Title

Investigating Designed Protein Nanocages as a Modular Platform for Synthetic Biology

Author

Lee, Eric Jinsuk

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Investigating Designed Protein Nanocages as a Modular Platform for Synthetic Biology

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy in Biochemistry, Molecular and Structural Biology

by

Eric Jinsuk Lee

2024

ABSTRACT OF THE DISSERTATION

Investigating Designed Protein Nanocages as a Modular Platform for Synthetic Biology

by

Eric Jinsuk Lee

Doctor of Philosophy in Biochemistry, Molecular and Structural Biology

University of California, Los Angeles, 2024

Professor Todd O. Yeates, Chair

Protein nanocages are abundant throughout nature, comprising many different shapes, sizes, and functions. However, despite their wide variety, these assemblies share key features, including a modifiable exterior surface, interior volume, and subunit-subunit interfaces. These characteristics, in conjunction with general stability and monodispersity, confer on nanocages an incredible capacity for synthetic biology applications. Already, both natural and designed protein nanocages have demonstrated exciting potential in therapeutics, imaging, and materials. The primary focus of the dissertation is to explore the utility of protein nanocages as modular molecular delivery platforms.

In the first chapter, I briefly review prior works that have studied protein nanocages as potential molecular carriers and delivery vehicles. The most prominent natural cages featured in the literature, ranging from small ferritins to large virus-like particles, are discussed. These examples encompass applications such as small molecule loading and delivery to more complex chemical and/or genetic modifications. The emergence of designed nanocages as an alternative to evolved nanocages for synthetic biology is also discussed. However, despite these incredible

advances, the general lack of specificity and modularity in protein cage-based encapsulation and delivery systems remains an ongoing challenge.

In the second chapter, I present the ligand-operable cage (LOC), a new type of protein cage that can "open" in response to a specific ligand. These designs aim to address the lack of a modular nanocage platform with a target-based disassembly mechanism. Uniquely, the opening mechanism of LOCs relies on the spatial orientation of surface-fused adaptors. In this study, we fused designed ankyrin repeat proteins (DARPins) to the surface of two different nanocages, sulfur oxygenase-reductase and the previously designed T33-51 cage. We also developed two readout assays indicating successful nanocage disassembly: one based on fluorescence unquenching of an encapsulated fluorophore, and one based on a split-NanoLuc luciferase assay.

In the third chapter, I discuss the design of three new cages to: 1) demonstrate the modularity of the LOC platform, and 2) explore their ability to bind cancer targets. The LOC variants, all based on the same T33-51 core, were designed to bind either BARD1, BRCA1, or KRAS, which are all prominently implicated in cancer. We modeled the overall structure of these designs and their adaptor orientations using AlphaFold-based ColabFold. We discuss the synthesis and purification of these LOC variants, characterize their assembly, and show preliminary results from target-triggered disassembly experiments.

In the fourth chapter, I aim to develop a combinatorial DNA-triggered bioswitch incorporating a protein nanocage. In a prior study, a NanoLuc-Cy3 BRET-beacon was designed to turn "on" or "off" in response to an oligonucleotide (ODN) trigger. We sought to expand this existing modular bioswitch by conditionally encapsulating the target ODN (i.e., the bioswitch trigger) inside of a ligand-operable cage. In this manner, we could add an additional layer of

modular control over the on-off state of the BRET-beacon. We detail the experimental scheme of this combinatorial switch, the synthesis of its components, and results from initial experiments.

In conclusion, the dissertation demonstrates the utility of protein nanocages as modular platforms for synthetic biology applications. Though prior works have exceptionally leveraged evolved cages, such as ferritin, VLPs, and other nonviral cages for molecular storage and delivery, we sought to address the general lack of a target-based, modular system. To this end, we demonstrate that our first LOC designs successfully released an encapsulated cargo upon specific target-binding using fluorescent and luminescent readout assays. Furthermore, we generated additional LOC variants that aimed to bind cancer-relevant proteins. Finally, we integrated a designed nanocage with a modular DNA sensor to create a combinatorial bioswitch. Taken together, these results, in consideration of experimental challenges and future directions, have exciting implications for future synthetic biology applications using protein nanocages.

This dissertation of Eric Jinsuk Lee is approved.

Feng Guo

Keriann Marie Backus

William Gelbart

Todd O. Yeates, Committee Chair

University of California, Los Angeles

2024

This work is dedicated to my loving family and friends

Thank you for supporting me all these years

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgements

I dedicate this dissertation to my amazing family and friends. Without their love and support, none of this work would have been possible. Furthermore, I would like to express my deepest gratitude to Professor Todd Yeates for offering me the opportunity to work in his laboratory for the past five years. I remember being first captivated by Professor Yeates' passion for his work during my graduate recruiting weekend at UCLA. My experience in the lab has helped me grow as both a scientist and person in so many ways and has further cultivated my passion for research.

I've also had the privilege of meeting many brilliant, kind peers who mentored and motivated me throughout my graduate career. In the Yeates lab, I am forever grateful for Dr. Kevin Cannon who was the first to guide me through my rotation and made me feel incredibly welcome in the lab. I must also thank my other lab mates for their friendship, helpful discussions, and fun nights out: Dr. Joshua Laniado, Dr. Jessica Ochoa, Dr. Justin Miller, Dr. Roger Castells, Dr. Kyle Meador, Matt Agdanowski, Morgan Gee, and our lab manager Inna Pashkov. I would be remiss not to mention UCLA staff scientists Dr. Mark Arbing, Dr. Duilio Cascio, Dr. Michael Sawaya, and Dr. Martin Phillips for their incredible mentorship and colorful daily banter. More, I must thank the wonderful mentors from my undergraduate research lab at Boston College who kickstarted my scientific journey: Dr. Laura-Anne Lowery, Dr. Sangmook Lee, Dr. Burcu Erdogan, Dr. Garrett Cammarata, Dr. Beth Bearce, Dr. Micaela Lasser, and the army of my undergraduate peers.

During my PhD candidacy, I also found an unexpected passion for teaching and mentorship. First and foremost, I thank Dr. Anne Hong-Hermesdorf for tolerating me for a total of five quarters while teaching undergrads how to perform basic biochemical techniques. I also

had the incredible opportunity to mentor Nika Gladkov, a brilliant undergraduate researcher who continues to accomplish great things. Without her help, much of this work would not have been possible.

Of course, I cannot forget my long-time friends who have kept my chin up these past several years. Thank you to Nick, Gary, Brian, Alex, Jojo, and Chris for helping me stay sane.

To Mom, Dad, and Nuna: I am so happy to make you proud everyday.

And finally, to Hannah, my beautiful, funny, and super-intelligent partner. I could write an entire chapter about how you've held my hand through this entire process.

In this dissertation, Chapter II is an adapted version of Lee EJ, Gladkov N, Miller JE & Yeates, TO. Design of Ligand-Operable Protein-Cages That Open Upon Specific Protein Binding. ACS Synth Biol. 13(1):157-167 (2024). DOI: 10.1021/acssynbio.3c00383. Original ligand-operable cages were designed by EJL and JEM. Experiments were performed by EJL and NG. Figures were prepared by EJL and TOY. The written manuscript was prepared by EJL and TOY. TOY was the principal investigator.

Vita

Education

2018, B.S. Biochemistry, Boston College

2020, M.S. Biochemistry and Molecular Biology, University of California, Los Angeles

Publications

Gladkov, N.; Scott, E. A.; Meador, K.; **Lee, E. J.**; Laganowsky, A. D.; Yeates, T. O.; Castells Graells, R. Design of a Symmetry-Broken Tetrahedral Protein Cage by a Method of Internal Steric Occlusion; preprint; Biochemistry, 2023. https://doi.org/10.1101/2023.11.08.566319.

