara PrimeScript 1st strand cDNA Synthesis Kit (5'-CATACTGTTGGTTGCTAGGC-3'). The skpp-132-F forward primer was used to amplify synNC genomes with qPCR (5'-TAGGATTACTGCTCGGTGAC-3'). The skpp-offset-R reverse primer was used to amplify synNC genomes and to later amplify miniprotein sequences out of synNC genomes with qPCR (5'-GTTGCTAGGCTCAGTGATGG-3'). A separate forward primer was used in combination with skpp-132-R to amplify the miniprotein sequences out of synNC genomes for sequencing (forward primer named c_lib_fwd, 5'-GTATCTTTGACGGCTCCGGT-3').

Sequencing analysis.

Sequencing analysis was performed by aligning the MiSeq output files with PEAR (a fast and accurate Illumina Paired-End reAd mergeR) and using custom Python scripts (*48*). Scripts are available upon request.

Flow cytometry binding study.

4T1 and BEND3 cells were separately cultured in RPMI 1640 with 10% FBS (Gibco 11875093) at 37°C at 5% CO_2. Cells were dissociated with StemPro Accutase (Gibco A1110501). Cells were stained with ZombieViolet viability stain (Biolegend 423113) in HEPES buffer (20 mM HEPES, 150 mM NaCl, pH 7.4). Fluorescent nanoparticles were incubated with cells in HEPES buffer with 1% w/v BSA for 30 minutes on ice. Cells were washed three times with HEPES-1% BSA prior to examination on an Attune NxT Flow Cytometer (Invitrogen). Data were analyzed using FlowJo.

Figures and statistical analysis.

Figures were made using Inkscape, GraphPad Prism, PyMOL, UCSF ChimeraX, Biorender, Python, and FlowJo. Statistical analysis was performed using GraphPad Prism 9.

Acknowledgements

The authors gratefully acknowledge Dr. Drew Sellers, Nataly Kacherovsky, Dr. Heather H. Gustafson, Dr. Karla Herpoldt, Brooke Fiala, Dr. Inna Goreshnik, Dr. David Peeler, Dr. Cassie Bryan, Dr. Ian Cardle, Dr. Gary Liu, Dr. Ivan Anishchanka, Dr. Ratika Krishnamurty, Dr. Barry Stoddard, Erin Yang, and Kandise Van Wormer for technical and scientific advice, feedback on the manuscript, lab and administrative support, and insightful conversations. The authors would also like to thank Xinting Li, Mila Lamb, and

Stephen Rettie for assistance with mass spectrometry. This work was supported by the National Cancer Institute (NCI) grants 1R21CA232430 (N.P.K. and S.H.P.), 1R01CA257563 (S.H.P) and T32CA080416 (A.O.). Defense Threat Reduction Agency Grant HDTRA1-18-1-0001 (D.B. and N.P.K.), the Bill and Melinda Gates Foundation INV-010680 (D.B. and N.P.K.), the Audacious Project at the Institute for Protein Design (N.P.K. and D.B.), and National Institute of Allergy and Infectious Disease (NIAID) grant U54AI170856 (N.P.K.).

References

1. A. Olshefsky, C. Richardson, S. H. Pun, N. P. King, Engineering Self-Assembling Protein Nanoparticles for Therapeutic Delivery. *Bioconjug. Chem.* **33**, 2018–2034 (2022).

2. M. J. Mitchell, M. M. Billingsley, R. M. Haley, M. E. Wechsler, N. A. Peppas, R. Langer, Engineering precision nanoparticles for drug delivery. *Nat. Rev. Drug Discov.* **20**, 101–124 (2021).

3. E. Blanco, H. Shen, M. Ferrari, Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nat. Biotechnol.* **33**, 941–951 (2015).

4. C. D. Sago, M. P. Lokugamage, F. Z. Islam, B. R. Krupczak, M. Sato, J. E. Dahlman, Nanoparticles That Deliver RNA to Bone Marrow Identified by in Vivo Directed Evolution. *J. Am. Chem. Soc.* **140**, 17095–17105 (2018).

5. Q. Cheng, T. Wei, L. Farbiak, L. T. Johnson, S. A. Dilliard, D. J. Siegwart, Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR–Cas gene editing. *Nat. Nanotechnol.* **15**, 313–320 (2020).

6. M. P. Lokugamage, D. Vanover, J. Beyersdorf, M. Z. C. Hatit, L. Rotolo, E. S. Echeverri, H. E. Peck, H. Ni, J.-K. Yoon, Y. Kim, P. J. Santangelo, J. E. Dahlman, Optimization of lipid nanoparticles for the delivery of nebulized therapeutic mRNA to the lungs. *Nat Biomed Eng*. **5**, 1059–1068 (2021).

7. M. Qiu, Y. Tang, J. Chen, R. Muriph, Z. Ye, C. Huang, J. Evans, E. P. Henske, Q. Xu, Lung-selective mRNA delivery of synthetic lipid nanoparticles for the treatment of pulmonary lymphangioleiomyomatosis. *Proc. Natl. Acad. Sci. U. S. A.* **119** (2022), doi:10.1073/pnas.2116271119.

8. S. G. Huayamares, M. P. Lokugamage, R. Rab, A. J. Da Silva Sanchez, H. Kim, A. Radmand, D. Loughrey, L. Lian, Y. Hou, B. R. Achyut, A. Ehrhardt, J. S. Hong, C. D. Sago, K. Paunovska, E. S. Echeverri, D. Vanover, P. J. Santangelo, E. J. Sorscher, E. Dahlman, High-throughput screens identify a lipid nanoparticle that preferentially delivers mRNA to human tumors in vivo. *J. Control. Release* (2023), doi:10.1016/j.jconrel.2023.04.005.

9. H. H. Gustafson, A. Olshefsky, M. Sylvestre, D. L. Sellers, S. H. Pun, Current state of in vivo panning technologies: Designing specificity and affinity into the future of drug targeting. *Adv. Drug Deliv. Rev.* **130**, 39–49 (2018).

10. K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, A new type of synthetic peptide library for identifying ligand-binding activity. *Nature*. **354**, 82–84 (1991).

11. R. C. Münch, H. Janicki, I. Völker, A. Rasbach, M. Hallek, H. Büning, C. J. Buchholz, Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer. *Mol. Ther.* **21**, 109–118 (2013).

12. G. L. Butterfield, M. J. Lajoie, H. H. Gustafson, D. L. Sellers, U. Nattermann, D. Ellis, J. B. Bale, S. Ke, G. H. Lenz, A. Yehdego, R. Ravichandran, S. H. Pun, N. P. King, D. Baker, Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature*. **552**, 415–420 (2017).

13. J. B. Bale, S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King, D. Baker, Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science*. **353**, 389–394 (2016).

14. K. Hertveldt, T. Beliën, G. Volckaert, General M13 phage display: M13 phage display in identification and characterization of protein-protein interactions. *Methods Mol. Biol.* **502**, 321–339 (2009).

15. T. G. W. Edwardson, T. Mori, D. Hilvert, Rational Engineering of a Designed Protein Cage for siRNA Delivery. *J. Am. Chem. Soc.* **140**, 10439–10442 (2018).

16. J. Marcandalli, B. Fiala, S. Ols, M. Perotti, W. de van der Schueren, J. Snijder, E. Hodge, M. Benhaim, R. Ravichandran, L. Carter, W. Sheffler, L. Brunner, M. Lawrenz, P. Dubois, A. Lanzavecchia, F. Sallusto, K. K. Lee, D. Veesler, C. E. Correnti, L. J. Stewart, D. Baker, K. Loré, L. Perez, N. P. King, Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell*. **176**, 1420–1431.e17 (2019).

17. D. Martinez-Cano, R. Ravichandran, H. Le, H. E. Wong, B. Jagannathan, E. J. Liu, W. Bailey, J. Yang, K. Matthies, H. Barkhordarian, B. Shah, N. Srinivasan, J. Zhang, A. Hsu, J. Wypych, J. Stevens, D. M. Piedmonte, L. P. Miranda, L. Carter, M. Murphy, N. P. King, N. Soice, Process Development of a SARS-CoV-2 Nanoparticle Vaccine. *Process Biochem.* (2023), doi:10.1016/j.procbio.2023.03.014.

18. A. Chevalier, D.-A. Silva, G. J. Rocklin, D. R. Hicks, R. Vergara, P. Murapa, S. M. Bernard, L. Zhang, K.-H. Lam, G. Yao, C. D. Bahl, S.-I. Miyashita, I. Goreshnik, J. T. Fuller, M. T. Koday, C. M. Jenkins, T. Colvin, L. Carter, A. Bohn, C. M. Bryan, D. A. Fernández-Velasco, L. Stewart, M. Dong, X. Huang, R. Jin, I. A. Wilson, D. H. Fuller, D. Baker, Massively parallel de novo protein design for targeted therapeutics. *Nature*. **550**, 74–79 (2017).

19. P. V. Baranov, R. F. Gesteland, J. F. Atkins, Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.* **3**, 373–377 (2002).

20. C. L. Sanders, J. F. Curran, Genetic analysis of the E site during RF2 programmed frameshifting. *RNA*. **13**, 1483–1491 (2007).

21. B. A. Pulaski, S. Ostrand-Rosenberg, *Curr. Protoc. Immunol.*, in press.

22. B. Schrörs, S. Boegel, C. Albrecht, T. Bukur, V. Bukur, C. Holtsträter, C. Ritzel, K. Manninen, A. D. Tadmor, M. Vormehr, U. Sahin, M. Löwer, Multi-Omics Characterization of the 4T1 Murine Mammary Gland Tumor Model. *Front. Oncol.* **10**, 1195 (2020).

23. M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: making protein folding accessible to all. *Nat. Methods*. **19**, 679–682 (2022).

24. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

25. L. Cao, B. Coventry, I. Goreshnik, B. Huang, W. Sheffler, J. S. Park, K. M. Jude, I. Marković, R. U. Kadam, K. H. G. Verschueren, K. Verstraete, S. T. R. Walsh, N. Bennett, A. Phal, A. Yang, L. Kozodoy, M. DeWitt, L. Picton, L. Miller, E.-M. Strauch, N. D. DeBouver, A. Pires, A. K. Bera, S. Halabiya, B. Hammerson, W. Yang, S. Bernard, L. Stewart, I. A. Wilson, H. Ruohola-Baker, J. Schlessinger, S. Lee, S. N. Savvides, K. C. Garcia, D. Baker, Design of protein-binding proteins from the target structure alone. *Nature*. **605**, 551–560 (2022).

26. E. Ruoslahti, Targeting tumor vasculature with homing peptides from phage display. *Semin. Cancer Biol.* **10**, 435–442 (2000).

27. A. J. Zurita, W. Arap, R. Pasqualini, Mapping tumor vascular diversity by screening phage display libraries. *J. Control. Release*. **91**, 183–186 (2003).

28. D. L. Sellers, J. M. Bergen, R. N. Johnson, H. Back, J. M. Ravits, P. J. Horner, S. H. Pun, Targeted axonal import (TAxI) peptide delivers functional proteins into spinal cord motor neurons after peripheral administration. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 2514–2519 (2016).

29. R. Pasqualini, E. Ruoslahti, Organ targeting in vivo using phage display peptide libraries. *Nature*. **380**, 364–366 (1996).

30. M. Sylvestre, C. P. Saxby, N. Kacherovsky, H. Gustafson, S. J. Salipante, S. H. Pun, Identification of a DNA Aptamer That Binds to Human Monocytes and Macrophages. *Bioconjug. Chem.* **31**, 1899–1907 (2020).

31. N. Kacherovsky, I. I. Cardle, E. L. Cheng, J. L. Yu, M. L. Baldwin, S. J. Salipante, M. C. Jensen, S. H. Pun, Traceless aptamer-mediated isolation of CD8+ T cells for chimeric antigen receptor T-cell therapy. *Nat Biomed Eng*. **3**, 783–795 (2019).

32. L. F. Yang, N. Kacherovsky, N. Panpradist, R. Wan, J. Liang, B. Zhang, S. J. Salipante, B. R. Lutz, S. H. Pun, Aptamer Sandwich Lateral Flow Assay (AptaFlow) for Antibody-Free SARS-CoV-2 Detection. *Anal. Chem.* **94**, 7278–7285 (2022).

33. J. Tubiana, L. Adriana-Lifshits, M. Nissan, M. Gabay, I. Sher, M. Sova, H. J. Wolfson, M. Gal, Funneling modulatory peptide design with generative models: Discovery and characterization of disruptors of calcineurin protein-protein interactions. *PLoS Comput. Biol.* **19**, e1010874 (2023).

34. J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, N. Hanikel, S. J. Pellock, A. Courbet, W. Sheffler, J. Wang, P. Venkatesh, I. Sappington, S. V. Torres, A. Lauko, V. De Bortoli, E. Mathieu, R. Barzilay, T. S. Jaakkola, F. DiMaio, M. Baek, D. Baker, Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models. *bioRxiv* (2022), p. 2022.12.09.519842.

35. H.-B. Pang, G. B. Braun, Z.-G. She, V. R. Kotamraju, K. N. Sugahara, T. Teesalu, E. Ruoslahti, A free cysteine prolongs the half-life of a homing peptide and improves its tumor-penetrating activity. *J. Control. Release*. **175**, 48–53 (2014).

36. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature*. **596**, 583–589 (2021).

37. N. Kasaraneni, A. M. Chamoun-Emanuelli, G. Wright, Z. Chen, Retargeting Lentiviruses via SpyCatcher-SpyTag Chemistry for Gene Delivery into Specific Cell Types. *MBio*. **8** (2017), doi:10.1128/mBio.01860-17.

38. A. Muik, J. Reul, T. Friedel, A. Muth, K. P. Hartmann, I. C. Schneider, R. C. Münch, C. J. Buchholz, Covalent coupling of high-affinity ligands to the surface of viral vector particles by protein trans-splicing mediates cell type-specific gene transfer. *Biomaterials*. **144**, 84–94 (2017).

39. E. N. Baker, C. J. Squire, P. G. Young, Self-generated covalent cross-links in the cell-surface

adhesins of Gram-positive bacteria. *Biochem. Soc. Trans.* **43**, 787–794 (2015).

40. P. D. Senter, E. L. Sievers, The discovery and development of brentuximab vedotin for use in relapsed Hodgkin lymphoma and systemic anaplastic large cell lymphoma. *Nat. Biotechnol.* **30**, 631–637 (2012).

41. S. Kang, M. Uchida, A. O'Neil, R. Li, P. E. Prevelige, T. Douglas, Implementation of p22 viral capsids as nanoplatforms. *Biomacromolecules*. **11**, 2804–2809 (2010).

42. J. O. Jeon, S. Kim, E. Choi, K. Shin, K. Cha, I.-S. So, S.-J. Kim, E. Jun, D. Kim, H. J. Ahn, B.-H. Lee, S.-H. Lee, I.-S. Kim, Designed nanocage displaying ligand-specific Peptide bunches for high affinity and biological activity. *ACS Nano*. **7**, 7462–7471 (2013).

43. G. Ueda, A. Antanasijevic, J. A. Fallas, W. Sheffler, J. Copps, D. Ellis, G. B. Hutchinson, A. Moyer, A. Yasmeen, Y. Tsybovsky, Y.-J. Park, M. J. Bick, B. Sankaran, R. A. Gillespie, P. J. Brouwer, P. H. Zwart, D. Veesler, M. Kanekiyo, B. S. Graham, R. W. Sanders, J. P. Moore, P. J. Klasse, A. B. Ward, N. P. King, D. Baker, Tailored design of protein nanoparticle scaffolds for multivalent presentation of viral glycoprotein antigens. *Elife*. **9** (2020), doi:10.7554/eLife.57659.

44. S. Tetter, N. Terasaka, A. Steinauer, R. J. Bingham, S. Clark, A. J. P. Scott, N. Patel, M. Leibundgut, E. Wroblewski, N. Ban, P. G. Stockley, R. Twarock, D. Hilvert, Evolution of a virus-like architecture and packaging mechanism in a repurposed bacterial protein. *Science*. **372**, 1220–1224 (2021).

45. A. Van de Steen, R. Khalife, N. Colant, H. Mustafa Khan, M. Deveikis, S. Charalambous, C. M. Robinson, R. Dabas, S. Esteban Serna, D. A. Catana, K. Pildish, V. Kalinovskiy, K. Gustafsson, S. Frank, Bioengineering bacterial encapsulin nanocompartments as targeted drug delivery system. *Synth Syst Biotechnol*. **6**, 231–241 (2021).

46. J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N. P. King, D. Baker, Robust deep learning-based protein sequence design using ProteinMPNN. *Science*. **378**, 49–56 (2022).

47. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).

48. J. Zhang, K. Kobert, T. Flouri, A. Stamatakis, PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*. **30**, 614–620 (2014).

Supporting Information

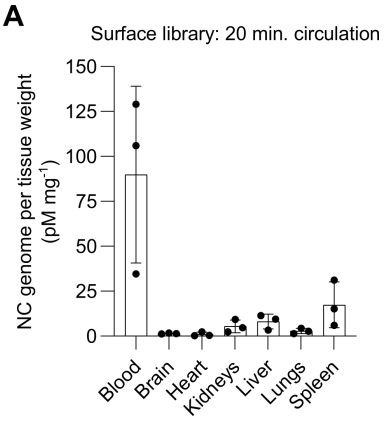


Figure S1. Biodistribution of I53-50-v4 surface library after the first round of selection.
(A) Biodistribution of the I.V.-injected surface library from selection round 1 after 20 minutes of circulation, measured by RT-qPCR and reported as picomolar (pM) of nucleocapsid (NC) genome per milligram (mg) of tissue. n=3 mice per group; error bars depict standard deviation.

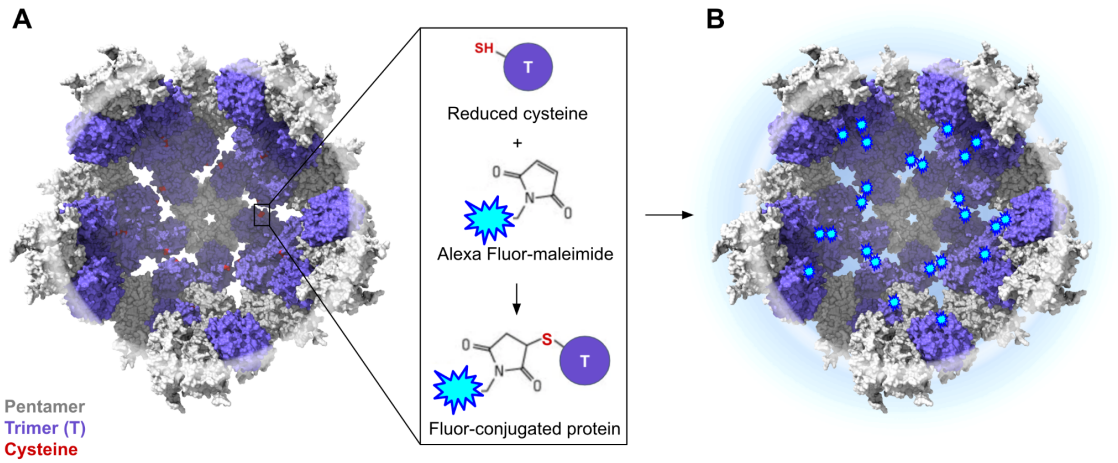


Figure S2. Maleimide conjugation scheme.
Protein nanoparticle half-shells show the interior cavity of the nanoparticle. **(A)** Unique cysteines (red) were engineered into the internal cavity of "I53-50-v4v0-Cys" (I53-50 nanoparticle, v4 evolved exterior, v0 non-mRNA encapsulating interior, with unique internal cysteines). The pores of the nanoparticle are small enough for maleimide-functionalized fluorophores to migrate in and react with the free sulfhydryls. **(B)** Depiction of I53-50-v4v0-Cys encapsulating covalently conjugated fluorophores.