Lee, E. J.; Gladkov, N.; Miller, J. E.; Yeates, T. O. Design of Ligand-Operable Protein-Cages That Open Upon Specific Protein Binding. *ACS Synth. Biol.* **2024**, *13* (1), 157–167. https://doi.org/10.1021/acssynbio.3c00383.

Chapter I: Introduction of protein nanocages and their emergent applications for synthetic biology as molecular carriers

Eric J. Lee[1]

[1] UCLA Department of Chemistry and Biochemistry, Los Angeles, CA 90095

ABSTRACT

Nature has evolved many extraordinary protein assemblies, including filaments, lattices and other complex architectures. Of these, protein nanocages have been a focus of many emergent synthetic biology and nanomaterials studies. Ranging from the relatively small, ubiquitous ferritin to larger, complex viral capsids, protein nanocages have already been used for many applications, such as therapeutics and imaging. Furthermore, advances in protein design have made it possible to create wholly new structures composed of non-native building blocks, expanding the design space for nanocages and their potential as a synthetic platform. This introductory chapter encompasses a brief history of protein nanocage research, with a focus on applications for molecular encapsulation and delivery. We discuss early works involving naturally evolved structures repurposed for various applications, emergent synthetic biology approaches, advent of designed cages, as well as ongoing challenges in the field.

INTRODUCTION

Nature is replete with complex supramolecular structures – biological assemblies comprising multiple smaller subunits. Architectures ranging from filaments to complex, asymmetric molecular machines have long inspired scientists to not only study and understand their structure and function, but also engineer them for new purposes.

A prominent class of supramolecular assembly is the protein nanocage. These self-assembling particles comprise repeats of one or more proteins in a symmetrical arrangement, taking the form of cubes and other polyhedra. Moreover, several characteristic features make them attractive as a platform for synthetic biology applications[1]. Firstly, protein nanocages possess an external surface that allows them to interact with their environment (Figure 1.1A,

left). Secondly, they contain an interior space that can encapsulate molecules of interest (Figure 1.1A, right). Finally, nanocages are composed of symmetrical, self-assembling subunits that cooperatively associate via interface sidechains (Figure 1.1C). In addition, each of these characteristic features can be further chemically or genetically modified. Protein nanocages are also robust, monodisperse, and biocompatible[2]. Prominent examples include the iron storage molecule ferritin and its homologues, as well as viral capsids. Advances in modern protein design have also made it possible to create entirely new, synthetic nanocages. Such designed assemblies comprise components that otherwise do not interact in nature.

Already, nanocages have been used for diverse studies in materials, therapeutics, and imaging (Figure 1.2)[3–8]. Notably, nanocages have displayed promise as platforms for molecular containment and delivery, ranging from the delivery of small molecule drugs to the encapsulation of whole proteins. Methods for functionalizing nanocages with non-native cargo range from co-assembly with non-native cargo in solution to more complicated genetic and/or chemical modifications. In this introductory chapter, I will discuss a brief overview of prior works exploring both natural and synthetic cages as delivery vehicles, as well as ongoing challenges in the field.

PRIOR WORKS

Ferritin

Of the many known evolved protein cages, ferritin is one of the most abundantly studied as a potential nanocarrier. This 24-mer octahedral assembly has been well-characterized in its structure and function as an iron storage vessel, as well as its ability to store other small molecules. One of the many early experiments leveraging ferritin as a potential therapeutic

vehicle was performed by Yang et al. in 2006. Their work utilized the pH-triggered assembly states of ferritin to encapsulate platinum-based cancer drugs[9]. At neutral pH, ferritin adopts its characteristic octahedral, enclosed state. However, at extremely acidic pH, the nanocage disassembles into its dissociated monomeric components. Such pH-dependent properties further support ferritin as an attractive platform for cancer-related delivery since environmental pH is often a differentiator between healthy and cancerous tissues. In a similar study Zhang et al. demonstrated the ability of ferritin to encapsulate anthocyanins, a class of molecule with high therapeutic potential, but low cellular stability[10]. This study importantly demonstrated the ability of nanocages to protect and delivery otherwise serum-unstable cargo, increasing cellular uptake. Additional work with ferritin by Chen et al. demonstrated that one could modify the cage's pH-dependent disassembly profile via cleavage of the C-terminal helix, allowing the cage to disassemble at a less acidic pH of 4[11]. The same group also demonstrated that DOX-loaded ferritin can cross the blood-brain barrier[12].

Viruses

Virus-like particles are the next prominent class used for nanocarrier studies. Viral capsids already demonstrate important biological functions regarding molecular storage, as they comprise up to hundreds of subunits made to transport viral genetic cargo. Adenoassociated virus-2, brome mosaic, MS2 phage, cowpea mosaic virus, cowpea chlorotic mosaic virus, papillomavirus L1 capsid, turnip yellow mosaic virus, cucumber mosaic virus, and human hepatitis B virus are some examples that have been studied for biotechnology and biomedical applications, namely the packaging of exogenous nucleic acids[13–17]. For example, Biddlecome et al. have recently demonstrated the therapeutic potential of mRNA-carrying VLPs.[18] Owing to

their size, viruses have been investigated as whole protein carriers in synthetic biology studies. Generally, these systems involve an anchoring or shuttling mechanism to load the nanocage with significantly greater efficiency. For example, Schoonen et al. demonstrated the sortase-mediated fusion of whole enzymes on the inside of a capsid protein[19]. Packaging of proteins into viral capsids has also been achieved via fusion to Gag proteins[20]. Similarly, though not viruses, bacterial microcompartments have also become a focus of synthetic biology experiments, with their ability to encapsulate whole sets of enzymes[21–23].

Notable non-viral cages

Vaults are relatively large ribonucleoproteins that have been used as nanocapsules[24]. Matsumoto et al. have demonstrated synthetically modified vault protein cages with stimuli-responsive polymers[24]. These findings have culminated in a new type of hybrid material, where the structure of the synthetic particle is based on the vault protein, but the stimuli response is dictated by the nature of the conjugated polymer. Vaults have also been identified as candidates for protein-based drug loading due to their size[26].

E2 protein cages are 60-subunit assemblies derived from pyruvate dehydrogenase complex**.** Dalmau et al. and Peng et al. performed separate experiments involving subunit-subunit interface mutations of a previously characterized E2 scaffold to create a pH-responsive assembly and disassembly system[27,28]. These studies led to disassembled components that were more soluble and potentially more useful in biologically relevant applications. Ren et al. also utilized the E2 protein cage to encapsulate maleimide-carrying fluorophore and doxorubicin molecules via internal cysteine residues, further demonstrating the utility of E2 cages delivery vehicles[29].

Lumazine synthase, another icosahedral cage, has also been studied as a molecular delivery vehicle. Lilivivat et al. engineered a variant that could encapsulate nucleic acids via inserted positive residues in the cage interior[30]. Ra et al. used lumazine synthases to deliver antigens to dendritic cells via recombinant fusion of antigenic peptides[31]. Tetter et al. used directed evolution to create larger lumazine synthase cages to package its own encoding RNA[32].

Other synthetic biology studies

Increasingly diverse and novel methods for encapsulation and/or cage assembly have emerged. For example, Hoersch et al. created a light-gated nanocage by engineering ATP-driven chaperonins[33]. This was achieved by incorporating photoswitchable azobenzene-based molecules into the structure of the protein complex. This culminated in a protein nanocage that underwent conformational changes in response to different wavelengths of light. Kang et al. incorporated a thrombin cleavage site into ferritin to create a protease-responsive delivery platform[34]. In this system, protease digestion of the incorporated cleavage site caused the "excision" of an internally encapsulated alpha helix. This helix was conjugated to a fluorophore, which was released from the cage interior post-cleavage. It was also demonstrated that hydrostatic pressure could be used as a trigger for cage disassembly[35]. Other methods such as metal coordination, disulfide formation, and allosteric elements have been studied as stimuli for engineered nanocage assembly and disassembly[36–39].

Designed nanocages

Recent advances in protein design have made it possible to create entirely new nanocages out of non-native protein building blocks. The first of these designed nanocages comprised

multiple copies of a recombinant fusion between two trimeric proteins, leading to a novel tetrahedral self-assembly[40]. This work revealed that independent proteins could be made to form a symmetric self-assembly provided that their recombinant fusion fit the necessary geometric constraints to form a cage. An alternative strategy also emerged where, instead of a fusion, computational interface design was used to make otherwise non-interacting oligomers bind each other at the correct geometry[41–46].

It was recently demonstrated that a designed fusion-type nanocage could be modified to be cleaved by proteases[47]. Basing their design on a previously designed cage core, Miller et al. incorporated a thrombin and asparagine protease cleavage site in separate design cases. This allowed them to cause proteolytic cleavage and subsequent disassembly of the cage. These findings illustrated that designed nanocages can be broken or opened in response to specific, disease-relevant proteases.

One disadvantage of designed nanocages may be that they are typically not designed with multiple, defined assembly configurations in mind. In fact, many designed nanocages have found their use in scaffolding or materials. On the other hand, with ever-emergent structure prediction methods and software, synthetic cages theoretically have significantly greater design space. New tools such as ProteinMPNN have even made it possible to create protein cage sequences using deep learning[48].

ONGOING CHALLENGES AND FUTURE DIRECTIONS

The primary challenge for the wider, general usage of nanocages as nanocarriers is the general lack of target-specific strategies for cargo release. Existing work primarily relies on drastic shifts in environmental conditions, which may not be amenable for certain biological or

biomedical studies. For example, a ferritin cage-based nanocarrier that is dependent upon pH changes to release its cargo may not be applicable in an environment that requires cargo to be both stored and released under physiological conditions. A more desirable nanocarrier would instead respond to a particular ligand (e.g., a protein of interest) pertinent to the environment of interest to release its cargo. The triggerable cages by Kang et al. and Miller et al. are examples that have sought to address this design challenge by making the cages respond specifically to certain proteases.

Moreover, there is no known nanocage-based platform that offers sufficient modularity for widespread applications in synthetic biology. For example, a nanocage engineered to open in response to a single stimulus may only find utility in said single application. In this case, entirely new cages would have to be laboriously designed or engineered for alternative applications. The ideal nanocage-based platform to mitigate this issue would be to base all designs off a "common" cage core, onto which minor, *intuitive* mutations and/or modifications can be made. This way, a single design can be readily "reprogrammed" to respond to many different stimuli.

Despite these current challenges, protein nanocages offer great potential in biotechnology and medicine as molecular vehicles. Their biocompatibility and multiple modification sites have made them a very attractive platform for scientists in creating novel materials and therapeutics. Both viral and non-viral cages of varying shapes and sizes have already demonstrated their ability to encapsulate cargo primarily via spontaneous coassembly or covalent linkage. Alternative methods for creating cages that can assemble or disassemble under unique conditions are ever emergent. Finally, recent studies have shown more targeted strategies for cage disassembly, such as specific protease degradation. Nevertheless, many challenges, particularly

regarding target specificity and modularity, remain. The following work discussed in the dissertation proposes design solutions to these important questions.

FIGURES

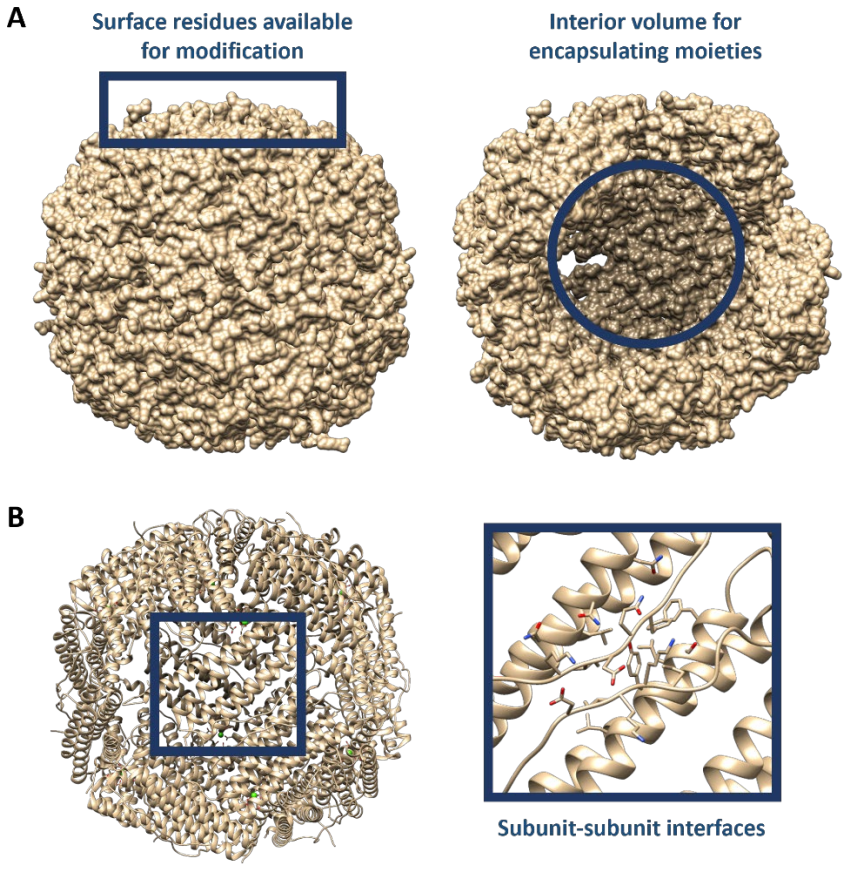

Figure 1.1. Structural overview of an example nanocage (ferritin) and its characteristics. A) Left, surface map of nanocage, available for chemical or genetic modifications; right, partial surface map exhibiting interior volume, available for encapsulating moieties. B) Ribbon model of nanocage to demonstrate interface residues (right inset) that contribute to overall cooperative assembly.