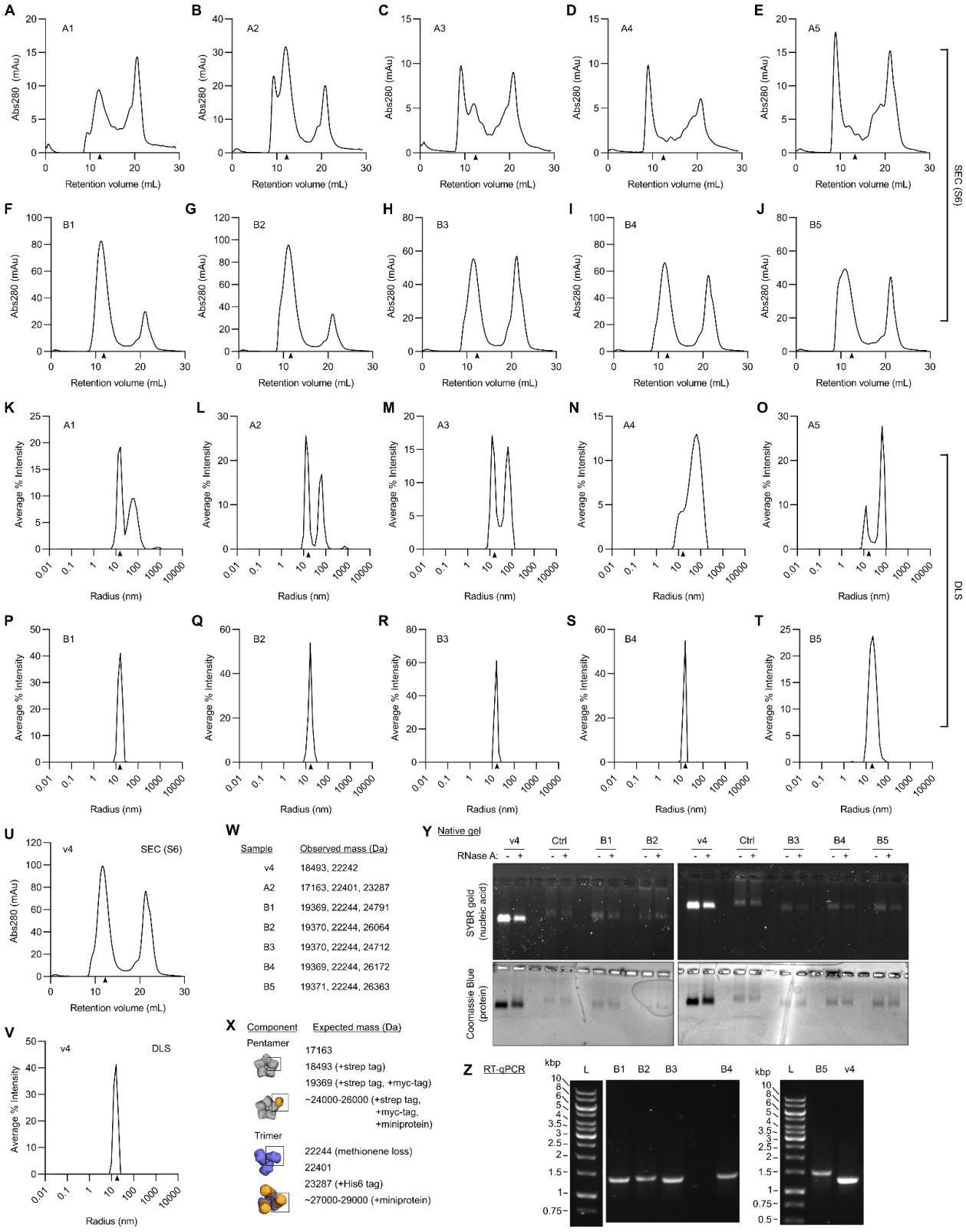


Figure S3. Miniprotein library individual mutant characterization.

Figure S3 Continued. **(A-E)** SEC of synNCs displaying one of five unique miniproteins on the N terminus of the trimer (A1-A5). Arrows on x axes at 11-12 mL denote the fractions collected for downstream analysis. The expected retention volume of the nanoparticles is 10-13 mL, and the expected retention volume of salts is 20-22 mL (conductivity peak). **(F-G)** Dynamic light scattering (DLS) of the SEC fractions collected in (A-E). **(K-O)** SEC of synNCs displaying one of five unique miniproteins on the C terminus of the pentamer at reduced valency (B1-B5). Arrows on x axes denote the fraction collected for downstream analysis. **(P-T)** DLS of the SEC fractions collected in (A-E). **(U)** SEC of I53-50-v4. **(V)** DLS of I53-50-v4. **(W)** Observed masses of samples v4, A2, B1-B5 by mass spectrometry. **(X)** Key for expected component masses. **(Y)** Native agarose gel of v4, an internal control, and B1-B5 with or without RNase A treatment prior to loading in the gel. The same gels are stained for nucleic acid (top) or protein (bottom). **(Z)** RT-qPCR of mRNA extracted from B1-B5 and v4.

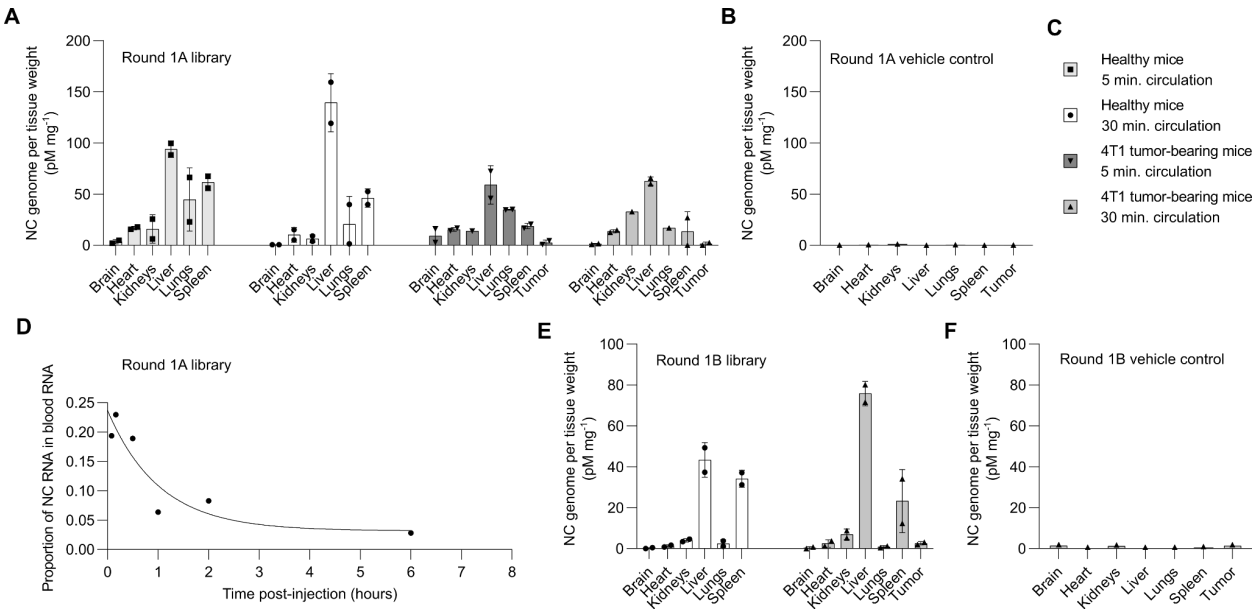
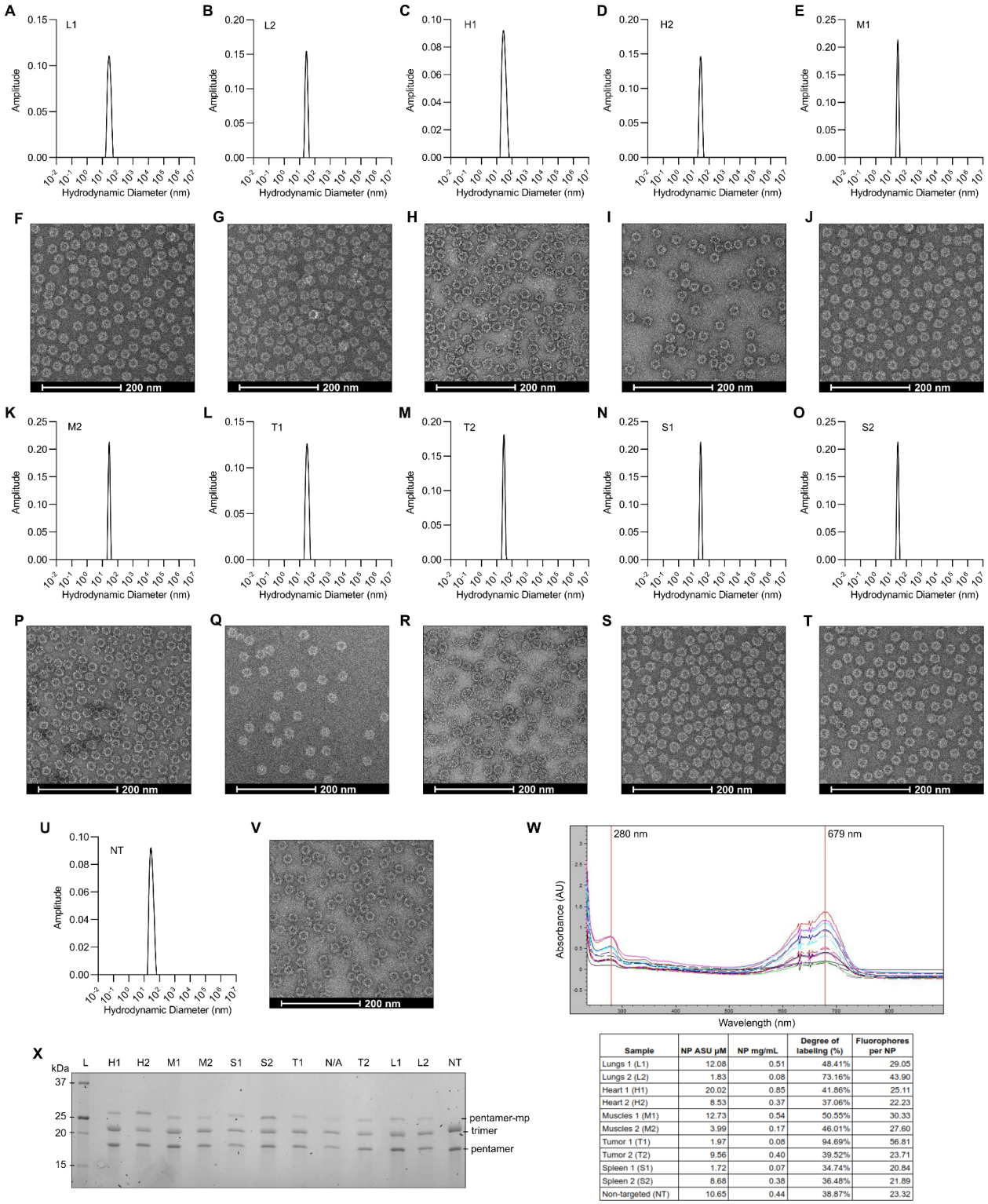


Figure S4. Miniprotein library biodistribution and presence in blood over time measured from Round 1A and Round 1B.

(A-B) Biodistribution of the I.V.-injected miniprotein library (A) or vehicle control (B) from selection round 1A after 5 or 30 minutes of circulation, measured by RT-qPCR and reported as picomolar (pM) nucleocapsid (NC) genome per milligram (mg) of tissue. n=2 mice per group in (A), n=1 mouse per group in (B); error bars depict standard error of the mean. **(C)** Symbol key. Squares: healthy mice - 5 min. circulation time, Circles: healthy mice - 30 min. circulation time, Triangles pointing down: 4T1 tumor-bearing mice - 5 min. circulation time, Triangles pointing up: 4T1 tumor-bearing mice - 30 min. circulation time. **(D)** Proportion of synNC RNA measured in blood samples over the total amount of RNA in the blood samples over time. The line plotted represents a nonlinear decay best fit. **(E)** Biodistribution of the I.V.-injected miniprotein library (E) or vehicle control (F) from selection round 1B after 30 minutes of circulation, measured by RT-qPCR and reported as picomolar (pM) nucleocapsid (NC) genome per milligram (mg) of tissue. n=2 mice per group in (E), n=1 mouse per group in (F); error bars depict standard error of the mean.



Sample	NP ASU µM	NP mg/mL	Degree of labeling (%)	Fluorophores per NP
Lungs 1 (L1)	12.08	0.51	48.41%	29.05
Lungs 2 (L2)	1.83	0.08	73.16%	43.90
Heart 1 (H1)	20.02	0.85	41.86%	25.11
Heart 2 (H2)	8.53	0.37	37.06%	22.23
Muscles 1 (M1)	12.73	0.54	50.55%	30.33
Muscles 2 (M2)	3.99	0.17	46.01%	27.60
Tumor 1 (T1)	1.97	0.08	94.69%	56.81
Tumor 2 (T2)	9.56	0.40	39.52%	23.71
Spleen 1 (S1)	1.72	0.07	34.74%	20.84
Spleen 2 (S2)	8.68	0.38	36.48%	21.89
Non-targeted (NT)	10.65	0.44	38.87%	23.32

Figure S5. Biochemical characterization of AF680-labeled nanoparticles post-freeze/thaw.

Figure S5 Continued. **(A-E)** DLS of AF680-labeled nanoparticles H1, H2, M1, M2, and L1. **(F-J)** Negative stain electron microscopy of the same nanoparticles from (A-E), respectively. **(K-O)** DLS of AF680-labeled nanoparticles L2, T1, T2, S1, and S2. **(P-T)** Negative stain electron microscopy of the same nanoparticles from (K-O),

respectively. **(U)** DLS of AF680-labeled non-targeted nanoparticles (NT). **(V)** Negative stain electron microscopy of the same nanoparticles from (U). **(W)** UV-Vis spectroscopy of AF680-labeled nanoparticles from (A-V). **(X)** Reducing SDS-PAGE of nanoparticles from (A-V).

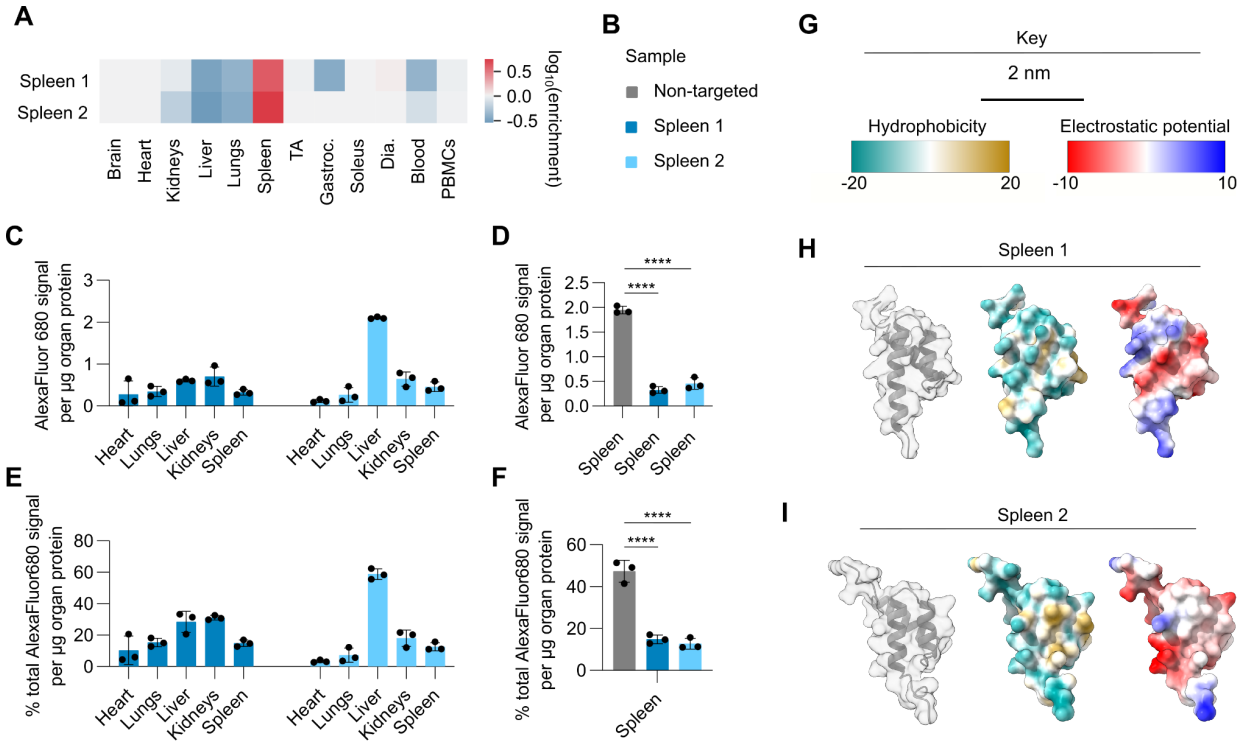


Figure S6. Library variants of interest in the spleen.
(A) Enrichment of sequences of interest in organs from healthy mice. (B) Key for C-F. (C-F) The raw (C-D) and normalized (E-F) AF680 signal per µg homogenized organ protein from mice IV-ROI-injected with AF680-labeled nanoparticles Spleen 1 (dark blue) and Spleen 2 (light blue) (% total Alexa Fluor 680 signal per µg organ protein) (C, E) and compared to non-targeted nanoparticle in the spleen only (gray, D, F). n=3 mice per group. Error bars depict standard deviations. Statistics were calculated using ordinary one-way ANOVA with Tukey's correction for multiple comparisons. **: $p \leq 0.01$, ns: not significant. Only statistically significant comparisons are indicated. (G) Key for H-I. (H-I) AlphaFold2 predictions of miniprotein structures (*23*). The secondary structure, surface hydrophobicity, and surface electrostatic potentials of each miniprotein structure were visualized in UCSF ChimeraX (*24*).

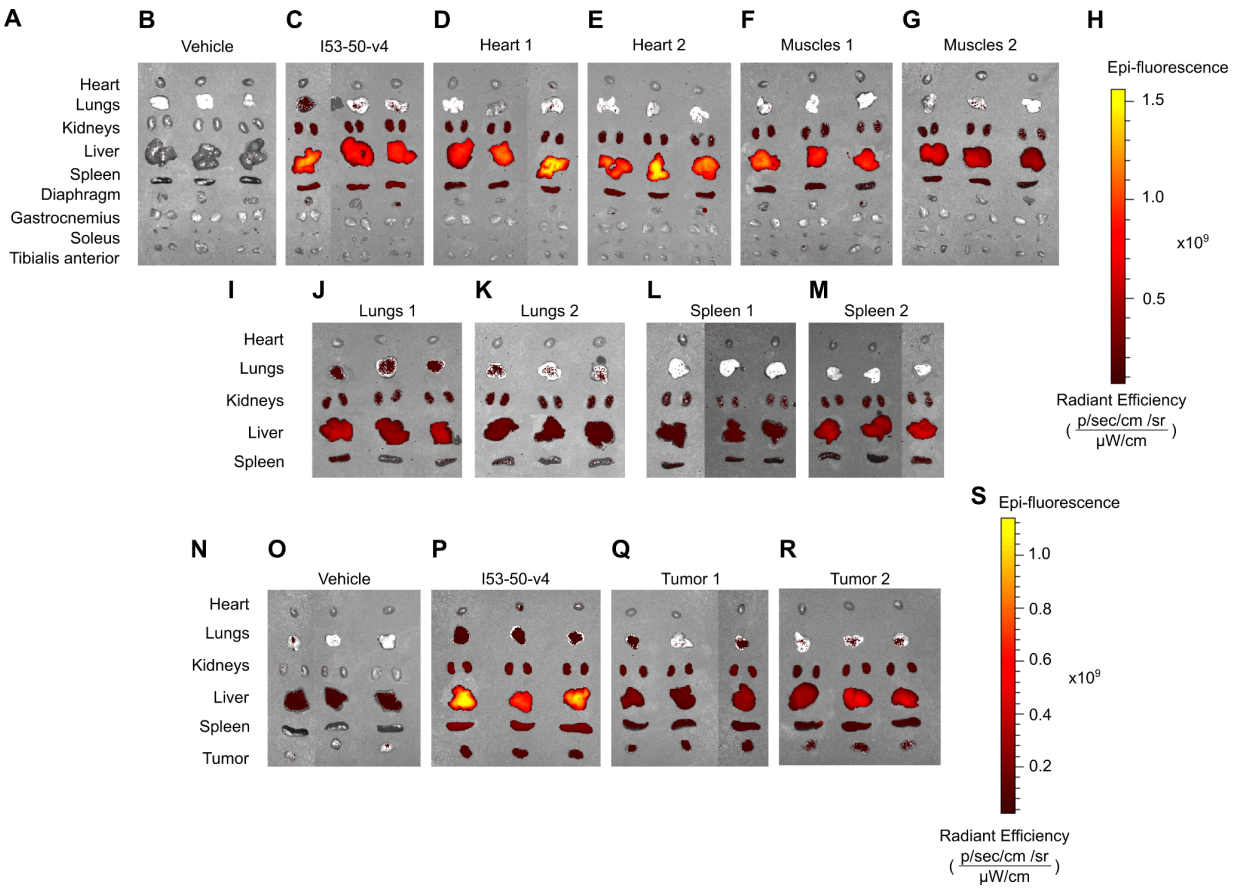


Figure S7. IVIS images of organs utilized in Figures 6 and S6.
(A-M) Organs from healthy mice I.V.-R.O.I.-injected with vehicle control (B) or AF680-labeled nanoparticles (C-G, J-M). **(A, I)** Organ keys. **(H)** Epi-fluorescence scale bar. Minimum: 6.8×10^{7}. Maximum: 1.56×10^{9}. **(N-S)** Organs from 4T1 tumor-bearing mice I.V.-R.O.I.-injected with vehicle control (O) or AF680-labeled nanoparticles (P-R). **(N)** Organ key. **(S)** Epi-fluorescence scale bar. Minimum: 2.4×10^{7}. Maximum: 1.14×10^{9}.

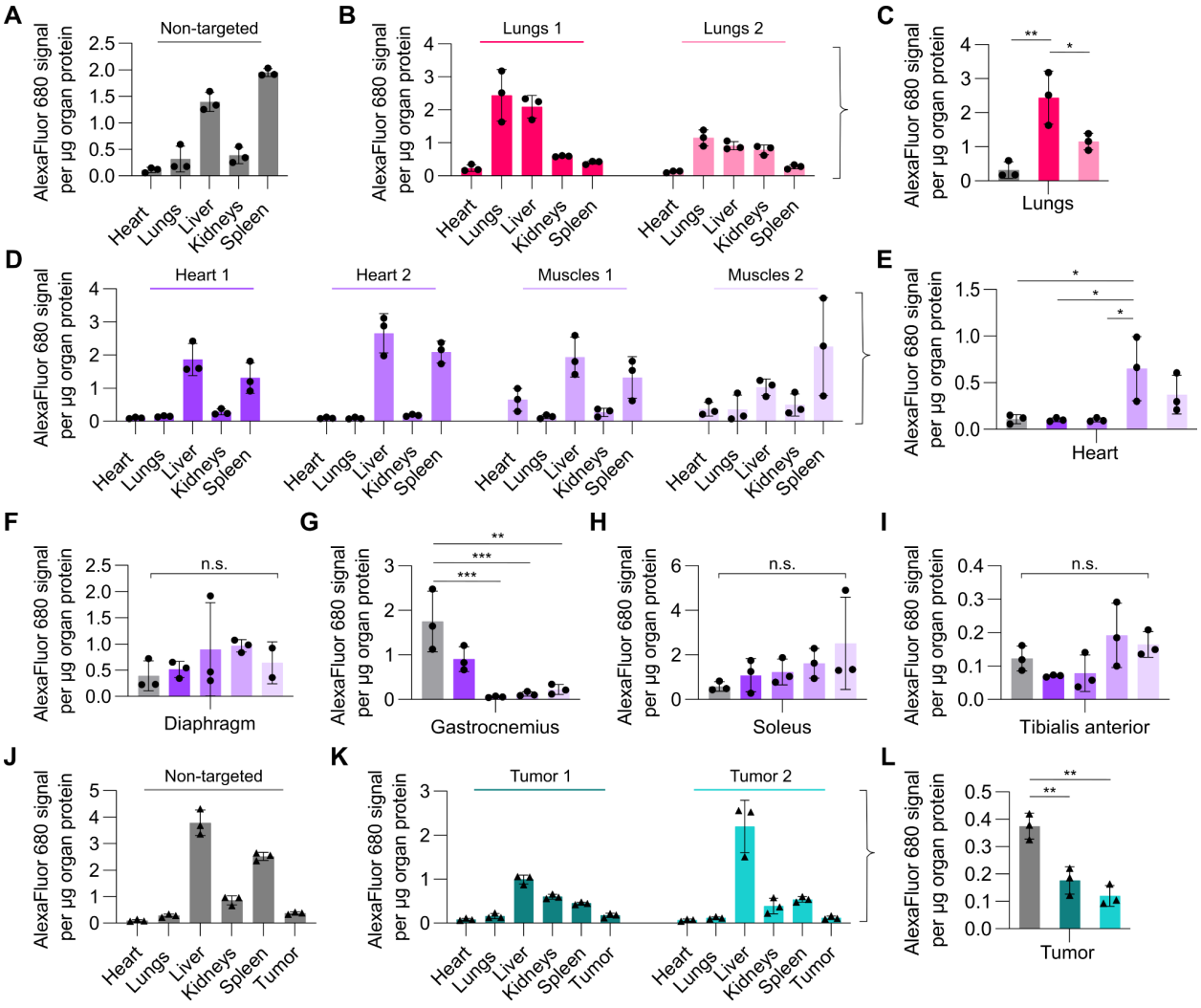


Figure S8. Fluorescence biodistribution of variants of interest (non-normalized).
Each graph shows the raw, non-normalized fluorescence signal per µg homogenized organ protein from mice injected with AF680-labeled nanoparticles (fluorescence per µg organ protein). **(A)** Non-targeted nanoparticle in healthy mice; **(B-C)** Lungs 1 and 2 library variants in all organs (D) and compared to non-targeted nanoparticle in healthy mice in the lungs only (E); **(D-E)** Heart 1, Heart 2, Muscles 1, and Muscles 2 library variants in all organs (D) and compared to non-targeted nanoparticle in healthy mice in the heart only (E); **(F-I)** Heart 1, Heart 2, Muscles 1, and Muscles 2 library variants compared to non-targeted in diaphragm (F), gastrocnemius (G), soleus (H), or tibialis anterior (I); **(J)** Non-targeted nanoparticle-AF680 in 4T1 tumor-bearing mice; **(K-L)** Tumor 1 and 2 library variants in all organs (K) and compared to non-targeted nanoparticle in the tumor only (L).. Circles depict individual values from healthy mice. Triangles depict individual values from 4T1 tumor-bearing mice. n=3 mice per group. Error bars depict standard deviations. Statistics: Ordinary one-way ANOVA with Tukey's correction for multiple comparisons. ***: $p \leq 0.001$; **: ; $p \leq 0.01$; *: $p \leq 0.05$; ns: not statistically significant. Only statistically significant comparisons are indicated in C, E-I, L unless otherwise noted.

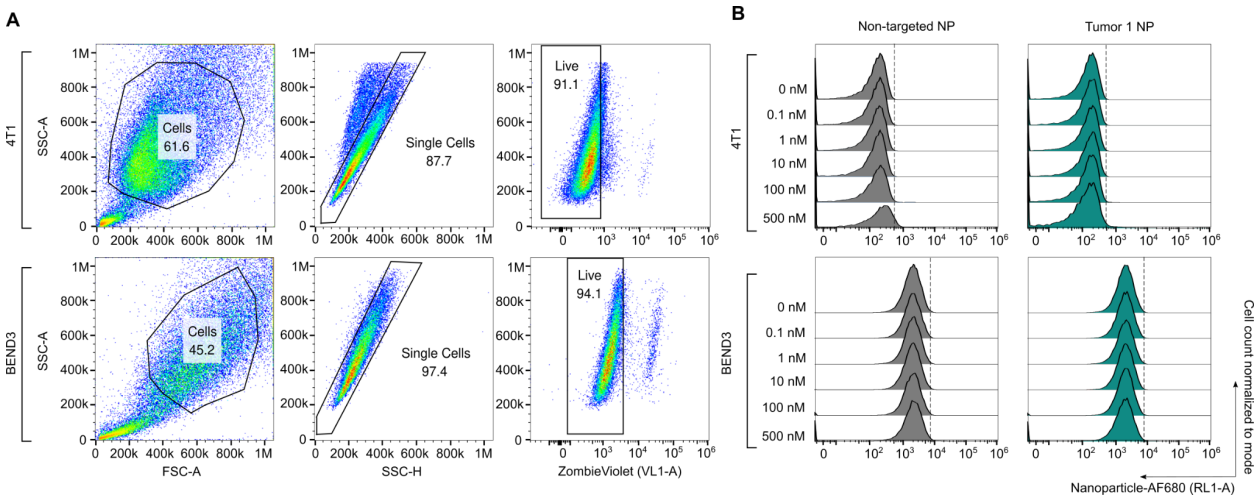


Figure S9. Flow cytometry of non-targeted or Tumor 1 AF680-labeled nanoparticles with 4T1 or BEND3 cells.

(A) Gating strategy for 4T1 cells (top) or BEND3 cells (bottom). **(B)** Titration of AF680-labeled nanoparticles incubated with 4T1 cells (top) or BEND3 cells (bottom) shows lack of binding to either cell type.

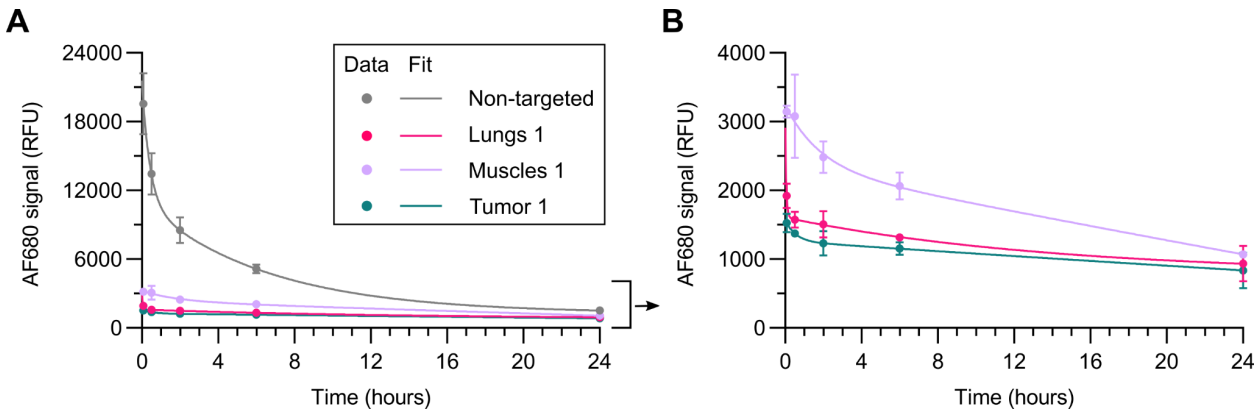


Figure S10. Serum circulation half-lives of AF680-labeled variants of interest in healthy mice.
(A) AF680 signal (relative fluorescence units, RFU) in serum over time of AF680-labeled nanoparticles (Non-targeted, Lungs 1, Muscles 1, and Tumor 1). **(B)** Zoomed in view of Lungs 1, Muscles 1, and Tumor 1 from (A). Corresponding values are in **Tab. S3**.

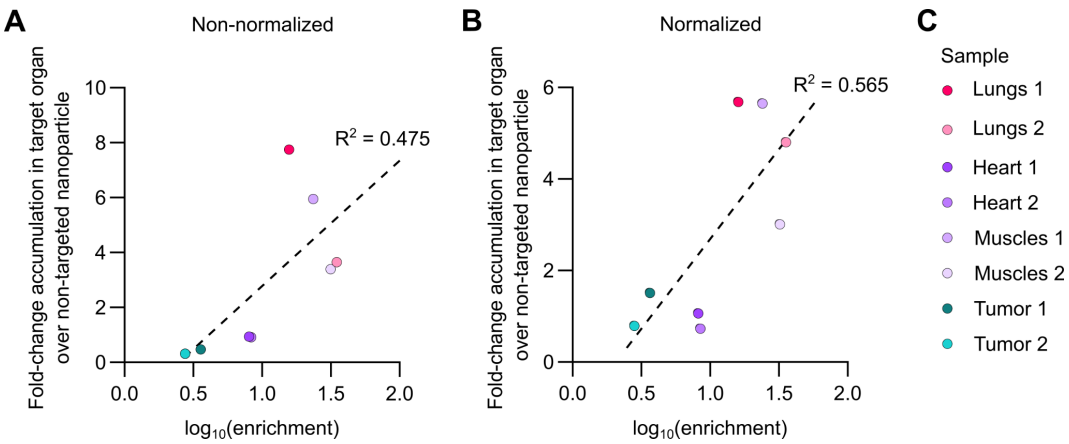


Figure S11. Correlation between enrichment score and relative organ accumulation fold-change compared to non-targeted nanoparticles.
(A-B) Log_{10}(enrichment) of specific sequence in target organ from mRNA-Seq (**Fig. 4D-E**, **Eq. 2**) versus fold-change non-normalized **(A)** and normalized **(B)** accumulation in target organs over non-targeted nanoparticles by fluorescence biodistribution show positive correlation (fluorescence biodistribution data from **Fig. S8** and **Fig. 6**, respectively). The dashed line is a linear fit to the data.

Table S1. DNA sequences from T7 promoter to T7 terminator in pet29b(+) plasmid in FASTA format and IUPAC nucleotide code.

Key:
Trimer
Trimer helical extension
Pentamer
Miniprotein adapters
--- Miniprotein placeholder
prfB
Bolded: Surface library mutation sites (encoding E243D or D100K)

>I53-50-v4_library_with_surface_variants
CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATATGACGATGGAGGAGCTATTCAAGCGCCACACGATAGTGGCAGTGCTGAGGGCGAATAGC
GTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTACCGTGCCGG
ATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGGATGGCGCGATTATTGGTGCGGGCACCGTTACCAGTGT
GVWACAGTGCVWAAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAAVWAATTAGCVWATTC
TGCVWAGAAVWAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCTGGGCCATG
ATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCGAATGTGAA
GTTCGTGCCAACCGGCGGTGTTAACCTGGATAATGTGTGCAAATGGTTTAAAGCCGGCGTGCTGGCGGTGGGCGTGGGC
AACGCCTTAGTTAAGGGCAACCCGGACAAAGTGCGTGAAAAAGCGAAAAAGTTTGTTAAGAAGATCCGTGGCTGCACCG
AAGGTTCTTAAAGAAGGAGATATCATATGAACCAGCACAGTMAYAAGGACMAYGAAACGGTAAGAATAGCGGTAGTGCG
TGCGSRRTGGCATGCGTTTATTGTGGACGCCTGTGTGAGTGCGTTTGAAGCCGCGATGCGTRANATTGGCGGCRANSRR
TTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAAAACCGGCSAAT
ATGGCGCGGTGTTAGGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGCGCGGTTATTGA
TGGCATGATGAATGTGCAGCTGMAAACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACTACGACAAGAGC
AACGCCAAAACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCTGCGTGGAGATTTTAG
CGGCCCGTGAAAAGATTGCGGCGGGTTCTCTCGAGCACCATCATCACCATCACTGAGCCTAGCAACCAACAGTATGGAT
CCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGG
CCTCTAAACGGGTCTTGAGGGGTTTTTTG

>I53-50-v4v0-Cys
CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATATGAAGATGGAAGAGCTGTTTAAGAAGCATAAAATAGTGGCAGTGCTGAGGGCGAATAGC
GTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTACCGTGCCGG
ATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGAAAGGCGCGATTATTGGTGCGGGCACCGTTACCAGTGT
GGATCAGGCCCGTAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAAGAAATTAGCCAGTTC
GCCAAGGAAAAAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCTGGGCCATG
ATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCGAATGTGAA
GTTCGTGCCAACCGGCGGTGTTTGCCTGGATAATGTGGCCGAGTGGTTTAAAGCCGGCGTGCTGGCGGTGGGCGTGGGC
AGCGCCTTAGTTAAGGGCACCCCGGACGAGGTGCGTGAAAAAGCGAAAGCGTTTGTTGAGAAGATCCGTGGCGCCACCG
AAGGTTCTTAAAGAAGGAGATATCATATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGG
CAGCGAACAAAAACTTATAAGTGAAGAGGACTTAGGCAGCAACCAGCACAGTCAGAAGGACCAGGAAACGGTAAGAATA
GCGGTAGTGCGTGCGCGCTGGCATGCGGAGATTGTGGACGCCGCTGTGAGTGCGTTTGAAGCCGCGATGCGTAAAATTG
GCGGCGAACGTTTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAA
AACCGGCCGTTATGGCGCGGTGTTAGGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGC
GCGGTTATTGATGGCATGATGAATGTGCAGCTGGATACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACT

ACGACGATAGCGATGCCCATACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCGCGGT
GGAGATTTTAGCGGCCCGTGAAAAGATTGCGGCGGGTTCTTGAGCCAATCTCGAGCACCATCATCACCATCACTGAGCC
TAGCAACCAACAGTATGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

>I53-50-v4v0-Cys_**E243D**

CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATATGAAGATGGAAGAGCTGTTTAAGAAGCATAAAATAGTGGCAGTGCTGAGGGCGAATAGC
GTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTACCGTGCCGG
ATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGAAAGGCGCGATTATTGGTGCGGGCACCGTTACCAGTGT
GGATCAGGCCCGTAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAA**GAT**ATTAGCCAGTTC
GCCAAGGAAAAAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCTGGGCCATG
ATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCGAATGTGAA
GTTCGTGCCAACCGGCGGTGTTTGCCTGGATAATGTGGCCGAGTGGTTTAAAGCCGGCGTGCTGGCGGTGGGCGTGGGC
AGCGCCTTAGTTAAGGGCACCCCGGACGAGGTGCGTGAAAAAGCGAAAGCGTTTGTTGAGAAGATCCGTGGCGCCACCG
AAGGTTCTTAAAGAAGGAGATATCATATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGG
CAGCGAACAAAAACTTATAAGTGAAGAGGACTTAGGCAGCAACCAGCACAGTCAGAAGGACCAGGAAACGGTAAGAATA
GCGGTAGTGCGTGCGCGCTGGCATGCGGAGATTGTGGACGCCGCTGTGAGTGCGTTTGAAGCCGCGATGCGTAAAATTG
GCGGCGAACGTTTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAA
AACCGGCCGTTATGGCGCGGTGTTAGGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGC
GCGGTTATTGATGGCATGATGAATGTGCAGCTGGATACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACT
ACGACGATAGCGATGCCCATACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCGCGGT
GGAGATTTTAGCGGCCCGTGAAAAGATTGCGGCGGGTTCTTGAGCCAATCTCGAGCACCATCATCACCATCACTGAGCC
TAGCAACCAACAGTATGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATA
ACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