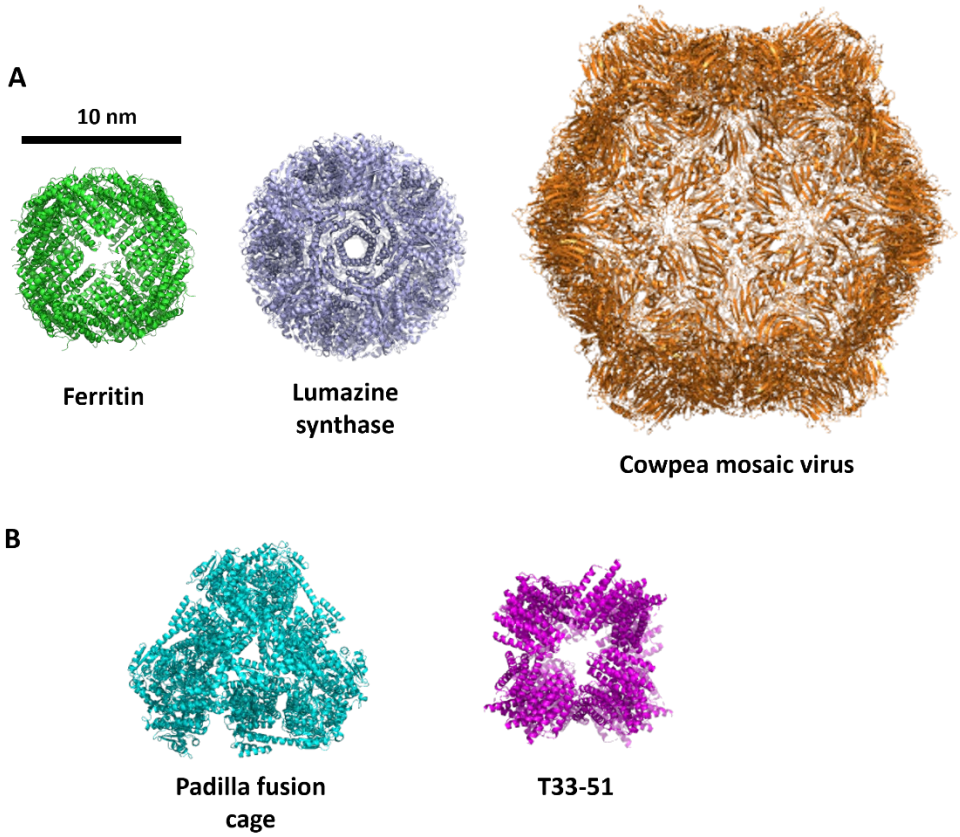

Figure 1.2. Select examples of nanocages utilized or proposed for encapsulation and delivery applications. A) Natural cages from left to right: human ferritin, lumazine synthase from *Aquifex Aeolicus*, cowpea mosaic virus; PDB: 4V6B; PDB: 1HQK; PDB: 1NY7. B) Designed cages from left to right: "Padilla" single-component nanocage, T33-51 two-component nanocage; PDB: 4QES; PDB: 5CY5. Structures are approximately scaled.

REFERENCES

(1) Cannon, K. A.; Ochoa, J. M.; Yeates, T. O. High-Symmetry Protein Assemblies: Patterns and Emerging Applications. *Curr. Opin. Struct. Biol.* **2019**, *55*, 77–84. https://doi.org/10.1016/j.sbi.2019.03.008.

(2) Naahidi, S.; Jafari, M.; Edalat, F.; Raymond, K.; Khademhosseini, A.; Chen, P. Biocompatibility of Engineered Nanoparticles for Drug Delivery. *J. Controlled Release* **2013**, *166* (2), 182–194. https://doi.org/10.1016/j.jconrel.2012.12.013.

(3) Marcandalli, J.; Fiala, B.; Ols, S.; Perotti, M.; de van der Schueren, W.; Snijder, J.; Hodge, E.; Benhaim, M.; Ravichandran, R.; Carter, L.; Sheffler, W.; Brunner, L.; Lawrenz, M.; Dubois, P.; Lanzavecchia, A.; Sallusto, F.; Lee, K. K.; Veesler, D.; Correnti, C. E.; Stewart, L. J.; Baker, D.; Loré, K.; Perez, L.; King, N. P. Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **2019**, *176* (6), 1420-1431.e17. https://doi.org/10.1016/j.cell.2019.01.046.

(4) McConnell, S. A.; Cannon, K. A.; Morgan, C.; McAllister, R.; Amer, B. R.; Clubb, R. T.; Yeates, T. O. Designed Protein Cages as Scaffolds for Building Multienzyme Materials. *ACS Synth. Biol.* **2020**, *9* (2), 381–391. https://doi.org/10.1021/acssynbio.9b00407.

(5) Liu, Y.; Gonen, S.; Gonen, T.; Yeates, T. O. Near-Atomic Cryo-EM Imaging of a Small Protein Displayed on a Designed Scaffolding System. *Proc. Natl. Acad. Sci.* **2018**, *115* (13), 3362–3367. https://doi.org/10.1073/pnas.1718825115.

(6) Liu, Y.; Huynh, D. T.; Yeates, T. O. A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold. *Nat. Commun.* **2019**, *10* (1), 1864. https://doi.org/10.1038/s41467-019-09836-0.

(7) 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. https://doi.org/10.1021/acschembio.1c00007.

(8) Hills, R. A.; Tan, T. K.; Cohen, A. A.; Keeffe, J. R.; Keeble, A. H.; Gnanapragasam, P. N. P.; Storm, K. N.; Hill, M. L.; Liu, S.; Gilbert-Jaramillo, J.; Afzal, M.; Napier, A.; James, W. S.; Bjorkman, P. J.; Townsend, A. R.; Howarth, M. Multiviral Quartet Nanocages Elicit Broad Anti-Coronavirus Responses for Proactive Vaccinology. *BioRxiv Prepr. Serv. Biol.* **2023**, 2023.02.24.529520. https://doi.org/10.1101/2023.02.24.529520.

(9) Yang, Z.; Wang, X.; Diao, H.; Zhang, J.; Li, H.; Sun, H.; Guo, Z. Encapsulation of Platinum Anticancer Drugs by Apoferritin. *Chem. Commun.* **2007**, No. 33, 3453. https://doi.org/10.1039/b705326f.

(10) Zhang, T.; Lv, C.; Chen, L.; Bai, G.; Zhao, G.; Xu, C. Encapsulation of Anthocyanin Molecules within a Ferritin Nanocage Increases Their Stability and Cell Uptake Efficiency. *Food Res. Int.* **2014**, *62*, 183–192. https://doi.org/10.1016/j.foodres.2014.02.041.

(11) Chen, H.; Zhang, S.; Xu, C.; Zhao, G. Engineering Protein Interfaces Yields Ferritin Disassembly and Reassembly under Benign Experimental Conditions. *Chem. Commun.* **2016**, *52* (46), 7402–7405. https://doi.org/10.1039/C6CC03108K.

(12) Chen, Z.; Zhai, M.; Xie, X.; Zhang, Y.; Ma, S.; Li, Z.; Yu, F.; Zhao, B.; Zhang, M.; Yang, Y.; Mei, X. Apoferritin Nanocage for Brain Targeted Doxorubicin Delivery. *Mol. Pharm.* **2017**, *14* (9), 3087–3097. https://doi.org/10.1021/acs.molpharmaceut.7b00341.

(13) Lee, L.; Wang, Q. Adaptations of Nanoscale Viruses and Other Protein Cages for Medical Applications. *Nanomedicine Nanotechnol. Biol. Med.* **2006**, *2* (3), 137–149. https://doi.org/10.1016/j.nano.2006.07.009.

(14) Fang, C.-Y.; Lin, P. Y.; Ou, W.-C.; Chen, P.-L.; Shen, C.-H.; Chang, D.; Wang, M. Analysis of the Size of DNA Packaged by the Human JC Virus-like Particle. *J. Virol. Methods* **2012**, *182* (1–2), 87–92. https://doi.org/10.1016/j.jviromet.2012.03.017.

(15) Georgens, C.; Weyermann, J.; Zimmer, A. Recombinant Virus Like Particles as Drug Delivery System. *Curr. Pharm. Biotechnol.* **2005**, *6* (1), 49–55. https://doi.org/10.2174/1389201053167202.

(16) Müller, O. J.; Kaul, F.; Weitzman, M. D.; Pasqualini, R.; Arap, W.; Kleinschmidt, J. A.; Trepel, M. Random Peptide Libraries Displayed on Adeno-Associated Virus to Select for Targeted Gene Therapy Vectors. *Nat. Biotechnol.* **2003**, *21* (9), 1040–1046. https://doi.org/10.1038/nbt856.

(17) Peng, K.-W.; Holler, P. D.; Orr, B. A.; Kranz, D. M.; Russell, S. J. Targeting Virus Entry and Membrane Fusion through Specific Peptide/MHC Complexes Using a High-Affinity T-Cell Receptor. *Gene Ther.* **2004**, *11* (15), 1234–1239. https://doi.org/10.1038/sj.gt.3302286.

(18) Biddlecome, A.; Habte, H. H.; McGrath, K. M.; Sambanthamoorthy, S.; Wurm, M.; Sykora, M. M.; Knobler, C. M.; Lorenz, I. C.; Lasaro, M.; Elbers, K.; Gelbart, W. M. Delivery of Self-Amplifying RNA Vaccines in in Vitro Reconstituted Virus-like Particles. *PLOS ONE* **2019**, *14* (6), e0215031. https://doi.org/10.1371/journal.pone.0215031.