>I53-50-v4v0-Cys_**D100K**

CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATATGAAGATGGAAGAGCTGTTTAAGAAGCATAAAATAGTGGCAGTGCTGAGGGCGAATAGC
GTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTACCGTGCCGG
ATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGAAAGGCGCGATTATTGGTGCGGGCACCGTTACCAGTGT
GGATCAGGCCCGTAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAAGAAATTAGCCAGTTC
GCCAAGGAAAAAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCTGGGCCATG
ATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCGAATGTGAA
GTTCGTGCCAACCGGCGGTGTTTGCCTGGATAATGTGGCCGAGTGGTTTAAAGCCGGCGTGCTGGCGGTGGGCGTGGGC
AGCGCCTTAGTTAAGGGCACCCCGGACGAGGTGCGTGAAAAAGCGAAAGCGTTTGTTGAGAAGATCCGTGGCGCCACCG
AAGGTTCTTAAAGAAGGAGATATCATATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGG
CAGCGAACAAAAACTTATAAGTGAAGAGGACTTAGGCAGCAACCAGCACAGTCAGAAGGACCAGGAAACGGTAAGAATA
GCGGTAGTGCGTGCGCGCTGGCATGCGGAGATTGTGGACGCCGCTGTGAGTGCGTTTGAAGCCGCGATGCGTAAAATTG
GCGGCGAACGTTTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAA
AACCGGCCGTTATGGCGCGGTG**AAA**GGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGC
GCGGTTATTGATGGCATGATGAATGTGCAGCTGGATACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACT
ACGACGATAGCGATGCCCATACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCGCGGT
GGAGATTTTAGCGGCCCGTGAAAAGATTGCGGCGGGTTCTTGAGCCAATCTCGAGCACCATCATCACCATCACTGAGCC
TAGCAACCAACAGTATGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATA
ACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

>I53-50-v4_trimer_miniprotein_display

CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATGGCGGAGGGTCGGCTTCGCATATG---CTCGAGGGTGGAGGTTCCGAACAAAAGgaaaag
gcggcgaaagcggaggaagcagcgcgtATGACGATGGAGGAGCTATTCAAGCGCCACACGATAGTGGCAGTGCTGAGGG
CGAATAGCGTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTAC
CGTGCCGGATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGGATGGCGCGATTATTGGTGCGGGCACCGTT
ACCAGTGTGGATCAGTGCCGTAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAAGAAATTA
GCCAGTTCTGCAAGGAAAAAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCT
GGGCCATGATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCG
AATGTGAAGTTCGTGCCAACCGGCGGTGTTAACCTGGATAATGTGTGCAAATGGTTTAAAGCCGGCGTGCTGGCGGTGG
GCGTGGGCAACGCCTTAGTTAAGGGCAACCCGGACAAAGTGCGTGAAAAAGCGAAAAAGTTTGTTAAGAAGATCCGTGG
CTGCACCGAAGGTTCTTGGTCTCATCCACAGTTCGAAAAGTAAAGAAGGAGATATCATATGGGCAGCAGCCATCATCAT
CATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCAACCAGCACAGTCARAAGGACCARGAAACGGTAAGAATAGCGG
TAGTGCGTGCGCGCTGGCATGCGTTTATTGTGGACGCCTGTGTGAGTGCGTTTGAAGCCGCGATGCGTAAAATTGGCGG
CGAACGTTTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAAAACC
GGCCGCTATGGCGCGGTGTTAGGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGCGCGG
TTATTGATGGCATGATGAATGTGCAGCTGGATACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACTACGA
CAAGAGCAACGCCAAAACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCTGCGTGGAG
ATTTTAGCGGCCCGTGAAAAGATTGCGGCGGGTTCTTGAGCCAATCTCGAGCACCATCATCACCATCACTGAGCCTAGC
AACCAACAGTATGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTA
GCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

>I53-50-v4_prfB_pentamer_miniprotein_display

CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATATGACGATGGAGGAGCTATTCAAGCGCCACACGATAGTGGCAGTGCTGAGGGCGAATAGC
GTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTACCGTGCCGG
ATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGGATGGCGCGATTATTGGTGCGGGCACCGTTACCAGTGT
GGATCAGTGCCGTAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAAGAAATTAGCCAGTTC
TGCAAGGAAAAAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCTGGGCCATG
ATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCGAATGTGAA
GTTCGTGCCAACCGGCGGTGTTAACCTGGATAATGTGTGCAAATGGTTTAAAGCCGGCGTGCTGGCGGTGGGCGTGGGC
AACGCCTTAGTTAAGGGCAACCCGGACAAAGTGCGTGAAAAAGCGAAAAAGTTTGTTAAGAAGATCCGTGGCTGCACCG
AAGGTTCTTAAAGAAGGAGATATCATATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCGAACAAAAACTTATAAGTGAAGAGGACTTAGGCAGCAACCAGCACAGTCAGAAGGACCAGGAAACGGTAAGAATA
GCGGTAGTGCGTGCGCGCTGGCATGCGTTTATTGTGGACGCCTGTGTGAGTGCGTTTGAAGCCGCGATGCGTAAAATTG
GCGGCGAACGTTTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAA
AACCGGCCGCTATGGCGCGGTGTTAGGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGC
GCGGTTATTGATGGCATGATGAATGTGCAGCTGGATACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACT
ACGACAAGAGCAACGCCAAAACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCTGCGT
GGAGATTTTAGCTGCCCGTGAAAAGATTGCGGCGGGTTCTCTcGAGGGTTCTAGGGGGTATCTTTGACGGCTCCGGTTC
CGGTTCTGGCGGAGGGTCGGCTTCGCATATG---CTCGAGGGTGGAGGTTCCGAACAAAAGTAACTCGAGCACCATCAT
CACCATCACTGAGCCTAGCAACCAACAGTATGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGC
CACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

>I53-50-v4v0-Cys_prfB_pentamer_miniprotein_display

CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGATAGGATTACTGCTCGGTG
ACAAGAAGGAGATATACATATGACGATGGAGGAGCTATTCAAGCGCCACACGATAGTGGCAGTGCTGAGGGCGAATAGC
GTGGAGGAAGCGATTGAGAAGGCCGTGGCCGTGTTCGCCGGTGGCGTTCATCTGATTGAAATTACCTTTACCGTGCCGG
ATGCGGATACCGTGATTAAGGCGCTGAGTGTGTTAAAAGAGGATGGCGCGATTATTGGTGCGGGCACCGTTACCAGTGT
GGATCAGTGCCGTAAAGCGGTGGAAAGTGGCGCCGAATTTATCGTTAGCCCGCATCTGGATGAAGAAATTAGCCAGTTC

TGCAAGGAAAAAGGCGTGTTCTATATGCCAGGCGTTATGACCCCAACGGAACTGGTTAAAGCGATGAAGCTGGGCCATG
ATATCCTGAAGTTATTTCCAGGTGAAGTGGTTGGTCCGCAGTTTGTGAAAGCCATGAAGGGCCCGTTCCCGAATGTGAA
GTTCGTGCCAACCGGCGGTGTTAACCTGGATAATGTGTGCAAATGGTTTAAAGCCGGCGTGCTGGCGGTGGGCGTGGGC
AACGCCTTAGTTAAGGGCAACCCGGACAAAGTGCGTGAAAAAGCGAAAAAGTTTGTTAAGAAGATCCGTGGCTGCACCG
AAGGTTCTTAAAGAAGGAGATATCATATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGG
CAGCGAACAAAAACTTATAAGTGAAGAGGACTTAGGCAGCAACCAGCACAGTCAGAAGGACCAGGAAACGGTAAGAATA
GCGGTAGTGCGTGCGCGCTGGCATGCGTTTATTGTGGACGCCTGTGTGAGTGCGTTTGAAGCCGCGATGCGTAAAATTG
GCGGCGAACGTTTTGCCGTGGATGTGTTTGATGTGCCGGGTGCGTACGAAATTCCACTGCATGCGCGTACCCTGGCGAA
AACCGGCCGCTATGGCGCGGTGTTAGGCACCGCCTTTGTGGTGAATGGTGGCATTTATCGTCACGAATTTGTGGCGAGC
GCGGTTATTGATGGCATGATGAATGTGCAGCTGGATACGGGCGTGCCAGTGTTAAGTGCCGTGCTGACCCCACACAACT
ACGACAAGAGCAACGCCAAAACCCTGCTGTTCTTAGCGCTGTTTGCGGTGAAAGGCATGGAAGCGGCGCGTGCCTGCGT
GGAGATTTTAGCTGCCCGTGAAAAGATTGCGGCGGGTTCTCTcGAGGGTTCTAGGGGGTATCTTTGACGGCTCCGGTTC
CGGTTCTGGCGGAGGGTCGGCTTCGCATATG---CTCGAGGGTGGAGGTTCCGAACAAAAGTAACTCGAGCACCATCAT
CACCATCACTGAGCCTAGCAACCAACAGTATGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGC
CACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

Table S2. Miniprotein sequences of interest

Miniprotein of interest	**Amino acid sequence** (excluding adapters from Tab. S1)
Lungs 1	PLIARLIAKTLLDLGFSEDQISEDLKRAGIDESEIRDALRYAR
Lungs 2	AKEVAVRLALAAIRFGDERFVQDAAQAANISPEELKKLARQVE
Heart 1	TLFQEDAIDFIRLLGYSEEEARRILKKANGNPLLAHEIAHRIG
Heart 2	PAADLAFYAWFADANGDRKTAKELLEKAGYSPEEAEQIIKDIE
Muscles 1	TDDQELAIFLIRGLGYSEEEARRILKKANGNPYLAREIADRIG
Muscles 2	TVDQEVAIVFIRHLGYSEEEARRILKKANGNPYLADEIARRIG
Tumor 1	TRFQELAIRRIRVLGYSEEEARRILKKANGNPDLAGEIAYRIG
Tumor 2	DKDSVEAAIELLLAAGDEDSAREAAKEAGLTEEDLRRLKKRVS
Spleen 1	RQEDAKKEAERAIRSGNDTKAFEILWHAGFDAFEALVIAEKLK
Spleen 2	TDEELKEAAKKLASAGNEEAAKLALLAAGLSPQQVEEWLQKFH

Table S3. Values from two phase, non-linear decay fit of serum circulation half-lives of AF680-labeled variants of interest in healthy mice (from Fig. S9).

Two phase decay	Parameter	Non-targeted NP-AF680	M1-AF680	L1-AF680	T1-AF680
Best-fit values	Half Life (Slow, Elimination)	4.518 h	239.6 h	9.869 h	214.5 h
	Half Life (Fast, Distribution)	0.2859 h	1.263 h	0.04086 h	0.3739 h
95% confidence interval	Half Life (Slow, Elimination)	1.512 to ? h	2.743 to ? h	2.256 to ? h	1.627 to ? h
	Half Life (Fast, Distribution)	? to 0.9638 h	? to 16.78 h	? to 5.193 h	? h
Goodness of Fit	R squared	0.962	0.8991	0.8394	0.7649

Table S4. Endotoxin values.

Sample	EU per mg	EU per dose
Surflib round 1 library	Exact values were not recorded. <1000 EU per mg was required at the time.	
Surflib round 2 library		
I53-50-v4v0-Cys-D100K_AF680		
I53-50-v4v0-Cys-E243D_AF680		
I53-50-v4v0-Cys_AF680		
Miniprotein library 1A	602.8	60.28
Miniprotein library 1B	<1000	<100
Miniprotein library 2	783.4	117.5
I53-50-v4v0-Cys-prfB-heart1_AF680	133.9	10.7
I53-50-v4v0-Cys-prfB-heart2_AF680	285.4	22.8
I53-50-v4v0-Cys-prfB-muscles1_AF680	110.3	8.8
I53-50-v4v0-Cys-prfB-muscles2_AF680	341.2	27.3
I53-50-v4v0-Cys-prfB-spleen1_AF680	673.8	53.9
I53-50-v4v0-Cys-prfB-spleen2_AF680	254.6	20.4
I53-50-v4v0-Cys-prfB-tumor1_AF680	387.9	31.0
I53-50-v4v0-Cys-prfB-tumor2_AF680	302.2	24.2
I53-50-v4v0-Cys-prfB-lungs1_AF680	256.6	20.5
I53-50-v4v0-Cys-prfB-lungs2_AF680	578.4	46.3
I53-50-v4v0-Cys_AF680 (non-targeted)	332.2	26.6

Chapter 3. Modular reprogramming of designed protein nanoparticles to selectively target solid tumors

Audrey Ellen Olshefsky*, Gabriel L. Butterfield*, Marc J. Lajoie*, Elizabeth Leaf, Catherine Treichel, Elise Nunes, Donna Cvet, Dong Mei Zhang, Xinting Li, Anindya Roy, Samuel P. Wrenn, Rebecca Cole, Rashmi Ravichandran, Sharon Ke, Justin Deccareau, Meilyn R. Sylvestre, Dan Ellis, Una Natterman, Issa Yousif, Natalie Brunette, Lauren Carter, Matthew Jones, Suzie H. Pun, David Baker, Adnan Abu-Yousif & Neil P. King

* Co-first authors.

Abstract

Targeting specific cell populations *in vivo* is a critical capability of any delivery platform and remains challenging for both synthetic and viral delivery systems. Our groups previously evolved a protein nanoparticle that encapsulates its own mRNA genome (i.e., synthetic nucleocapsid) to exhibit a similar *in vivo* circulation time and genome packaging efficiency as viruses, while maintaining a level of modularity resembling that of synthetic nanoparticles. Here we demonstrate synthetic nucleocapsids can be modularly retargeted by simple genetic fusion of a binding domain targeting a cell surface receptor. Further, the cargo can be modularly reassigned by mutation of interior surface residues to allow either RNA packaging or chemical conjugation to small molecules. The modularity of both the cargo and the binding domain positions synthetic nucleocapsids as a promising platform technology for targeted delivery.

Introduction

Targeting cells of interest based on surface markers is a fundamental feature of effective biomedical technologies. Solving this problem in a modular and robust way is the key to therapeutics ranging from T cells modified with chimeric antigen receptors to antibody drug conjugates, which allow straightforward targeting of cytotoxins to tumor cells and tumor-infiltrating leukocytes (*1*). Viruses and nanoparticles are particularly promising platforms for therapeutic delivery due to their amenability to engineering characteristics of interest, such as cargo encapsulation and *in vivo* tissue homing (*2*–*5*). However, the problem of targeted delivery has proven challenging for both viral vectors and lipid and polymeric delivery systems (*6*). Although examples of AAV capsids modified with targeting ligands exist, most AAV vector retargeting has still required significant effort due to the structural and functional complexity of viruses (*7*, *8*). Similarly, non-viral nanoparticles have relied heavily on passive targeting or local nanoparticle delivery (e.g., intratumorally injected) rather than active targeting through a receptor-ligand interaction (*9*, *10*). There remains a need for nanoparticle-based delivery platforms capable of both active targeting and drug loading in a modular, straightforward manner.

Self-assembling protein nanoparticles are a growing class of nanomaterials that possess favorable characteristics for drug delivery applications (*11*, *12*). Self-assembling protein nanoparticles, herein referred to as protein nanoparticles, are built from either naturally existing or *de novo* designed symmetric oligomers with protein-protein interfaces that cause the formation of atomically precise assemblies (e.g., ferritin, encapsulin, lumazine synthase, and computationally designed nanoparticles). These capsid-like assemblies can be altered through either rational design or evolution to introduce functionalities to the exterior surface or interior cavity. We recently reported the design and evolution of an I53-50-based self-assembling protein nanoparticle that encapsulates its own mRNA genome via electrostatic protein-mRNA interactions (*13*). The evolved version of this synthetic nucleocapsid (synNC), I53-50-version 4 or -v4, efficiently protects encapsidated mRNA and exhibits an *in vivo* circulation

half-life of over 4.5 hours in BALB/c mice. However, intravenously injected I53-50-v4 did not exhibit any specific tissue accumulation, so we set out to develop a modular platform for displaying targeting domains on synNCs that could enable specific localization to cell types and tissues of interest. We identified 6 previously validated binders against tumor-associated markers with various topologies: a de novo designed miniprotein binder against αvβ6 integrin (dnαvβ6), an anti-EGFR designed ankyrin repeat protein (DARPin), and an anti-Her2 DARPin (*14–16*). We also developed a version of the evolved synthetic nucleocapsid that does not encapsulate nucleic acid, but rather contains unique, internally facing cysteines that are available to participate in maleimide conjugation to small molecules like fluorophores, monomethyl auristatin D (MMAD), and monmethyl auristatin E (MMAE). Encapsulating fluorophores enables visualization of cell targeting both *in vitro* and *in vivo*, while encapsulation of MMAD or MMAE leads to the functional output of cell cytotoxicity by blocking the polymerization of tubulin in cells (*17–20*).

It has been previously demonstrated that protein nanoparticle-based therapeutic delivery vehicles can be built by fusing molecules such as EGFR binding domains to target over-expressing tumor cells and loading cargoes such as doxorubicin to kill cancer cells (*21*, *22*). However, these protein nanoparticle therapeutics have yet to reach the clinic due to manufacturing or toxicity challenges. It is incompletely explored whether different protein nanoparticle platforms could be more amenable to *in vivo* solid tumor delivery of chemotherapeutics. Here, we work towards tumor-targeted MMAD or MMAE delivery using a version of I53-50, a similar version of which is approved for use in humans for vaccine applications (*23*).