(19) Schoonen, L.; Nolte, R. J. M.; van Hest, J. C. M. Highly Efficient Enzyme Encapsulation in a Protein Nanocage: Towards Enzyme Catalysis in a Cellular Nanocompartment Mimic. *Nanoscale* **2016**, *8* (30), 14467–14472. https://doi.org/10.1039/C6NR04181G.

(20) 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.* **2011**, *108* (41), 16998–17003. https://doi.org/10.1073/pnas.1101874108.

(21) Quin, M. B.; Perdue, S. A.; Hsu, S.-Y.; Schmidt-Dannert, C. Encapsulation of Multiple Cargo Proteins within Recombinant Eut Nanocompartments. *Appl. Microbiol. Biotechnol.* **2016**, *100* (21), 9187–9200. https://doi.org/10.1007/s00253-016-7737-8.

(22) Tsai, S. J.; Yeates, T. O. Bacterial Microcompartments. In *Progress in Molecular Biology and Translational Science*; Elsevier, 2011; Vol. 103, pp 1–20. https://doi.org/10.1016/B978-0-12-415906-8.00008-X.

(23) Chowdhury, C.; Sinha, S.; Chun, S.; Yeates, T. O.; Bobik, T. A. Diverse Bacterial Microcompartment Organelles. *Microbiol. Mol. Biol. Rev.* **2014**, *78* (3), 438–468. https://doi.org/10.1128/MMBR.00009-14.

(24) Kedersha, N. L.; Rome, L. H. Isolation and Characterization of a Novel Ribonucleoprotein Particle: Large Structures Contain a Single Species of Small RNA. *J. Cell Biol.* **1986**, *103* (3), 699–709. https://doi.org/10.1083/jcb.103.3.699.

(25) Matsumoto, N. M.; Prabhakaran, P.; Rome, L. H.; Maynard, H. D. Smart Vaults: Thermally-Responsive Protein Nanocapsules. *ACS Nano* **2013**, *7* (1), 867–874. https://doi.org/10.1021/nn3053457.

(26) Han, M.; Kickhoefer, V. A.; Nemerow, G. R.; Rome, L. H. Targeted Vault Nanoparticles Engineered with an Endosomolytic Peptide Deliver Biomolecules to the Cytoplasm. *ACS Nano* **2011**, *5* (8), 6128–6137. https://doi.org/10.1021/nn2014613.

(27) Dalmau, M.; Lim, S.; Wang, S.-W. pH-Triggered Disassembly in a Caged Protein Complex. *Biomacromolecules* **2009**, *10* (12), 3199–3206. https://doi.org/10.1021/bm900674v.

(28) Peng, T.; Lim, S. Trimer-Based Design of pH-Responsive Protein Cage Results in Soluble Disassembled Structures. *Biomacromolecules* **2011**, *12* (9), 3131–3138. https://doi.org/10.1021/bm2005438.

(29) Ren, D.; Kratz, F.; Wang, S.-W. Protein Nanocapsules Containing Doxorubicin as a pH-Responsive Delivery System. *Small* **2011**, *7* (8), 1051–1060. https://doi.org/10.1002/smll.201002242.

(30) Lilavivat, S.; Sardar, D.; Jana, S.; Thomas, G. C.; Woycechowsky, K. J. *In Vivo* Encapsulation of Nucleic Acids Using an Engineered Nonviral Protein Capsid. *J. Am. Chem. Soc.* **2012**, *134* (32), 13152–13155. https://doi.org/10.1021/ja302743g.

(31) Ra, J.-S.; Shin, H.-H.; Kang, S.; Do, Y. Lumazine Synthase Protein Cage Nanoparticles as Antigen Delivery Nanoplatforms for Dendritic Cell-Based Vaccine Development. *Clin. Exp. Vaccine Res.* **2014**, *3* (2), 227. https://doi.org/10.7774/cevr.2014.3.2.227.

(32) Tetter, S.; Terasaka, N.; Steinauer, A.; Bingham, R. J.; Clark, S.; Scott, A. J. P.; Patel, N.; Leibundgut, M.; Wroblewski, E.; Ban, N.; Stockley, P. G.; Twarock, R.; Hilvert, D. Evolution of a Virus-like Architecture and Packaging Mechanism in a Repurposed Bacterial Protein. *Science* **2021**, *372* (6547), 1220–1224. https://doi.org/10.1126/science.abg2822.

(33) Hoersch, D.; Roh, S.-H.; Chiu, W.; Kortemme, T. Reprogramming an ATP-Driven Protein Machine into a Light-Gated Nanocage. *Nat. Nanotechnol.* **2013**, *8* (12), 928–932. https://doi.org/10.1038/nnano.2013.242.

(34) Kang, Y. J.; Park, D. C.; Shin, H.-H.; Park, J.; Kang, S. Incorporation of Thrombin Cleavage Peptide into a Protein Cage for Constructing a Protease-Responsive Multifunctional Delivery Nanoplatform. *Biomacromolecules* **2012**, *13* (12), 4057–4064. https://doi.org/10.1021/bm301339s.

(35) Le Vay, K.; Carter, B. M.; Watkins, D. W.; Dora Tang, T.-Y.; Ting, V. P.; Cölfen, H.; Rambo, R. P.; Smith, A. J.; Ross Anderson, J. L.; Perriman, A. W. Controlling Protein Nanocage Assembly with Hydrostatic Pressure. *J. Am. Chem. Soc.* **2020**, *142* (49), 20640–20650. https://doi.org/10.1021/jacs.0c07285.

(36) 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.; Wu, D.; Wrobel, T. P.; Fineberg, A.; Kushwah, M. S.; Kelemen, M.; Vavpetič, P.; Pelicon, P.; Kukura, P.; Benesch, J. L. P.; Iwasaki, K.; Heddle, J. G. An Ultra-Stable Gold-Coordinated Protein Cage Displaying Reversible Assembly. *Nature* **2019**, *569* (7756), 438–442. https://doi.org/10.1038/s41586-019-1185-4.

(37) Zang, J.; Chen, H.; Zhang, X.; Zhang, C.; Guo, J.; Du, M.; Zhao, G. Disulfide-Mediated Conversion of 8-Mer Bowl-like Protein Architecture into Three Different Nanocages. *Nat. Commun.* **2019**, *10* (1), 778. https://doi.org/10.1038/s41467-019-08788-9.

(38) Xu, M.; Zeng, R.; Xiang, J.; Yan, Q. Self-Assembly of Switchable Protein Nanocages via Allosteric Effect. *CCS Chem.* **2021**, *3* (8), 2223–2232. https://doi.org/10.31635/ccschem.020.202000437.

(39) Ni, T. W.; Tezcan, F. A. Structural Characterization of a Microperoxidase inside a Metal-Directed Protein Cage. *Angew. Chem. Int. Ed Engl.* **2010**, *49* (39), 7014–7018. https://doi.org/10.1002/anie.201001487.

(40) 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.* **2001**, *98* (5), 2217–2221. https://doi.org/10.1073/pnas.041614998.

(41) 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. https://doi.org/10.1038/nature13404.

(42) Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C.; Cascio, D.; Yeates, T. O.; Gonen, T.; King, N. P.; Baker, D. Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes. *Science* **2016**, *353* (6297), 389–394. https://doi.org/10.1126/science.aaf8818.

(43) Yeates, T. O. Geometric Principles for Designing Highly Symmetric Self-Assembling Protein Nanomaterials. *Annu. Rev. Biophys.* **2017**, *46*, 23–42. https://doi.org/10.1146/annurev-biophys-070816-033928.

(44) Laniado, J.; Yeates, T. O. A Complete Rule Set for Designing Symmetry Combination Materials from Protein Molecules. *Proc. Natl. Acad. Sci.* **2020**, *117* (50), 31817–31823. https://doi.org/10.1073/pnas.2015183117.

(45) Laniado, J.; Cannon, K. A.; Miller, J. E.; Sawaya, M. R.; McNamara, D. E.; Yeates, T. O. Geometric Lessons and Design Strategies for Nanoscale Protein Cages. *ACS Nano* **2021**, *15* (3), 4277–4286. https://doi.org/10.1021/acsnano.0c07167.

(46) Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N. Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science* **2013**, *340* (6132), 595–599. https://doi.org/10.1126/science.1233936.

(47) Miller, J. E.; Srinivasan, Y.; Dharmaraj, N. P.; Liu, A.; Nguyen, P. L.; Taylor, S. D.; Yeates, T. O. Designing Protease-Triggered Protein Cages. *J. Am. Chem. Soc.* **2022**, *144* (28), 12681–12689. https://doi.org/10.1021/jacs.2c02165.

(48) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; De Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. Robust Deep Learning–Based Protein Sequence Design Using ProteinMPNN. *Science* **2022**, *378* (6615), 49–56. https://doi.org/10.1126/science.add2187.

Chapter II. Design of Ligand-Operable Protein-Cages That Open Upon Specific Protein Binding

Eric J. Lee[1], Nika Gladkov[1], Justin E. Miller[2], Todd O. Yeates[1,2,3]

[1] UCLA Department of Chemistry and Biochemistry, Los Angeles, CA 90095

[2] UCLA Molecular Biology Institute, Los Angeles, CA 90095

[3] UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095

ABSTRACT

Protein nanocages have diverse applications in medicine and biotechnology, including molecular delivery. However, although numerous studies have demonstrated the ability of protein nanocages to encapsulate various molecular species, limited methods are available for subsequently opening a nanocage for cargo release under specific conditions. A modular platform with a specific protein-target-based mechanism of nanocage opening is notably lacking. To address this important technology gap, we present a new class of designed protein cages, the Ligand-Operable Cage (LOC). LOCs primarily comprise a protein nanocage core and a fused surface binding adaptor. The geometry of the LOC is designed so that binding of a target protein ligand (or multiple copies thereof) to the surface binder is sterically incompatible with retention of the assembled state of the cage. Therefore, the tight binding of a target ligand drives cage disassembly by mass action, subsequently exposing the encapsulated cargo. LOCs are modular; direct substitution of the surface binder sequence can reprogram the nanocage to open in response to any target protein ligand of interest. We demonstrate these design principles using both a natural and a designed protein cage as the core, with different proteins acting as the triggering ligand and with different reporter readouts ─ fluorescence unquenching and luminescence ─ for cage disassembly. These developments advance the critical problem of targeted molecular delivery and detection.

INTRODUCTION

Advances in rational protein design methods have made it possible to create novel proteins, enzymes, and higher-order complexes.[1-6] Exploiting principles of symmetric self-assembly, sophisticated supramolecular architectures in the shapes of cubes and other polyhedra

can be produced from much simpler protein building blocks.[6-16] Comparable supramolecular structures are ubiquitous throughout nature, as exemplified by microtubules, polyhedral viruses, S-layer proteins, etc.[17–21] Recent design studies have led to wholly new sets of novel protein architectures, opening vast avenues of exploration for synthetic biology.

Chief among these supramolecular classes is the protein nanocage. As an engineering platform, nanocages offer two key features ─ a symmetrical surface for polyvalent display and an interior volume for molecular containment ─ making them versatile for many synthetic biology studies.[22] Already, nanocages have been leveraged for exciting applications in biomaterials, therapeutics, and imaging.[23–28] Notably, nanocages have demonstrated potential as molecular carriers, able to encapsulate functional moieties ranging from small drugs to macromolecules.[29–32]

While considerable prior work has focused on creating robustly assembling protein cages, designing highly targeted mechanisms of cage disruption remains a challenge, thereby limiting the use of nanocages as molecular delivery vehicles. Previously explored methods of nanocage opening include changes in pH, hydrostatic pressure, reducing agents, light, and chelators.[33–41] Such approaches are suitable when the desired application admits control of bulk properties and external conditions. Interest in higher levels of specificity ─ e.g., discriminating between different cell states or cell types ─ calls for new ideas. Two recent works have explored the design of protein cages that are sensitive to (i.e., destroyed by) the presence of specific proteases.[42,43] This method offers delivery and response mechanisms for systems in which specific proteases are naturally present or where they can be expressed. However, even more specific systems for the targeted opening could be transformative. A platform with ultimate

specificity would come from a protein cage that could be readily engineered to open in response to the presence of essentially any protein of interest.

Here, we present Ligand-Operable Cages (LOCs), a new and modular class of protein nanocages with a capacity for target-based disruption. LOCs are primarily composed of an assembled nanocage core with fused binding adaptors on their surface. The binding adaptors are configured in such a way that binding of a cognate protein to the binding adaptor is incompatible with retention of the overall nanocage assembly. This feature leads to cage opening upon binding a specific target protein. We describe key design principles for LOCs and present two prototype cases, one based on a natural protein cage, and another on a designed nanocage core.[44,45] We demonstrate the generality of the system with respect to the nanocage core, the programmed target ligand, and the molecular mechanism for signaling disassembly.

RESULTS AND DISCUSSION

LOC Design Principles

Our design ideas aim to create protein nanocages with external binding adaptors specifically situated so that the nanocage opens upon binding a cognate protein ligand. The platform comprises two essential components: (1) a symmetric self-assembling protein cage core built from multiple copies of one or sometimes two distinct subunit types and (2) an adaptor protein for binding the protein ligand of interest (Figure 2.1A). Suitable cage cores typically take the form of a cubic shape (or other Platonic solid), with an interior cavity.

Connecting the adaptor to the nanocage core subunit by using genetic fusion leads to assemblies with defined oligomeric compositions and configurations. Multiple copies of the binding adaptor are presented on the cage surface. The specific 3-dimensional configuration of

the protruding adaptors plays a critical role in the behavior of the nanocage. In the absence of the target ligand, the unbound form of the nanocage suffers from no steric collisions. In contrast, and by design, the binding of the cognate ligand, or many copies thereof, is sterically incompatible with retention of cage assembly (Figure 2.1B). This creates a geometric conflict. According to established principles of thermodynamics, addition of a high-affinity ligand to such a system ─ assuming a sufficient degree of dynamics are at play ─ will drive the system to disassemble by mass-action. Somewhat related ideas involving steric collision have been employed in other molecular systems.[46–48] A third, optional, component is a passenger molecule or other inward-facing molecular moiety ─ e.g., a fluorescent molecule or other bioactive marker ─ which can be used as a readout for assembly disruption. The incorporation of specific interior components endows LOCs with diverse utilities.

For our design studies, we used Designed Ankyrin Repeat Proteins (DARPins) as our binding adaptor protein. Similar to nanobodies and related immunoglobulin-based protein domains, DARPins have proven highly amenable to sequence selection in their loop regions ─ producing variants that bind tightly to practically any protein ligand of interest.[49–51] DARPins, however, offer a distinctive feature: their highly alpha helical composition suits them for making continuous alpha helical connections to other protein domains that also have terminal helices.[25,26,52–56] Such an unbroken helix spanning both components contributes to rigidity. We reasoned that connecting a DARPin adaptor domain to the cage core using a continuous alpha helical linker would help physical forces (originating from sterically restricted ligand binding) propagate from the adaptors to the cage core. Based on these ideas, we hypothesized that a LOC platform could be programmed to open in response to diverse protein ligands, with disassembly then signaling through various reporter mechanisms.