Results

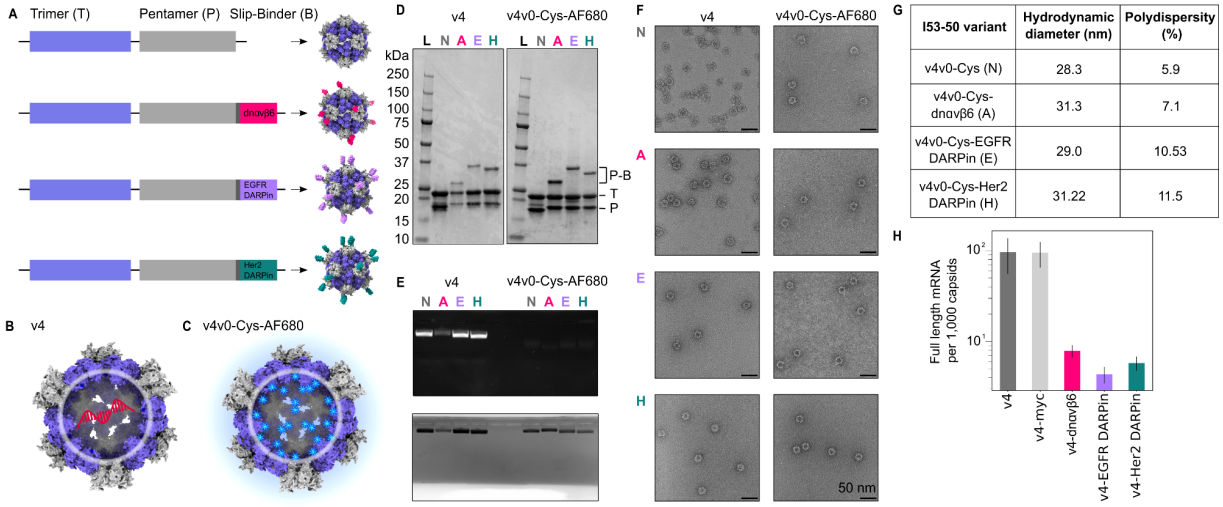
The structural integrity of synthetic nucleocapsids is retained upon surface and cargo modifications

We designed targeted synNCs that display EGFR, Her2, or αvβ6 integrin targeting domains by genetic fusion to the C-terminus of I53-50-v4B, the pentameric subunit of the two-component protein nanoparticle (**Fig. 1A**) (*14–16*). The C-terminus of the pentamer faces outward from the synNCs, allowing the binders to interact with their targets. We also sought to modulate display valency to reduce potential aggregation and nonspecific binding. The C-terminal genetic fusion enables the incorporation of a programmed ribosomal frameshifting sequence (prfB), which shifts the open reading frame +1 to avoid a stop codon and instead introduce a flexible polypeptide linker leading into the subsequent encoded protein (*24*, *25*). The resulting targeted synNCs outwardly display binders at partial valency.

To enable the encapsulation of additional classes of molecules, we designed synNCs with interior surfaces that do not associate with mRNA (v4 exterior and v0 interior, termed v4v0 synNCs) (**Fig. 1B-C**). We mutated all pre-existing cysteines to alanines and introduced a unique cysteine at the C-terminus of I53-50-v4v0A, the trimer subunit of the two-component protein nanoparticle. The resulting v4v0-Cys synNCs contain cysteines on the interior surface that are available to participate in maleimide conjugation to small molecules.

All proteins were produced in *E. coli* and purified by affinity chromatography and size-exclusion chromatography (*13*). By SDS-PAGE, we observed the binding domains fused at partial valency for both v4 and v4v0 and calculated about 40% display valency in ImageJ (**Fig. 1D**). By native agarose gel after RNase treatment, all v4 synNCs appear to encapsulate and protect nucleic acids, while v4v0-Cys synNCs do not exhibit association with nucleic acids (**Fig. 2E**). We also observed efficient covalent conjugation to maleimide-functionalized AlexaFluors. All v4- and v4v0-Cys synNCs formed intact nanoparticles of the expected sizes, as observed by electron microscopy and dynamic light scattering (**Fig. 1F-G**).

Furthermore, By RT-qPCR, we measured successful encapsidation of full-length genomes in both the presence and absence of targeting domain display, although we observed an order of magnitude decrease in the packaging efficiency of v4-targeted synNCs compared to v4-non-targeted synNCs (**Fig. 1H**). These data demonstrate the ability to modularly design and functionalize both exterior and interior synNC surfaces. Overall, the structural integrity of the synthetic nucleocapsids appears to be retained upon surface and cargo modifications.



I53-50 variant	Hydrodynamic diameter (nm)	Polydispersity (%)
v4v0-Cys (N)	28.3	5.9
v4v0-Cys-dnαvβ6 (A)	31.3	7.1
v4v0-Cys-EGFR DARPin (E)	29.0	10.53
v4v0-Cys-Her2 DARPin (H)	31.22	11.5

Figure 1. Genetic and biochemical characterization of nanoparticles encapsulating mRNA or AlexaFluor680. **(A)** To modularly retarget I53-50-v4 to specific cancer cell receptors, we fused several different binding domains to the C-terminus of the pentamer component. We incorporated a ribosomal slip sequence, resulting in partial valency of the targeting domains. **(B-C)** The nanoparticle interior can be mutated for mRNA encapsulation (B) or small molecule encapsulation via maleimide conjugation to a unique cysteine (C). **(D)** The protein subunits are the expected size as measured by SDS-PAGE for both mRNA-encapsulating nanoparticles ("v4" interior, left gel) and AlexaFluor 680-encapsulating nanoparticles ("v4v0-Cys" interior, right gel). L: ladder, N: non-targeted nanoparticle, A: αvβ6-targeted nanoparticle, E: EGFR-targeted nanoparticle, H: Her2-targeted nanoparticle. N, A, E, H abbreviations used in subsequent panels. **(E)** Native agarose gel of mRNA-encapsulating nanoparticles ("v4" interior, left 4 samples) and AlexaFluor 680-encapsulating nanoparticles ("v4v0-Cys" interior, right 4 samples) stained for both nucleic acid (top) and protein (bottom). **(F)** Negative stain transmission electron micrographs of mRNA-encapsulating nanoparticles ("v4" interior, left gel) and AlexaFluor 680-encapsulating nanoparticles ("v4v0-Cys" interior, right gel) nanoparticles (scale bars = 50 nm). **(G)** Hydrodynamic diameters and polydispersity of AlexaFluor 680-encapsulating nanoparticles ("v4v0-Cys" interior) measured by dynamic light scattering. **(H)** Full-length self-mRNA genome packaging efficiency of nanoparticles with v4 interior, displaying no targeting domain (dark gray), myc-tag (light gray), dnαvβ6 binder (pink), EGFR DARPin (blue), or Her2 DARPin (yellow). n=2 experimental replicates, error bars denote standard error of the mean.

Targeted nanoparticles selectively associate with target receptor-expressing tumor cell populations regardless of cargo

After biochemical and structural validation, we compared the *in vitro* targeting behavior of targeted synNCs and non-targeted synNCs. We mixed V4-synNCs, which encapsulate mRNA, with either wild-type K-562 cells or K-562 cells that were genetically engineered to express an individual target and a fluorescent reporter (αvβ6 integrin and mCherry, EGFR and iRFP, or Her2 and eGFP). After incubation at 37°C for 30 minutes, we washed the cells three times to prevent nonspecific RNA recovery and extracted RNA for RT-qPCR. We detected targeted synNC RNA recovered from the corresponding target cells, while we did not measure any RNA above background in samples containing off-target cells or non-targeted synNCs (**Fig. 2A**).

We next asked whether targeted synNCs could preferentially bind to target cells in mixed cell populations composed of both target-positive and target-negative cells. After determining ligand depletion prevented the measurement of K_d (*26*), we performed all flow cytometry experiments in saturating conditions (**Fig S1**). We incubated fluorescently labeled v4v0-Cys-targeted nanoparticles with wild-type K-562 cells and K-562 cells genetically modified with target receptors and fluorescent protein markers for 30 minutes at 37°C and 5% CO_2 (**Fig. 2B**). In anticipation of *in vivo* tumor targeting experiments, we also examined the targeting behavior of synNCs in pre-clinically relevant human cancer cell lines. We identified the following cell lines with which to perform flow cytometry and microscopy experiments: A-431 (αvβ6 integrin-high, EGFR-high, Her2-low), Calu-3 (αvβ6 integrin-high, EGFR-low, Her2-high), and HEK-293-T (αvβ6 integrin-low, EGFR-low, Her2-low). By flow cytometry, we did not observe any significant binding above background by non-targeted synNCs to any cell type (**Fig. 2B, F, I, L**). All three targeted synNCs showed preferential binding against target cell lines with minimal binding above background against target-negative cells (**Fig. 2C-N**). We also observed preferential binding by the corresponding monomeric binding domains (**Fig. S1**). Together with the mRNA recovery assay, these

data provide evidence that targeted synthetic nucleocapsids are able to preferentially associate with target-expressing cells, regardless of the type of encapsulated cargo.

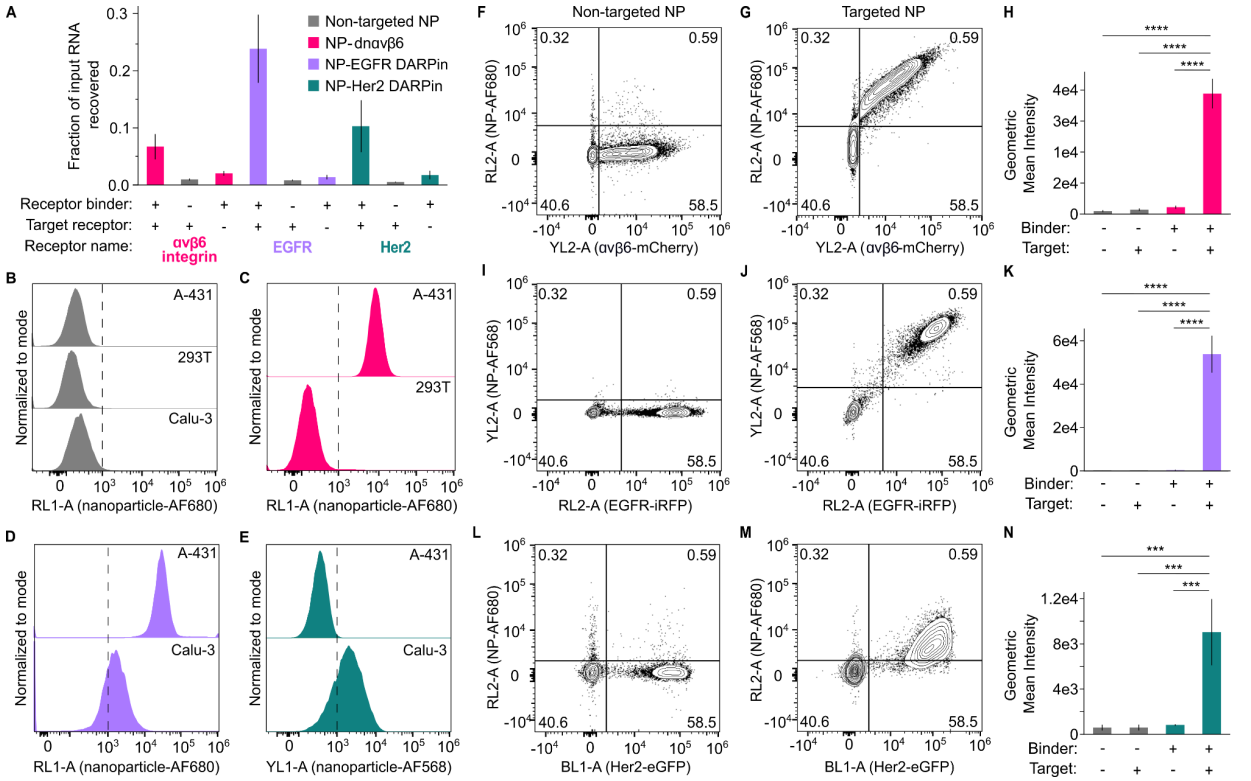


Figure 2. Ligand-receptor-dependent interaction between nanoparticles and cells observed by mRNA recovery assay and flow cytometry binding studies.
(A) Fraction of input synthetic nucleocapsid RNA recovered from K-562 cells while varying targeting moiety displayed by the synthetic nucleocapsid (dnαvβ6 binder, EGFR-binding DARPin, Her2-binding DARPin), presence or absence of of receptor on cell surface, and identity of receptor (αvβ6 integrin, EGFR, or Her2). The color key located at the top right of this panel herein denotes the sample used in each plot: gray = non-targeted nanoparticle, pink = αvβ6-targeted nanoparticle, lavender = EGFR-targeted nanoparticle, teal = Her2-targeted nanoparticle. **(B-E)** Flow cytometry of adherent cells incubated with AlexaFluor-encapsulating nanoparticles. **(F, G, I, J, L, M)** Representative flow cytometry contour plots of AlexaFluor-encapsulating nanoparticles incubated with mixed, genetically engineered K-562 cell populations with fluorescent protein reporter-linked receptor expression as denoted in (A). **(H, K, N)** Quantification of geometric mean intensities from (F, G, I, J, L, M) from 3 experimental replicates. Error bars denote standard deviations. Two-way ANOVA with Tukey correction for multiple comparisons. Only significant comparisons are shown; the remaining comparisons are not significant. ***: $P \leq 0.001$. ****: $P \leq 0.0001$.

We next sought to examine the subcellular localization of targeted and non-targeted NPs. By confocal microscopy, we observed that EGFR-targeted synthetic nucleocapsids are internalized by target cell lines when incubated at 37°C, but not at 4°C (**Fig. 3**). We did not observe significant binding by non-targeted nanoparticles.

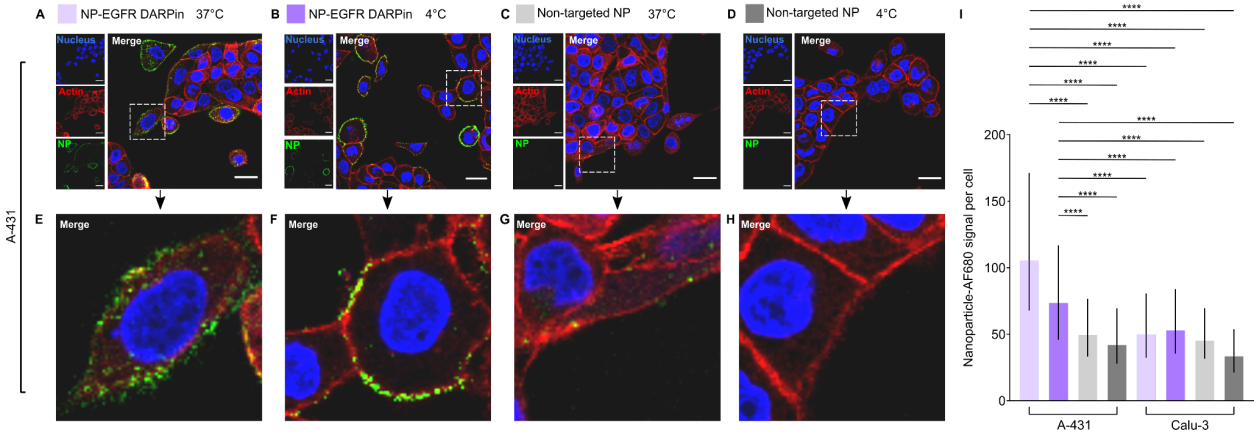


Figure 3. Targeted protein nanoparticles are internalized by receptor-expressing cancer cells.
(A-D) Confocal microscopy of EGFR-targeted nanoparticles (A-B, purple) or non-targeted nanoparticles (C-D, gray) in A-431 cells incubated for 45 minutes at 37°C (A, C, lighter shades) or 4°C (B, D, darker shades). Boxes with white, dashed lines in (A-D) denote the zoomed in regions shown in **(E-H).** Blue: NucBlue/Nucleus. Red: CellMask Actin Red/Actin. Green: AlexaFluor488/Nanoparticle (NP). **(I)** Quantification of the same microscopy samples by high-throughput fluorescence microscopy (10,000-100,000 cells per treatment from 3 experimental replicates. Bar values denote median signal. Error bars denote the interquartile range. ****: p <= 0.0001, Kruskal-Wallis test for non-Gaussian distributions with Dunn correction for multiple comparisons. For readability, only the error bars comparing the first two groups are shown) (Tab. S2). Scale bars: 10 μm.

AF680-labeled nanoparticles show target receptor-dependent accumulation in mice with human xenograft tumors

To examine the tumor targeting behavior of synNCs *in vivo*, we generated heterotopic, bilateral, human xenograft tumor models in nu/nu mice. We tested *in vivo* targeting in two separate treatment groups to maximize the difference in target receptor expression profiles (**Fig. S2 A-B, F-G**). To control for potential effects of blood flow variation on accumulation in left-side or right-side tumors, we inoculated half of the mice in each group with the target tumor on the left side and half with the target tumor on the right side, however, we did not observe any side-specific effects. Mice in group A were inoculated with A-431 cells and HEK-293-T cells on opposite flanks, as A-431 cells express higher levels of αvβ6 integrin than HEK-293-T cells. To compare the biodistribution and tumor accumulation of non-targeted and targeted synNCs, mice in group 1 were intravenously injected with I53-50-v4v0-Cys-AF680 (non-targeted NP), I53-50-v4v0-Cys-AF680-dnαvβ6 (targeted NP), or vehicle control (**Fig. S2A-B**). Mice in group 2 were

inoculated with A-431 cells and Calu-3 cells on opposite flanks, as A-431 cells exhibit anti-EGFR binding capacity and lower anti-Her2 binding capacity compared to Calu-3 cells. To compare the biodistribution and tumor accumulation of two different targeted synNCs, mice in group 2 were intravenously injected with v4v0-Cys-AF680-EGFR_DARPin, v4v0-Cys-AF680-Her2_DARPin, or vehicle control (**Fig. S2F-G**)

To qualitatively examine biodistribution over time, mice were imaged at 30 minutes, 1 hour, 6 hours, and 24 hours. Mice injected with dnαvβ6-targeted NPs were the only group that showed discernable tumor-specific accumulation out to 24 hours (**Fig. S2A-B, F-G**). After 24 hours, mice were euthanized, and blood and major organs were harvested for *ex vivo* imaging and semi-quantitative fluorescence measurement (**Fig. S2C, E, H, I**). The majority of non-targeted synNCs appeared to remain in circulation (**Fig. S2 D, E**). This contributed to high radiant efficiency in all the non-perfused major organs, which were still harboring blood. Overall, we did not observe effects based on which side the target tumor was on in any treatment group. We observed fluorescence signals above vehicle control in the tumors, liver, spleen as expected; in mice treated with Her2-targeted nanoparticles we also saw significant fluorescence signal in the heart and lungs. As mice injected with the monomeric Her2-DARPin alone on show fluorescent signal in the kidneys (as expected based on the small size and low isoelectric point of the protein), and we did not measure any mouse cross-reactivity by the Her2-DARPin by biolayer interferometry, we expect that the Her2 nanoparticles likely aggregated in circulation after administration (**Fig. S3**). Still, in all mice except EGFR-A-4, we observed preferential accumulation of synNCs in the target tumors. These preliminary data indicate that displaying specific tumor antigen binding domains on the surface of synthetic nucleocapsids might be an effective strategy to bias accumulation in target tumors *in vivo*.

***In vitro*-assembled nanoparticles retain stability and cell target selectivity upon conjugation to small molecules.**

After qualitatively observing the most selective targeting by dnαvβ6-targeted NPs *in vivo,* for efficiency and experimental feasibility, we moved forward with just this construct and the non-targeted construct. Our first aim was to develop a process for *in vitro*-assembled nanoparticle production with the cys-less pentamer and the single-cysteine trimer component. We experienced significant instability and low soluble yield of the pentamer component upon purification of that individual component, so we performed homology-guided redesign to introduce four core mutations that appeared to more efficiently pack the hydrophobic core of the pentamer component (**Fig. 4A-B**). We did not show whether these core mutations actually improved stability or resulted in some other phenomenon that happened to improve solubility (e.g., lower expression levels leading to increased solubility via decreased aggregation). Nonetheless, the cys-less pentamer with four core mutations yielded complete assemblies in the desired architectures, so we moved forward with this component (**Fig. 4C-I**). We separately produced trimers and pentamers, assembled the nanoparticles, (**Fig. 4C-D**), conjugated Alexa Fluor 680 (AF680) to the nanoparticle interiors, and examined their stability upon freeze-thaw. Both nanoparticles exhibited monodisperse populations by HPLC (**Fig. 4E-F**) and negative stain electron microscopy (nsEM) (**Fig. 4G-I**). Additionally, the nanoparticles retained selective targeting of only target-positive cells (A-431) by dnαvβ6-targeted NPs (**Fig. 4H, J**). Although we had not yet scaled up the manufacturing of the material to the order of magnitude needed for *in vivo* studies, these preliminary results demonstrate the successful production of low-endotoxin, *in vitro*-assembled protein nanoparticles capable of covalently encapsulating small molecules (AF680 in this case).