Design of a LOC Based on a Natural Protein Cage

We elected to begin our LOC design study using a natural nanocage core. We employed the oligomeric enzyme sulfur oxygenase reductase (SOR), which forms a 24-subunit cube-shaped assembly with octahedral symmetry.[44] Among a small set of natural protein assemblies with cubic symmetry, SOR was an appropriate candidate because it presents C-terminal helices on its surface, permitting fusion to the N-terminal helix of a DARPin adaptor.[22] For this trial, we used a previously reported DARPin sequence that binds superfolder green fluorescent protein (sfGFP).[26,52,57,58]

We generated a model of the SOR-based LOC (SOR-LOC) using PyMOL. To achieve this, each SOR monomer's C-terminal helix was aligned with and then extended into an ideal polyalanine helix. The N-terminal helix of the sfGFP-binding DARPin was then aligned with the opposite end of the ideal helix. This resulted in a final construct comprising an SOR monomer fused to the sfGFP-binding DARPin via an Ala–Gln linker. The DARPins were configured such that their binding sites would face inward, e.g., toward an axis of symmetry or another cage protein subunit.

As modeled, the full SOR-LOC assembly is composed of a native octahedral core structure with 24 protruding DARPin domains (Figure 2.2A, right). A closer view at one of the four-fold axes of the SOR-LOC octahedral assembly shows no steric clashing in the absence of the cognate target, sfGFP (Figure 2B, left). However, in a hypothetical sfGFP-bound form, major steric clashes would occur between each DARPin-bound sfGFP molecule and the neighboring DARPin (Figure 2.2B, right). Thus, the sfGFP-bound form would be incompatible with retention

of the cage core assembly. Therefore, we expected the high affinity binding of sfGFP to drive disassembly of the SOR-LOC.

We performed size-exclusion chromatography (SEC) to verify the assembly and subsequent sfGFP-induced disassembly of SOR-LOC. SEC analysis of SOR-LOC alone resulted in an elution volume profile featuring a prominent peak corresponding to the full assembly, approximately 1.26 MDa, in addition to some partial subassemblies (Figure 2.2C). After the addition of 50 μM sfGFP (2× molar excess of SOR-LOC DARPins), we observed that essentially all SOR-LOC and sfGFP in the sample aggregated. Following sample centrifugation and clarification of aggregates, the size-exclusion chromatogram of SOR-LOC incubated with sfGFP showed little to no A280 signal corresponding to assembled cages or partial subassemblies. This indicated that sfGFP binding caused almost complete disassembly and subsequent aggregation of the assembled SOR-LOC cages.

Furthermore, we isolated the peak fraction corresponding to assembled SOR-LOC cages and examined the sample via negative-stain electron microscopy. We observed relatively monodisperse assemblies of approximately the correct size (15 nm diameter) at high particle density (Figure 2.2D, left). However, after we mixed the isolated SOR-LOC cages with sfGFP (2× molar excess of SOR-LOC DARPins), we observed an apparent reduction in monodispersity and particle density as well as an apparent increase of aggregates (Figure 2.2D, right).

Lastly, we performed a dynamic light scattering (DLS) experiment on samples containing isolated SOR-LOC cages after incubation with and without sfGFP. We observed a relatively monodisperse signal peak at an approximately 7–8 nm radius for samples containing isolated SOR-LOC cages (Figure 2.3A). In contrast, SOR-LOC cages incubated with sfGFP yielded

essentially no signal at 7–8 nm radius, which supported results obtained by SEC and negative staining EM (Figure 2.3B).

Creation of a LOC Based on a Designed Protein Cage Core

We sought to extend the LOC concept using a designed protein assembly as the nanocage core. For the designed protein nanocage, we chose a tetrahedral assembly (known as T33-51) composed of two distinct subunit types, A and B, with stoichiometry A12B12 (4 trimers of each A and B), known for its solubility and stable assembly.[45] Furthermore, we wanted to test the modularity of our LOC design approach. Therefore, instead of using the sfGFP-binding DARPin employed above, we used a maltose binding protein (MBP)-binding DARPin as the binding adaptor to demonstrate that LOCs can respond to diverse protein targets.[59]

We generated a model of the T33-51-based LOC (T33-51-LOC) in the same manner as that described above. We aligned an ideal helix onto the C-terminus of the A subunit of the T33-51 cage. We then aligned the MBP-binding DARPin with the ideal helix. This resulted in a modified T33-51 A subunit with the MBP-binding DARPin fused via an Ala–Gln linker. As before, the DARPins were arranged such that the MBP-binding sites faced toward any axis of symmetry. To permit internal chemical labeling, we inserted a cysteine residue at the N-terminus of the A subunit. The B subunit of the T33-51 cage was left unchanged.

As modeled, T33-51-LOC comprises its original tetrahedral core assembly with 12 protruding DARPin domains (Figure 2.4A). Based on the positioning of the DARPin binding sites, individual MBP molecules could be bound without steric collision (Figure 2.4B, left and middle panels), but binding of two (or more) MBP molecules at a single vertex would lead to

steric collisions at the 3-fold axis (Figure 2.4B, right panel). We therefore hypothesized that the high-affinity binding of MBP to the surface DARPins would drive the disruption of T33-51-LOC via mass action.

We performed SEC and native polyacrylamide gel electrophoresis (native-PAGE) to characterize the T33-51-LOC assembly and MBP-binding. SEC of T33-51-LOC yielded a prominent peak at an elution volume of correct size, approximately 680 kDa, in the absence of the MBP ligand (Figure 2.4C). Addition of 75 μM MBP (3× molar excess of T33-51-LOC DARPins) caused a slight leftward shift of the assembled cage peak, indicating an increased molecular weight due to MBP binding (Figure 2.4C). However, when we performed native-PAGE on the SEC fractions corresponding to assembled cages and cages bound to MBP, we observed a significant decrease in the band intensity corresponding to assembled cages (Figure 2.4E). While minor assembled species remained (likely representing T33-51-LOC cages with only one MBP bound at any vertex), the majority were disrupted.

Additionally, we examined isolated T33-51-LOC cages incubated with and without MBP via negative stain EM. Cage samples incubated with only saline displayed monodisperse particles of correct size, approximately 10 nm, and apparent tetrahedral symmetry (Figure 2.4F, left). Upon adding MBP to T33-51-LOC, we observed an increase in polydispersity, a loss of symmetry, and fewer ordered particles (Figure 2.4F, right).

Finally, we performed a DLS experiment on isolated T33-51-LOC cages incubated with and without MBP. Cages without MBP (i.e., mixed with mock buffer) demonstrated that most of the sample ─ in terms of number and mass ─ exhibited a radius in the correct range of 6–7.5 nm (Figure 2.5A). Cages mixed with MBP exhibited significantly reduced numbers and masses of assembled T33-51-LOC, supporting results from native-PAGE and electron microscopy (Figure

2.5B). With these initial observations confirming our anticipated disruption of cages upon the addition of MBP, we expanded the system to incorporate alternative readouts of nanocage disruption and concomitant cargo exposure.

A LOC That Signals by Disruption-Based Fluorescence Unquenching

We designed a first LOC-based readout experiment using fluorescence as an indicator of disruption. To do this, we chemically labeled the interior cysteine of T33-51-LOC with Oregon Green 488 maleimide (labeled form referred to as og-T33-51-LOC). In the assembled state, multiple Oregon Green dye molecules exhibit a self-quenching property due to mutual proximity of the dyes (Figure 2.6A).[60] When the distance between the dye molecules increases, the quenching effect is attenuated, leading to an increase in the fluorescence intensity. Therefore, we expected to observe an increase in fluorescence intensity at 535 nm upon incubation of og-T33-51-LOC with increasing concentrations of MBP, as the high-affinity binding of MBP to the surface DARPins drives disassembly of the og-T33-51-LOC nanocage.