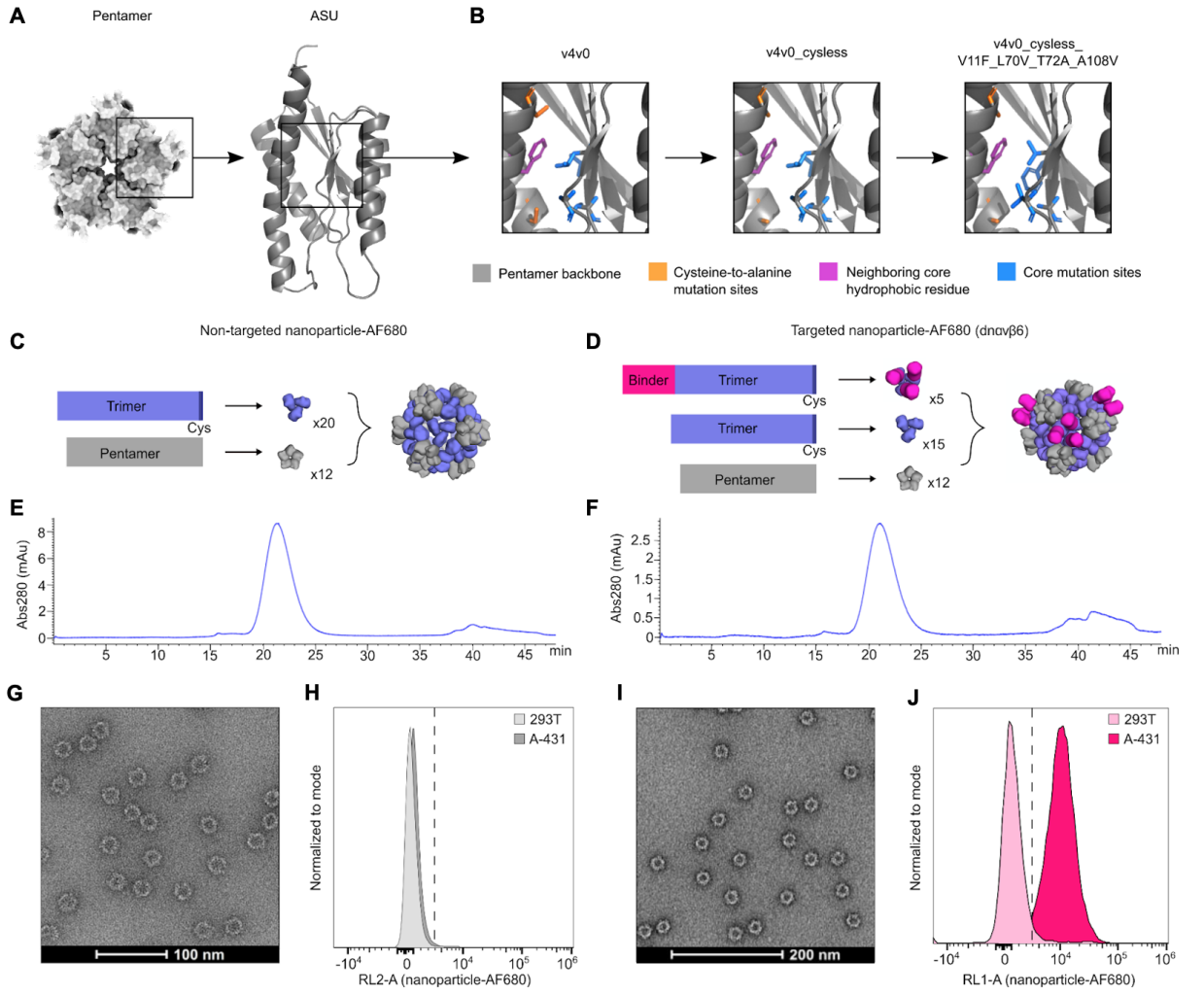


Figure 4. Following homology-guided redesign of the cys-less pentamer component, in vitro-assembled, AF680-labeled nanoparticles retain size, monodispersity, morphology, and target selectivity.
(A-B) The core of the pentamer component (A) was redesigned to be free of cysteines and to increase hydrophobic packing (B). The cysteine-to-alanine and core mutations were mimicked from Brucella abortus lumazine synthase (PDB ID 1T13), a homolog of Mesorhizobium loti lumazine synthase (PDB ID 2OBX), the structure from which this pentamer was designed for nanoparticle assembly. **(C-D)** *In vitro* assembly scheme. I53-50-based nanoparticles are composed of 20 trimer subunits and 12 pentamer subunits. Trimers can either be bare (A) or fused to a targeting domain (B). Trimers contain an internally-facing cysteine for maleimide conjugation to small molecules such as AlexaFluor 680 (AF680). **(E-F)** High-performance liquid chromatography (HPLC) of non-targeted and targeted nanoparticles, respectively, each show monodisperse peaks at the expected retention time. **(G, I)** Negative stain electron microscopy (nsEM) of non-targeted and targeted nanoparticles, respectively, shows monodisperse nanoparticles of the expected sizes and morphologies. **(H, J)** Flow cytometry of AF680-labeled non-targeted nanoparticles (H, grays) and targeted nanoparticles (J, pinks) shows selective binding of targeted nanoparticles with target cells (A-431, dark pink) and lack of binding in all other conditions (i.e., no binding of any nanoparticle to 293T cells (light shades) and no binding of non-targeted nanoparticle (grays) to either cell line).

We next sought to examine the *in vitro* cytotoxicity of non-targeted and targeted nanoparticles loaded with MMAD instead of AF680. We performed these assays with two different nanoparticle lots that had different MMAD conjugation and purification efficiencies (**Tab. S3**). We observed different EC50 of MMAD depending on the quantity of free MMAD in solution (batch A in **Fig. 5A** versus batch B in **Fig. 5B**). We also observed that A-431 cells are more sensitive to free MMAD-linker (MMAD-Val-Cit-PAB-Maleimide) than HEK293T cells (**Fig. 5B-C**). However, we also observed that for both batches A and B, the rank-order EC50 of targeted nanoparticles with targeted cells is lower than any off-target combination (targeted nanoparticles with off-target cells, non-targeted nanoparticles with either cell line) (**Tab. S4**). These data suggest that targeted nanoparticles are more potent at killing the corresponding target cell population than non-targeted nanoparticles.

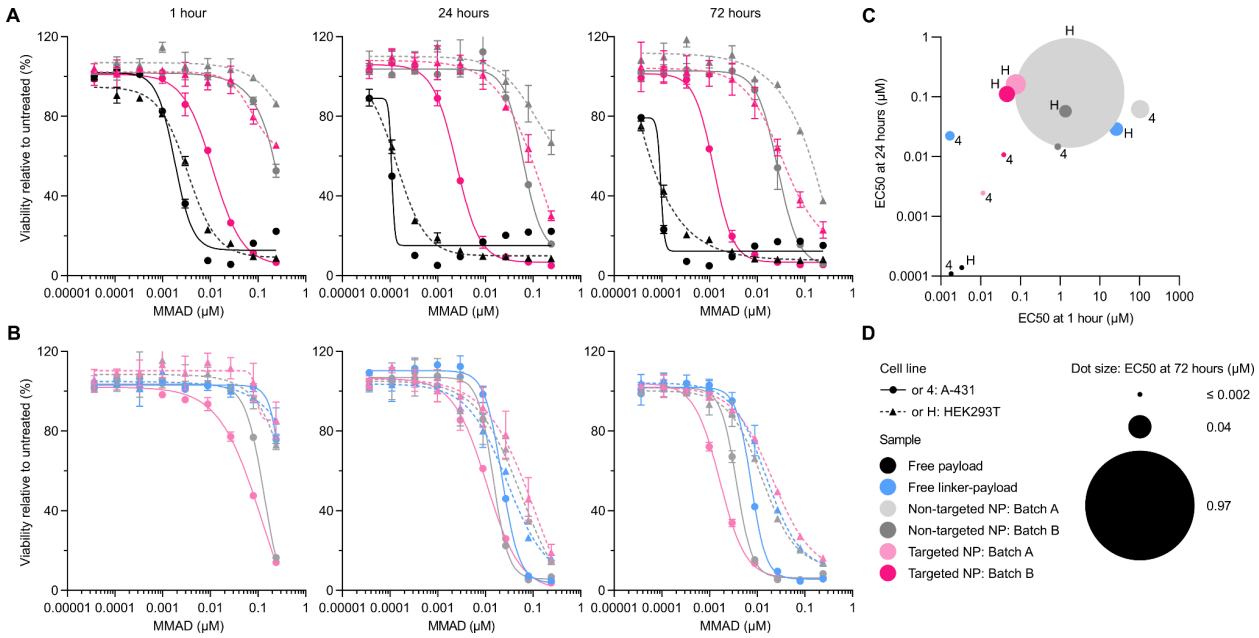


Figure 5. MMAD cell killing assays in HEK293T cells and A-431 cells.
(**A-B**) A-341 cells (solid lines) or HEK293T cells (dashed lines) were treated with payload (free payload: black, free linker-payload: blue) or nanoparticle-MMAD conjugates (non-targeted: gray, targeted: pink). Cell viability relative to untreated cells was measured in an Incucyte after being treated with nanoparticle-MMAE conjugates from Batch A (A) or Batch B (B) after 1 hour (left), 24 hours (middle), or 72 hours (right). (**C**) Multivariable graph of EC50 of cells treated with free payload (MMAD, black), free linker-payload (MMAD-VC-PAB-Mal, blue), non-targeted nanoparticle-MMAD conjugate (gray), or targeted nanoparticle-MMAD conjugate (pink). Dot size: scaled according to the EC50 at 72 hours. 4: A-431 cells. H: HEK293T cells. (**D**) Legend. Error bars: standard deviation from three technical replicates.

With the ultimate goal of *in vivo* tumor reduction studies, we first performed a pilot toxicity study to measure the maximum tolerated dose (MTD) of protein-MMAE conjugates in healthy, athymic mice (*27*). We first examined the tolerability of dnαvβ6_MMAE, non-targeted NP_MMAE, and dnαvβ6-targeted NP_MMAE at and below the known MTD of free MMAE in mice (0.5 mg/kg) (*28*, *29*). Since MMAE is rapidly eliminated from the systemic circulation after IV administration in tumor-bearing mice but remains in tissues for 6-7 days before complete clearance, we measured mouse health over the course of 8 to 9 days (*29*). We did not observe any signs of toxicity and on average, the mice continued to increase in weight over time (**Fig. 6**, **Fig. S4**).

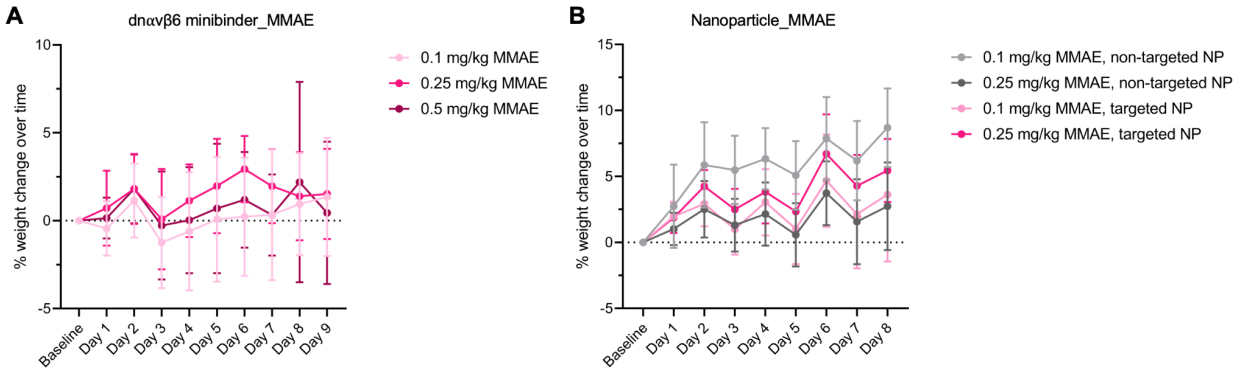


Figure 6. Nanoparticle-MMAE and minibinder-MMAE are tolerated in nu/nu mice at the known maximum tolerated dose (MTD) of free MMAE.
Percent weight change over time of nude mice treated with **(A)** dnαvβ6_MMAE (minibinder conjugated to MMAE) dosed at 0.1 mg/kg MMAE, 0.25 mg/kg MMAE, or 0.5 mg/kg MMAE, or **(B)** non-targeted NP_MMAE or dnαvβ6-targeted NP_MMAE at 0.1 mg/kg MMAE or 0.25 mg/kg MMAE.

Discussion

Here, we show that the cargo of I53-50 can be modularly reassigned by mutation of interior surface residues to allow either RNA packaging (v4) or chemical conjugation to small molecules of interest for imaging or anti-tumor activity (v4v0-Cys). We also show that nanoparticles can be loaded with MMAD to achieve cell killing *in vitro*. Furthermore, we observed no toxicity of dnαvβ6_MMAE, non-targeted

NP_MMAE, and dnαvβ6-targeted NP_MMAE at the known MTD of MMAE in mice (0.25-0.5 mg/kg). Future studies will examine higher doses of MMAE-conjugated proteins, such as doses used by antibody-drug conjugate and lipid nanoparticles (*28*, *29*), and will measure the therapeutic effect of MMAE-conjugated proteins in mouse tumor models. Overall, the modularity of both targeting domain and payload make synthetic nucleocapsids a promising platform technology for a variety of targeted delivery problems.

Brief Materials and Methods

Protein production, biochemical characterization, and RNA recovery

Synthetic nucleocapsids were grown as previously described in Butterfield & Lajoie et al. (Nature, 2017) and Olshefsky et al. (Chapter 2 here) unless otherwise noted (*13*). Synthetic nucleocapsids displaying the alpha v beta 6 integrin binder were grown by autoinduction. Final protein assemblies were purified into 25 mM HEPES, 150 mM NaCl, pH 7.4.

Flow cytometry

Cells were grown until mid-growth phase in media containing 10% v/v FBS. A-431, Calu-3, HEK-293T, and K-562 cells were grown in DMEM, EMEM, DMEM, and RPMI, respectively. Adherent cells were lifted from growth flasks with Stem Pro Accutase. Staining for flow cytometry was performed in the following binding buffer: 20 mM Tris, 150 mM NaCl, 1 mM MgCl2, 1 mM CaCl2, pH 7.4. Cells were washed twice with binding buffer, stained with Zombie Violet viability stain according to manufacturer's instructions, washed twice with binding buffer 1% BSA, and incubated with 100 nM protein in binding buffer 1% BSA for 30-60 minutes at 4 degrees C. K-562 cells were not washed before flow cytometry. Adherent cells were washed twice before flow cytometry. Samples were analyzed on an Attune NxT Flow Cytometer (Invitrogen).

Microscopy

Cells cultured as described above. Cells were seeded in Corning 10-well plates or ibidi µ-slide 8 well plates and grown for 1-2 days. Cells were washed with fresh, complete media, then incubated with 10 nM protein ASU in complete media for 45 minutes at either 37 or 4 degrees C. Cells were washed twice, fixed in 4% paraformaldehyde, stained with NucBlue Live Cell Ready Probe and Cell Mask Actin Stain, and washed three more times. Samples were imaged on a Leica SP8X Confocal Microscope and a GE IN Cell Analyzer 2200.

In vivo targeting

Tumors were injected subcutaneously on the flanks of nu/nu mice by the Takeda team. 1 nmole of fluorescently labeled nanoparticles that were frozen in 2.5% glycerol and subsequently thawed on ice were injected intravenously through the tail vein. Live mouse images were acquired on an IVIS system over time. Mice were euthanized and organs and blood were extracted for ex vivo imaging. All experiments were performed according to approved protocols by the office of animal welfare.

Alexa Fluor conjugation

Small molecules were added between 2- and 10-fold molar ratio over protein, incubated for 1-2 hours at room temperature or overnight at 4 degrees C, and purified with dialysis and a desalting column followed by SEC.

Acknowledgments

We are grateful for funding from NIH (National Cancer Institute NCI) 1R21CA232430-01, T32CA080416, and the Audacious Project grant for Drug Delivery Vehicles. Thanks to Ratika Krishnamurty and __(Takeda)__ for managing the collaboration between UW and Takeda. Thanks to Dr. Karla Herpoldt, Brooke Fiala, Dr. Drew Sellers, Isaac Sappington, Adam Wargacki, Erin Yang, Nataly Kacherovsky, Dr. Ian Cardle, Dr. Gary Liu, Mike Murphy, Maggie Alrichs, and Kandise Van Wormer for technical and scientific advice, feedback on the manuscript, lab support, and insightful conversations.

References

1. M. P. Kumar, J. Du, G. Lagoudas, Y. Jiao, A. Sawyer, D. C. Drummond, D. A. Lauffenburger, A. Raue, Analysis of Single-Cell RNA-Seq Identifies Cell-Cell Communication Associated with Tumor Characteristics. *Cell Rep.* **25**, 1458–1468.e4 (2018).

2. M. J. Mitchell, M. M. Billingsley, R. M. Haley, M. E. Wechsler, N. A. Peppas, R. Langer, Engineering precision nanoparticles for drug delivery. *Nat. Rev. Drug Discov.* **20**, 101–124 (2021).

3. M. E. Davis, Z. G. Chen, D. M. Shin, Nanoparticle therapeutics: an emerging treatment modality for cancer. *Nat. Rev. Drug Discov.* **7**, 771–782 (2008).

4. X. Sun, Y. Hong, Y. Gong, S. Zheng, D. Xie, Bioengineered Ferritin Nanocarriers for Cancer Therapy. *Int. J. Mol. Sci.* **22** (2021), doi:10.3390/ijms22137023.

5. D. H. Bryant, A. Bashir, S. Sinai, N. K. Jain, P. J. Ogden, P. F. Riley, G. M. Church, L. J. Colwell, E. D. Kelsic, Deep diversification of an AAV capsid protein by machine learning. *Nat. Biotechnol.* **39**, 691–696 (2021).

6. Z. Cheng, A. Al Zaki, J. Z. Hui, V. R. Muzykantov, A. Tsourkas, Multifunctional nanoparticles: cost versus benefit of adding targeting and imaging capabilities. *Science*. **338**, 903–910 (2012).

7. R. C. Münch, H. Janicki, I. Völker, A. Rasbach, M. Hallek, H. Büning, C. J. Buchholz, Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer. *Mol. Ther.* **21**, 109–118 (2013).

8. B. E. Deverman, P. L. Pravdo, B. P. Simpson, S. R. Kumar, K. Y. Chan, A. Banerjee, W.-L. Wu, B. Yang, N. Huber, S. P. Pasca, V. Gradinaru, Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat. Biotechnol.* **34**, 204–209 (2016).

9. M. P. Lokugamage, C. D. Sago, Z. Gan, B. R. Krupczak, J. E. Dahlman, Constrained Nanoparticles Deliver siRNA and sgRNA to T Cells In Vivo without Targeting Ligands. *Adv. Mater.* **31**, e1902251 (2019).

10. X. Huang, J. Z. Williams, R. Chang, Z. Li, C. E. Burnett, R. Hernandez-Lopez, I. Setiady, E. Gai, D. M. Patterson, W. Yu, K. T. Roybal, W. A. Lim, T. A. Desai, DNA scaffolds enable efficient and tunable functionalization of biomaterials for immune cell modulation. *Nat. Nanotechnol.* **16**, 214–223 (2021).

11. A. Olshefsky, C. Richardson, S. H. Pun, N. P. King, Engineering Self-Assembling Protein Nanoparticles for Therapeutic Delivery. *Bioconjug. Chem.* **33**, 2018–2034 (2022).

12. Y. Wang, T. Douglas, Bioinspired Approaches to Self-Assembly of Virus-like Particles: From Molecules to Materials. *Acc. Chem. Res.* **55**, 1349–1359 (2022).

13. G. L. Butterfield, M. J. Lajoie, H. H. Gustafson, D. L. Sellers, U. Nattermann, D. Ellis, J. B. Bale, S. Ke, G. H. Lenz, A. Yehdego, R. Ravichandran, S. H. Pun, N. P. King, D. Baker, Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature*. **552**, 415–420 (2017).

14. A. Roy, L. Shi, A. Chang, X. Dong, J. Li, R. V. Winegar, G. M. Cherf, D. Slocum, P. D. Poulson, M. Miranda, H. Bai, Y. Kipnis, T. Priya, L. Carter, R. Ravichandran, C. M. Chow, C. Overed-Sayer, D.

K. Finch, D. Lowe, L. J. Stewart, A. K. Bera, G. Matute-Bello, T. P. Birkland, F. DiMaio, G. Raghu, J. R. Cochran, P. M. Van Ry, T. Springer, D. Baker, *De novo* Design of a Highly Specific Inhibitor of Integrin Avβ6 for Pulmonary Fibrosis (2021), , doi:10.2139/ssrn.3802854.

15. D. Steiner, P. Forrer, A. Plückthun, Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. *J. Mol. Biol.* **382**, 1211–1227 (2008).

16. C. Zahnd, E. Wyler, J. M. Schwenk, D. Steiner, M. C. Lawrence, N. M. McKern, F. Pecorari, C. W. Ward, T. O. Joos, A. Plückthun, A designed ankyrin repeat protein evolved to picomolar affinity to Her2. *J. Mol. Biol.* **369**, 1015–1028 (2007).

17. A. Maderna, M. Doroski, C. Subramanyam, A. Porte, C. A. Leverett, B. C. Vetelino, Z. Chen, H. Risley, K. Parris, J. Pandit, A. H. Varghese, S. Shanker, C. Song, S. C. K. Sukuru, K. A. Farley, M. M. Wagenaar, M. J. Shapiro, S. Musto, M.-H. Lam, F. Loganzo, C. J. O'Donnell, Discovery of cytotoxic dolastatin 10 analogues with N-terminal modifications. *J. Med. Chem.* **57**, 10527–10543 (2014).

18. A. Maderna, C. A. Leverett, Recent advances in the development of new auristatins: structural modifications and application in antibody drug conjugates. *Mol. Pharm.* **12**, 1798–1812 (2015).

19. V. Chudasama, A. Maruani, S. Caddick, Recent advances in the construction of antibody–drug conjugates. *Nat. Chem.* **8**, 114–119 (2016).

20. A. C. Conibear, A. Schmid, M. Kamalov, C. F. W. Becker, C. Bello, Recent Advances in Peptide-Based Approaches for Cancer Treatment. *Curr. Med. Chem.* **27**, 1174–1205 (2020).

21. H. Jun, E. Jang, H. Kim, M. Yeo, S. G. Park, J. Lee, K. J. Shin, Y. C. Chae, S. Kang, E. Kim, TRAIL & EGFR affibody dual-display on a protein nanoparticle synergistically suppresses tumor growth. *J. Control. Release*. **349**, 367–378 (2022).

22. M. Liang, K. Fan, M. Zhou, D. Duan, J. Zheng, D. Yang, J. Feng, X. Yan, H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14900–14905 (2014).

23. SK bioscience and GSK's adjuvanted COVID-19 vaccine candidate meets coprimary objectives in a phase III study; Biologics License Application submitted for SKYCovione™(GBP510/GSK adjuvant) in South Korea (2022), (available at https://www.gsk.com/en-gb/media/press-releases/sk-bioscience-and-gsk-s-adjuvanted-covid-19-vaccine-candidate-meets-coprimary-objectives-in-a-phase-iii-study/).

24. P. V. Baranov, R. F. Gesteland, J. F. Atkins, Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.* **3**, 373–377 (2002).

25. C. L. Sanders, J. F. Curran, Genetic analysis of the E site during RF2 programmed frameshifting. *RNA*. **13**, 1483–1491 (2007).

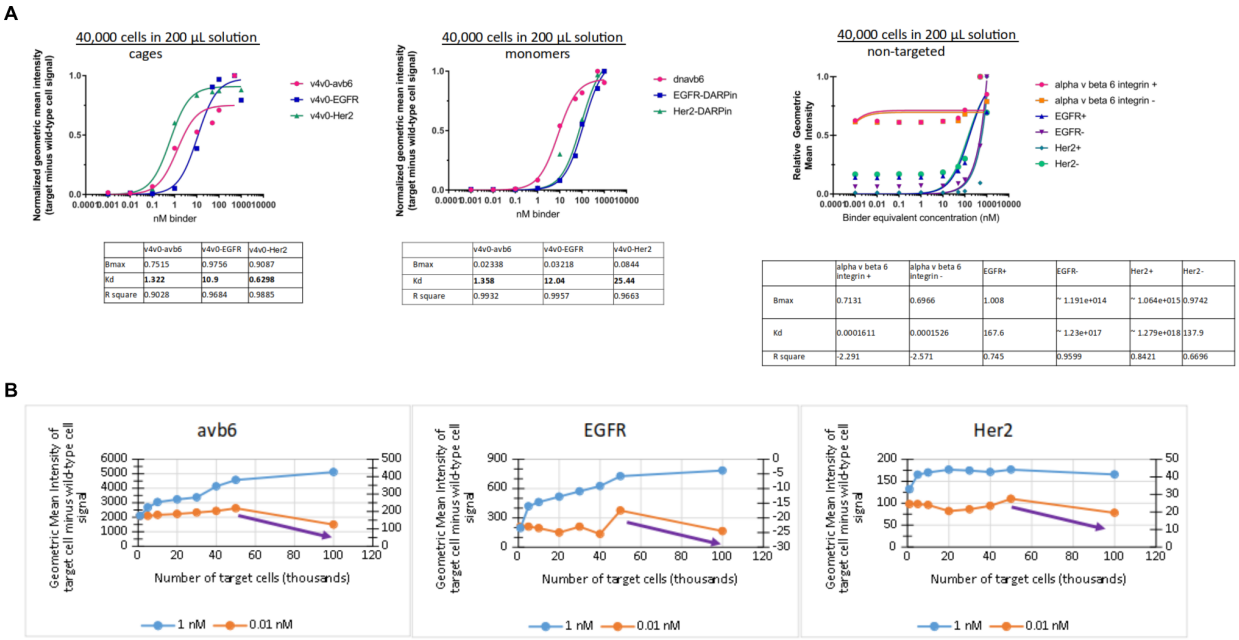
26. S. A. Hunter, J. R. Cochran, Cell-Binding Assays for Determining the Affinity of Protein-Protein Interactions: Technologies and Considerations. *Methods Enzymol.* **580**, 21–44 (2016).

27. Q. Zhang, S. X. Zeng, H. Lu, Determination of Maximum Tolerated Dose and Toxicity of Inauhzin in Mice. *Toxicol Rep*. **2**, 546–554 (2015).

28. R. Qi, Y. Wang, P. M. Bruno, H. Xiao, Y. Yu, T. Li, S. Lauffer, W. Wei, Q. Chen, X. Kang, H. Song, X. Yang, X. Huang, A. Detappe, U. Matulonis, D. Pepin, M. T. Hemann, M. J. Birrer, P. P. Ghoroghchian, Nanoparticle conjugates of a highly potent toxin enhance safety and circumvent platinum resistance in ovarian cancer. *Nat. Commun.* **8**, 2166 (2017).

29. H. P. Chang, Y. K. Cheung, D. K. Shah, Whole-Body Pharmacokinetics and Physiologically Based Pharmacokinetic Model for Monomethyl Auristatin E (MMAE). *J. Clin. Med. Res.* **10** (2021), doi:10.3390/jcm10061332.

Supporting information



	v4v0-avb6	v4v0-EGFR	v4v0-Her2
Bmax	0.7515	0.9756	0.9087
Kd	**1.322**	**10.9**	**0.6298**
R square	0.9028	0.9684	0.9885

	v4v0-avb6	v4v0-EGFR	v4v0-Her2
Bmax	0.02338	0.03218	0.0844
Kd	**1.358**	**12.04**	**25.44**
R square	0.9932	0.9957	0.9663

	alpha v beta 6 integrin +	alpha v beta 6 integrin -	EGFR+	EGFR-	Her2+	Her2-
Bmax	0.7131	0.6966	1.008	~ 1.191e+014	~ 1.064e+015	0.9742
Kd	0.0001611	0.0001526	167.6	~ 1.23e+017	~ 1.279e+018	137.9
R square	-2.291	-2.571	0.745	0.9599	0.8421	0.6696

Figure S1. Validation of cell target binding by flow cytometry.

Titrations of concentration of binder **(A)** or number of cells per well **(B)**. K-562 wild-type cells were mixed with equivalent numbers of K-562-αvβ6_mCherry, K-562-EGFR_iRFP, or K-562-Her2-eGFP cells. Mixed cell populations were incubated with v4v0-dnαvβ6, v4v0-EGFR_DARPin, v4v0-Her2_DARPin, dnαvβ6 monomeric binder, EGFR monomeric DARPin, Her2 monomeric DARPin, or non-targeted nanoparticle. Cells in (B) expressed the target receptor as indicated at the top of each plot (αvβ6 integrin, EGFR, or Her2) and were incubated with nanoparticles displaying the corresponding targeting domain.

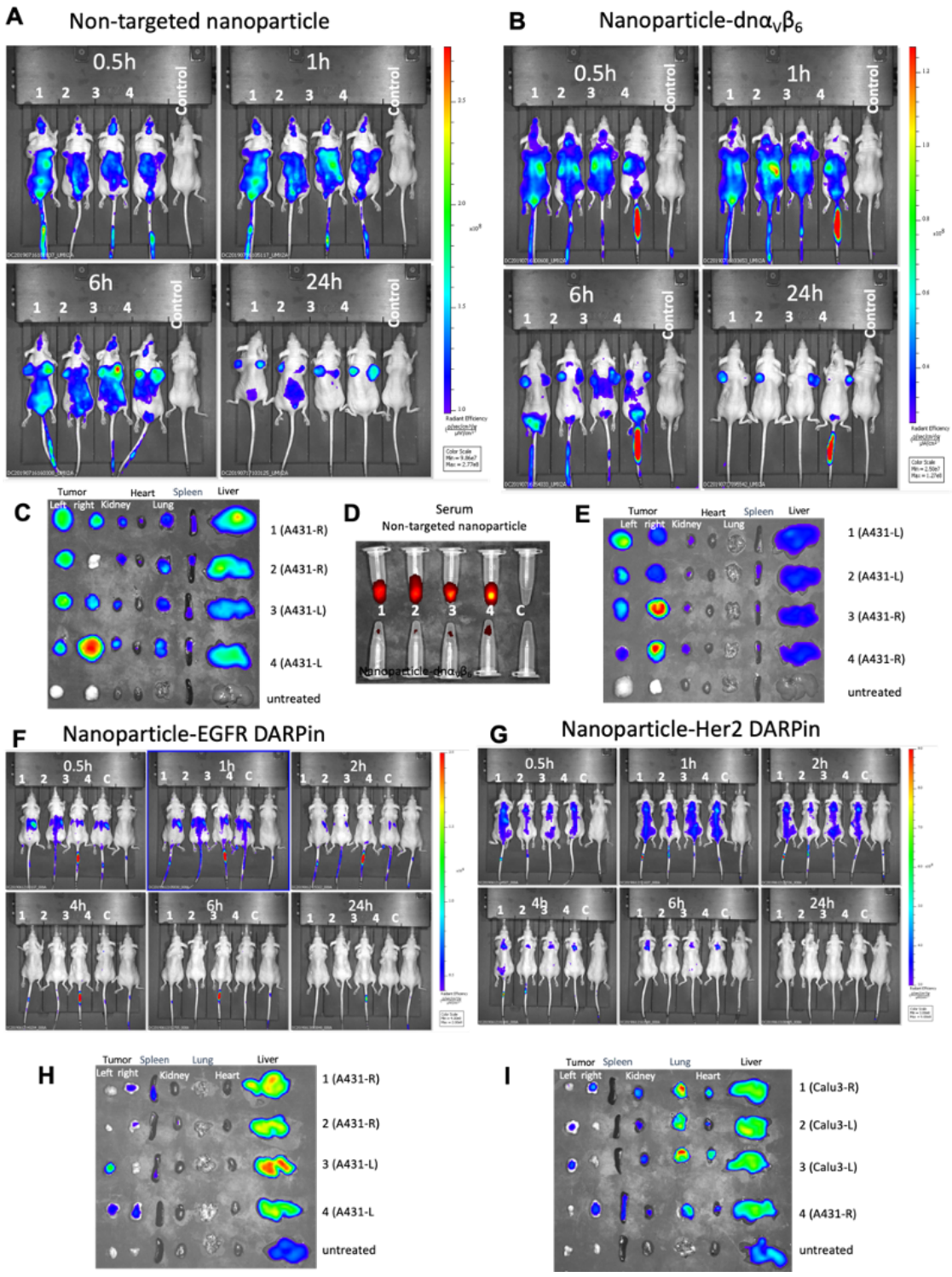


Figure S2. Biodistribution of targeted and non-targeted nanoparticles in bilateral in vivo tumor models.
Mice with bilateral, heterotopic tumors were injected with 1 nmol Alexa Fluor 680-labeled nanoparticles and imaged over time. At 24 hours, mice were euthanized, and organs were collected for imaging. Tumor identities are indicated in the organ subfigures. **(A, C)** Non-targeted nanoparticle and **(B, E)** avb6 integrin-targeted nanoparticle in mice with A-431 and 293T tumors. **(D)** Serum from mice treated with non-targeted nanoparticle (top) or targeted nanoparticle. **(F-H)** EGFR-targeted nanoparticle and **(G-I)** Her2-targeted nanoparticle in mice with A-431 and Calu-3 tumors, as indicated in H and I.

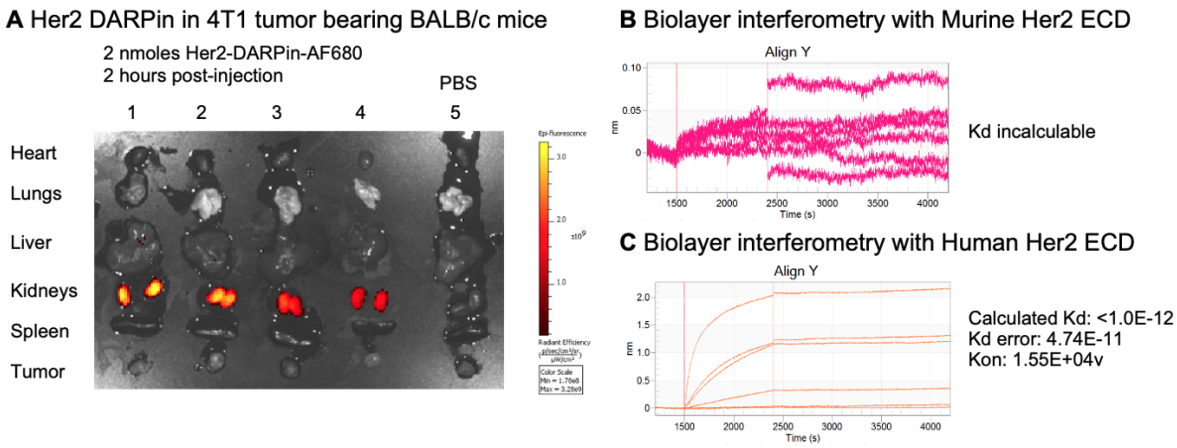


Figure S3. Testing species cross-reactivity of the Her2 DARPin and I53-50 nanoparticles displaying the Her2 DARPin.

(A) 2 nmoles of AF680-labeled DARPin were injected into BALB/c mice bearing heterotopic 4T1 mammary tumors and imaged with IVIS. **(B-C)** Biolayer interferometry of I53-50 nanoparticles displaying the Her2 DARPin with murine **(B)** or human **(C)** Her2 ectodomain (ECD).

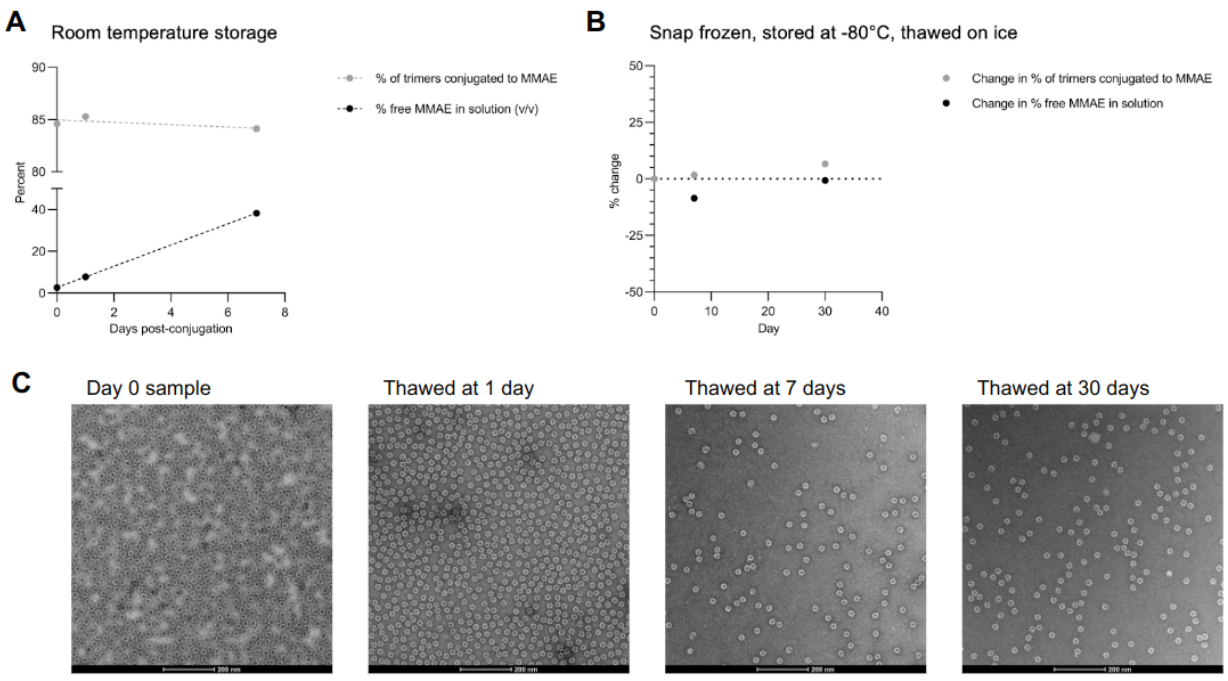


Figure S4. Nanoparticle-MMAE conjugation and biochemical characterization.
(A) Nanoparticle-MMAE conjugation and biochemical characterization shows high conjugation efficiency by LC-MS. **(A-B)** MMAE is released over time from nanoparticle-Cys:MC-Val-Cit-PAB-MMAE conjugate at room temperature, but not upon storage at -80°C. **(C)** Nanoparticles retain stability and structure after storage at -80 degrees C for at least 30 days, as measured by LC-MS (B) and negative stain electron microscopy (C).

Table S1. Endotoxin values.

Sample (dnαvβ6-targeted unless noted)	**Corresponding figure(s)**	**EU per dose**
Non-targeted NP_MMAD	Fig. 5A	78.95
Targeted NP_MMAD	Fig. 5A	111.56
Non-targeted NP_MMAD	Fig. 5B	107.05
Targeted NP_MMAD	Fig. 5B	114.05
Minibinder, 0.1 mg/kg MMAE	Fig. 6A	2.0
Minibinder, 0.25 mg/kg MMAE	Fig. 6A	5.0
Minibinder, 0.5 mg/kg MMAE	Fig. 6A	10.0
Non-targeted NP, 0.1 mg/kg MMAE	Fig. 6B	7.0
Non-targeted NP, 0.25 mg/kg MMAE	Fig. 6B	45.3
Targeted NP, 0.1 mg/kg MMAE	Fig. 6B	19.6
Targeted NP, 0.25 mg/kg MMAE	Fig. 6B	66.3
All other samples	Exact values were not recorded. <1000 EU per mg was required at the time.	

Table S2. Statistics from Figure 3: Kruskal-Wallis test with Dunn's correction for multiple comparisons.

Number of families	1			
Number of comparisons per family	28			
Alpha	0.05			
Dunn's multiple comparisons test	**Mean rank diff.**	**Significant?**	**Summary**	**Adjusted P Value**
a431_targeted_37 vs. a431_targeted_4	6922	Yes	****	<0.0001
a431_targeted_37 vs. a431_nontargeted_37	14400	Yes	****	<0.0001
a431_targeted_37 vs. a431_nontargeted_4	16881	Yes	****	<0.0001
a431_targeted_37 vs. calu3_targeted_37	13937	Yes	****	<0.0001
a431_targeted_37 vs. calu3_targeted_4	12804	Yes	****	<0.0001
a431_targeted_37 vs. calu3_nontargeted_37	15997	Yes	****	<0.0001
a431_targeted_37 vs. calu3_nontargeted_4	20913	Yes	****	<0.0001
a431_targeted_4 vs. a431_nontargeted_37	7478	Yes	****	<0.0001
a431_targeted_4 vs. a431_nontargeted_4	9959	Yes	****	<0.0001
a431_targeted_4 vs. calu3_targeted_37	7015	Yes	****	<0.0001
a431_targeted_4 vs. calu3_targeted_4	5882	Yes	****	<0.0001
a431_targeted_4 vs. calu3_nontargeted_37	9076	Yes	****	<0.0001
a431_targeted_4 vs. calu3_nontargeted_4	13991	Yes	****	<0.0001
a431_nontargeted_37 vs. a431_nontargeted_4	2482	Yes	****	<0.0001
a431_nontargeted_37 vs. calu3_targeted_37	-462.3	No	ns	>0.9999
a431_nontargeted_37 vs. calu3_targeted_4	-1596	Yes	***	0.0004
a431_nontargeted_37 vs. calu3_nontargeted_37	1598	Yes	***	0.0008
a431_nontargeted_37 vs. calu3_nontargeted_4	6513	Yes	****	<0.0001
a431_nontargeted_4 vs. calu3_targeted_37	-2944	Yes	****	<0.0001
a431_nontargeted_4 vs. calu3_targeted_4	-4078	Yes	****	<0.0001

a431_nontargeted_4 vs. calu3_nontargeted_37	-883.7	No	ns	0.581
a431_nontargeted_4 vs. calu3_nontargeted_4	4032	Yes	****	<0.0001
calu3_targeted_37 vs. calu3_targeted_4	-1134	No	ns	0.2277
calu3_targeted_37 vs. calu3_nontargeted_37	2060	Yes	****	<0.0001
calu3_targeted_37 vs. calu3_nontargeted_4	6976	Yes	****	<0.0001
calu3_targeted_4 vs. calu3_nontargeted_37	3194	Yes	****	<0.0001
calu3_targeted_4 vs. calu3_nontargeted_4	8109	Yes	****	<0.0001
calu3_nontargeted_37 vs. calu3_nontargeted_4	4916	Yes	****	<0.0001

Table S3. MMAD-Nanoparticle conjugation efficiencies and related values.

Batch	Sample	Concentration (μM)	Concentration (mg/mL)	Conjugation efficiency & "DAR" (MMAD per NP)	Concentration Conjugated MMAD	% Free MMAD
A	Non-targeted_ MMAD	3.015	0.125	100% (60)	3.02	0.27
	Targeted_ MMAD	6.178	0.289	100% (60)	6.2	20.15
B	Non-targeted_ MMAD	12.060	0.498	100% (60)	12.06	2.44
	Targeted_ MMAD	7.061	0.331	100% (60)	7.06	0.72

Table S4. Rank order EC50 at 72 hours from Figure 5.

	Batch A			Batch B		
Rank order (lowest to highest)	**Cell line**	**Sample**	**EC50 at 72 hours (µM)**	**Cell line**	**Sample**	**EC50 at 72 hours (µM)**
1	HEK293T	Free payload	1.63E-06	HEK293T	Free payload	1.63E-06
2	A-431	Free payload	9.50E-05	A-431	Free payload	9.50E-05
3	A-431	Targeted: Batch A	1.22E-03	A-431	Targeted: Batch B	1.65E-03
4	A-431	Free linker-payload	7.45E-03	A-431	Non-targeted: Batch B	3.69E-03
5	HEK293T	Free linker-payload	1.53E-02	A-431	Free linker-payload	7.45E-03
6	A-431	Non-targeted: Batch A	2.70E-02	HEK293T	Non-targeted: Batch B	1.28E-02
7	HEK293T	Targeted: Batch A	3.29E-02	HEK293T	Free linker-payload	1.53E-02
8	HEK293T	Non-targeted: Batch A	9.71E-01	HEK293T	Targeted: Batch B	2.11E-02

Table S5. Values from Figure 6 toxicity study.

Protein sample	mg/kg MMAE per dose	mg/kg protein per dose	% free MMAE by volume	Protein-MMAE conjugation efficiency (%)
Minibinder	0.1	0.2	88.6	16.2
	0.25	0.6	88.6	16.2
	0.5	1.1	88.6	16.2
Non-targeted NP	0.1	0.4	2.0	85.0
	0.25	2.6	42.0	79.0
Targeted NP	0.1	2.3	1.3	65.0
	0.25	7.8	38.0	75.0

Chapter 4. Proposed future directions

Audrey Ellen Olshefsky, Kefan Song, Xinru Wang, Christian Richardson, Neil P. King and Suzie H. Pun

There are many avenues and opportunities stemming from the work reported here that we have not yet fully resolved, explored, or completed. I will focus on the three subjects: (1) the insulin minibinders and other sequences of interest we discovered in the *in vivo* library selection project, (2) ideas for extending this work to *in vivo* selection for delivery readouts (beyond biodistribution), and (3) the manufacturing difficulties we have experienced with the MMAE delivery project.

Identification of insulin receptor minibinders and brain-targeting sequences by in vivo library selection based on synthetic nucleocapsids

In Chapter 2, we developed a novel ligand display platform to identify miniproteins that shift nanoparticle biodistribution. After empirically determining the cutoff for successful organ targeting to be $\log_{10}$(enrichment) of 1.0 (Chapter 2, Figure S11), we re-examined our sequencing data. Notably, a common motif enriched in the muscles and brain emerged, which turned out to be an insulin receptor minibinder designed by Derrick Hicks (**Fig. 1A-B**). Surface plasmon resonance (SPR) showed that these minibinders, when synthesized (separately from the nanoparticle), do indeed bind insulin receptor ectodomain with micromolar affinity (**Fig. 1C**). These preliminary results suggest that our library platform can be used in a similarly modular way as phage display, where minibinders identified with this method can be utilized on other delivery vehicles to confer *in vivo* targeting by the receptor-ligand specificity.

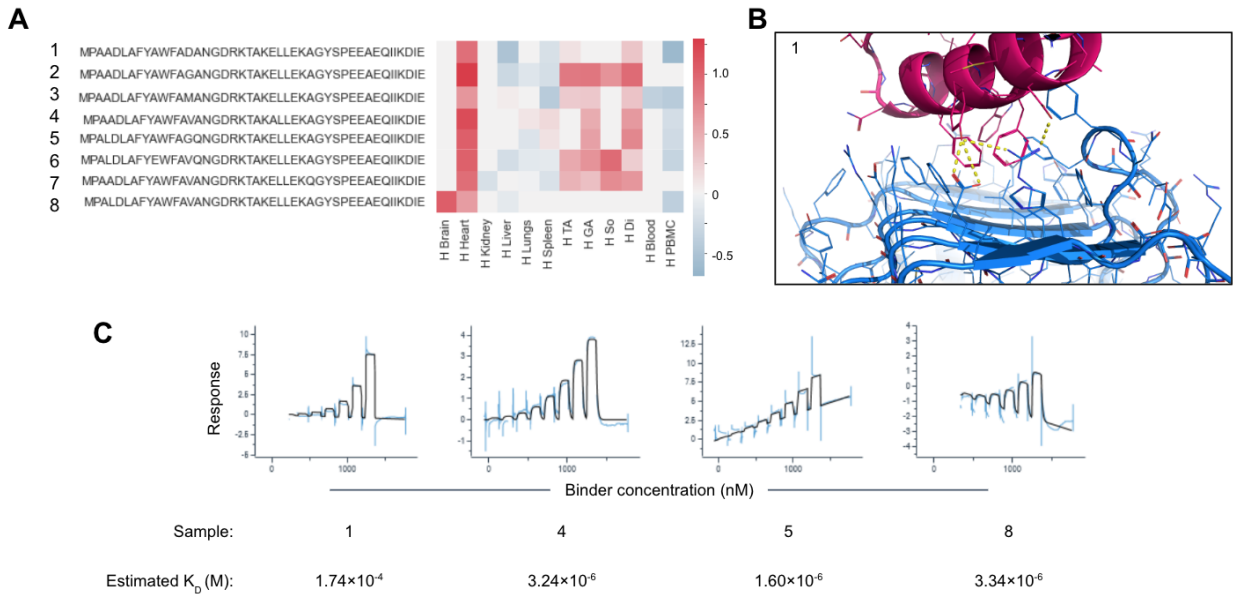


Figure 1. Identification and surface plasmon resonance (SPR) of insulin receptor minibinders.
(A) Eight sequences that were enriched in the muscles or brain, known to be designed to bind insulin receptor, were identified. Scale bar: $\log_{10}$(enrichment). **(B)** Representative AlphaFold2 prediction of Sample 1 (magenta) forming hydrophobic and hydrophilic (yellow dashes) molecular interactions with insulin receptor L1 domain (blue). **(C)** SPR of samples 1, 4, 5, 8 and estimated K_D with insulin receptor ectodomain. Blue: data, black: fit.

In examining sequence enrichment patterns above our empirically determined cutoff, we also identified several sequences enriched in the brain (**Fig. 2**). Though we do not yet know what these miniproteins were designed for, future directions could include searching for the molecular target and examining *in vivo* biodistribution of these miniproteins in a similar manner to Chapter 2, Figure 6. If successful, these miniproteins could be examined for the delivery of therapeutic molecules to the brain.

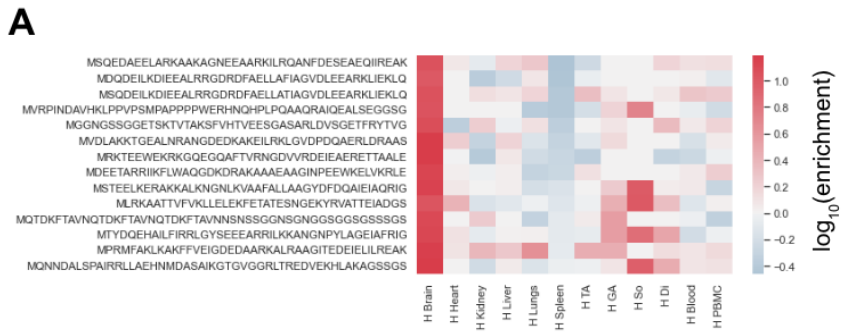


Figure 2. Identification of sequences highly enriched in the mouse brain.

***In vivo* selection of protein nanoparticles for functional delivery and genetic-based readouts**

While we reported therapeutically-relevant shifts in organ targeting in Chapter 2, and identified a validated insulin receptor minibinder, we did not show any functional delivery in this work. Additionally, due to the unstable nature of synthetic nucleocapsid mRNA beyond about 2 hours *in vivo*, we were unable to recover library mRNA from magnetically-isolated cell populations (data not shown). The ability to both screen for functional delivery (e.g., cargo release and genetic activation) and screen for localization to specific cell types (e.g., tumor-infiltrating leukocytes) would bring this platform technology closer to clinical utility.

Perhaps the most promising extension of this work to screen for functional delivery would be to apply the learnings from our colleagues in lipid nanoparticle engineering (Sago et al. 2018). Instead of encapsulating self-mRNA, we could engineer I53-50-v4v0 to encapsulate Cre recombinase mRNA or protein, enabling Cre-lox recombination. Cre is an enzyme that translocates to cell nuclei and interacts with genetically engineered loxP elements in DNA, introducing a site-specific recombination between loxP and a gene of interest (Madisen et al. 2010). Thus successful Cre-lox recombination necessitates the release of Cre mRNA or protein from the nanoparticle interior into the cells or tissues of interest. Since peak gene expression after Cre-lox recombination does not occur until at least 48 hours post-delivery, and I53-50-encapsulated mRNA is degraded within 2 hours following endocytosis (data not shown), we propose encapsulation of Cre protein for this study. To enable intracellular Cre delivery, we propose the incorporation of pH-responsive proteins and peptides into the nanoparticle scaffold (Boyken et al. 2019; Peeler et al. 2019). To enable the linkage between library variant and function, we propose using a liquid chromatography tandem-mass spectrometry (LC-MS/MS) barcoding scheme to link protein sequence to *in vivo* phenotype (instead of the typical DNA- or RNA-based barcoding to link genotype to phenotype). Short peptides could be appended to the Cre enzyme, then cleaved off and identified and relatively quantified by LC-MS/MS (by the unique fingerprint of peptide hydrophobicity and mass-to-charge ratio) (Egloff et al. 2019). The Cre-barcoding scheme could be used to identify variants of I53-50-v4v0 that

successfully deliver functional biomolecules to cells or tissues of interest *in vivo*, enabling future studies for therapeutic delivery (**Fig. 3**).

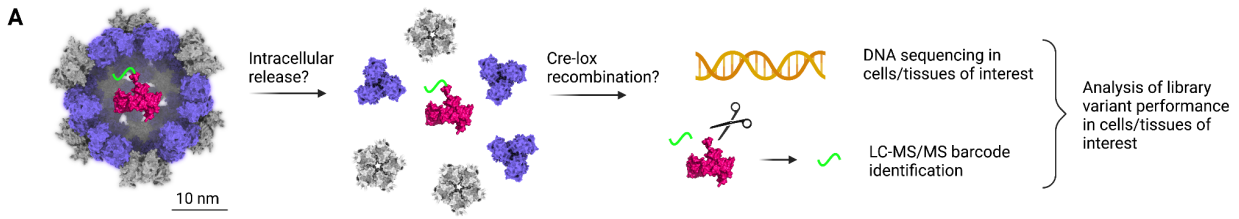


Figure 3. Encapsulation of Cre in I53-50 libraries for intracellular delivery.
(A) A half-shell view into the cavity of I53-50 (gray and slate) shows encapsulated Cre recombinase enzyme (pink, PDB ID 1NZB) with a peptide barcode (green). By co-expressing I53-50 and Cre in *E. coli*, analogous to Chapter 2 Figure 1, the barcode fused to Cre will correspond to the I53-50 library variant. The libraries will be screened for intracellular release of Cre and subsequent Cre-lox recombination in cells or tissues of interest via DNA sequencing and peptide barcode identification in parallel.

A major pitfall of the Cre-barcoding approach is insufficient Cre delivery efficiency. Though our groups have shown 0.5% delivery efficiency of a version of I53-50 engineered with pH responsive elements encapsulating prime editor guide RNA (data not shown), the *in vivo* delivery efficiency of our library variants might not be sufficient to measure meaningful genetic differences *in vivo*. If delivery efficiency remains prohibitively low, we could adapt or design from scratch a new nanoparticle with controlled disassembly properties and the ability to display exterior targeting domains. The nanoparticle could be scaffolded on de novo protein components or on functional enzymes themselves, like Cre or Cas9.

Overcoming the manufacturing difficulties of I53-50-v4v0-Cys_MMAE

As discussed in Chapter 1, *in vitro* assembly of protein nanoparticles enables precise control over nanoparticle composition and purity. Namely, *in vitro* assembly enables ultra-low endotoxin levels in the protein samples and partial display valency of targeting ligands like dnαvβ6 in Chapter 3 Figure 4D. Unfortunately, manufacturing *in vitro* assembled I53-50-v4v0-Cys_MMAE has required significant design and process development efforts over several years due to pentamer component instability. A fundamental solution to this hurdle could be returning to a bicistronic expression system, where the

pentamer and trimer of I53-50-v4v0-Cys are co-expressed in *E. coli* as in Chapter 2 Figure 1. By modifying the trimer component to add mass filling the interior cavity, such that free internal Cysteines are still available for MMAE conjugation, but no extra space is available in the cavity, it would be possible to physically preclude *E. coli* contaminants and endotoxin from getting trapped in the nanoparticle cavity. Enabling low-endotoxin production of bicistronic I53-50-v4v0-Cys would remove the need for separate expression and purification of protein components, thus sidestepping the petnamer component instability issue and alleviating manufacturing difficulties.

References

1. Boyken, Scott E., Mark A. Benhaim, Florian Busch, Mengxuan Jia, Matthew J. Bick, Heejun Choi, Jason C. Klima, et al. 2019. “De Novo Design of Tunable, pH-Driven Conformational Changes.” *Science* 364 (6441): 658–64.

2. Egloff, Pascal, Iwan Zimmermann, Fabian M. Arnold, Cedric A. J. Hutter, Damien Morger, Lennart Opitz, Lucy Poveda, et al. 2019. “Engineered Peptide Barcodes for in-Depth Analyses of Binding Protein Libraries.” *Nature Methods* 16 (5): 421–28.

3. Madisen, Linda, Theresa A. Zwingman, Susan M. Sunkin, Seung Wook Oh, Hatim A. Zariwala, Hong Gu, Lydia L. Ng, et al. 2010. “A Robust and High-Throughput Cre Reporting and Characterization System for the Whole Mouse Brain.” *Nature Neuroscience* 13 (1): 133–40.

4. Peeler, David J., Salina N. Thai, Yilong Cheng, Philip J. Horner, Drew L. Sellers, and Suzie H. Pun. 2019. “pH-Sensitive Polymer Micelles Provide Selective and Potentiated Lytic Capacity to Venom Peptides for Effective Intracellular Delivery.” *Biomaterials* 192 (February): 235–44.

5. Sago, Cory D., Melissa P. Lokugamage, Kalina Paunovska, Daryll A. Vanover, Christopher M. Monaco, Nirav N. Shah, Marielena Gamboa Castro, et al. 2018. “High-Throughput in Vivo Screen of Functional mRNA Delivery Identifies Nanoparticles for Endothelial Cell Gene Editing.” *Proceedings of the National Academy of Sciences of the United States of America* 115 (42): E9944–52.