Indeed, we observed that the fluorescence intensity increased as og-T33-51-LOC was incubated with higher concentrations of MBP (Figure 2.6B). While all concentrations of MBP led to a gradual increase in fluorescence over time at room temperature, we observed larger differences in the presence of increasing amounts of MBP, as well as a prominent secondary increase in fluorescence signal at around 40 to 50 min, unique to conditions with bound ligand. Furthermore, at higher concentrations of MBP, this second signal spike occurred earlier. We were able to determine an approximate half-saturation threshold of binding in the low micromolar range (Figure 2.6C), taking the background signal from the unbound form as the baseline for subtraction. In addition to the main effect of ligand-driven disruption, the

experiments showed additional complex behaviors, which were reproducible. These are not understood at present, but likely reflect the complex multicomponent nature of the underlying system.

A LOC That Signals by Complementation-Based Luminescence

Finally, we sought to couple a LOC disruption event to a bioactive readout. To achieve this, we modified our initial T33-51-LOC design by replacing the internal cysteine residue with a split-NanoLuc luciferase complementation peptide. This peptide is a C-terminal fragment of an engineered split-NanoLuc luciferase enzyme, previously used to study protein–protein interactions via luminescence.[61] This modified version of our design is referred to as the nL-T33-51-LOC.

In this system, the assembled state of nL-T33-51-LOC does not sterically permit NanoLuc complementation because the peptide is encapsulated and not surface-exposed. We expected a cage opening event to expose the encapsulated peptide and permit complementation of split-NanoLuc, reconstituting luminescent enzyme activity (Figure 2.7A). That is, we expected an increase in luminescent light production after incubating nL-T33-51-LOC with higher concentrations of MBP, which is the cognate ligand.

We performed SEC to verify the assembly of nL-T33-51-LOC. A prominent peak at the correct elution volume (approximately 683 kDa assembly) was observed following the purification of nL-T33-51-LOC by SEC (Figure 2.7B). Upon incubation of nL-T33-51-LOC with increasing concentrations of MBP, we observed an increase in luminescence compared to samples incubated with a mock or noncognate ligand (lysozyme) (Figure 2.7C). Concentrations of the cognate target protein (MBP) in the low micromolar range were sufficient to generate a

significant increase in signal. In principle, other split enzyme systems would be amenable to use in signaling. And of course, other, nonenzymatic, biological motifs could be employed to signal upon exposure by binding to downstream targets.

CONCLUSIONS

Our studies describe key design principles and initial demonstrations for ligand-operable cages (LOCs): protein cages tailored to open or disassemble in response to highly specific molecular binding events. We demonstrated the scope and modularity of the design approach by using two different protein cage cores (one natural and one designed), two different protein ligands as the effector (sfGFP and MBP), and two different mechanisms for readout of cage opening (one based on fluorescence unquenching, and one based on split-enzyme complementation). Many other design schemes are possible for LOCs. For applications in broad areas of synthetic biology and molecular medicine, these systems provide a facile way to detect a specific protein of interest (e.g., using a spectroscopic readout) or to generate specific biological effects in response to a protein of interest (e.g., by exposure of inward-facing biological motifs). While the LOC framework opens broad new possibilities, it also exposes interesting areas for deeper exploration on both fundamental and practical fronts. Regarding practical applications, further developments are needed to investigate cell-based utility for LOCs.

Binding affinities and protein ligand concentrations will also be critical concerns, as these determine the driving forces for cage disassembly. Certain cell-surface applications could be well-suited given the high local concentrations of receptors in many natural systems. Practical applications will also have to consider relevant time scales. In our initial experiments, the time scales for significant cage disassembly were in the range of several minutes (Figure 2.6B). The

particular LOC examples we demonstrate here can be converted readily to respond to other protein ligands and to signal in different ways. But further protein engineering could be important in tuning LOCs to disassemble at different concentrations of a ligand. For instance, mutations that slightly destabilize the subunit interfaces in the cage core tend to lower the ligand concentration required for disassembly. Further design improvements and variations could be important in cases in which well-behaved disassembly is needed. In some applications, the protein aggregation observed upon cage disassembly could be compatible with the release of triggering ligands, while the production of soluble subcomponents upon disassembly could be important in other situations.

A number of theoretical issues are also worth exploring, particularly concerning the interplay of competing thermodynamic processes in LOCs (i.e., binding/unbinding vs assembly/disassembly). One notable observation is that the concentration of ligand required to disassemble an LOC is substantially higher than the KD for binding of the protein ligand to the adaptor protein in isolation. For the og-T33-51 LOC, half-disruption of the cage occurred for ligand (MBP) concentrations in the high nanomolar to low micromolar range, whereas the K_D of the DARPin for MBP in isolation is reported to be in the low nanomolar range.[59] This is consistent with expectation, considering the basic principles of competing forces. A deeper, perhaps even predictive, understanding of these systems might be possible with a more detailed modeling. In the presence of the protein ligand of interest, LOC modeling would involve complex ensembles of species (e.g., different patterns of binding site occupancy, different patterns of interface disruption and cage breaking, etc.). Questions could be addressed concerning whether these molecular systems might tend to behave cooperatively or anticooperatively as a function of ligand concentration; experiments tailored to detect populated

disassembly intermediates could be informative on that point. This could be important, given that cooperative behavior is often advantageous for response tuning and circuit design. There are also open questions concerning design. Our experiments showed success under particular design choices that we took to be favorable, a notable one being our use of DARPins as the binding adaptors to exploit their capacity for continuous alpha helical fusions at their termini. Our studies did not address whether viable LOC designs could be realized using more flexible modes of adaptor attachment, for example, if nanobodies were employed as the adaptors. Further work should shed light on these open questions.

MATERIALS AND METHODS

LOC Design and Modeling

Modeling and visualization of test LOCs were performed by using PyMOL and UCSF ChimeraX.

SOR-LOC: A Design Based on a Natural Protein Cage Core

First, the N-terminus of an ideal polyalanine helix was superimposed onto the C-terminal helix of an SOR subunit. Any residues that composed the actual C-terminus of the full-length SOR sequence but were not helically structured were ignored for the purposes of alignment and subsequently deleted to accommodate a fully helical extension. Then, the N-terminal helix of sfGFP-binding DARPin was superimposed onto the C-terminus of the same ideal helix. This SOR-helix-DARPin model "monomer" was then symmetry-expanded, allowing us to model the DARPin orientations on the surface of a full, 24-mer SOR-helix-DARPin assembly. Modeled DARPins were rotated by adjusting the superimposed length along the ideal helix, and all

DARPins on the 24-mer assembly were subjected to the same rotations. Bound sfGFP molecules were modeled by superimposing the structures of sfGFP-DARPin complexes (PDB code: 6NHV) with unbound SOR-LOC DARPins. SfGFP-bound DARPin rotations were performed in the same way until we obtained a configuration that fit the following criteria: (1) in the absence of bound sfGFP, no clashes were observed between the DARPins and/or the SOR subunits, and (2) in the sfGFP-bound form, clashes were observed between any sfGFP molecule, neighboring sfGFP molecule, neighboring DARPin, and/or cage surface. Once the final configuration was chosen, residues of the ideal helix that overlapped with terminal SOR or DARPin residues were discarded and any remaining ideal helix residues comprised the linker sequence. Linker sequences were then designed such that no additional clashes were observed between the linker residues and any part of the cage, avoiding prolines and glycines in the sequence. Our final model comprised the SOR fused to DARPin by an Ala–Gln linker. SOR-LOC was also tagged with a C-terminal polyhistidine motif for affinity purification.

T33-51-LOC: A Design Based on a Designed Protein Cage Core

First, the N-terminus of an ideal polyalanine helix was superimposed onto the C-terminal helix of T33-51-A. Any residues that composed the actual C-terminus of the full-length T33-51-A sequence but were not helically structured were ignored for the purposes of alignment and subsequently deleted to accommodate a fully helical extension. Then, the N-terminal helix of MBP-binding DARPin (PDB code: 1SVX) was superimposed onto the C-terminus of the same ideal helix. T33-51-A-DARPin and T33-51-B were then symmetry-expanded, allowing us to model the DARPin orientations on the surface of a full A12B12 T33-51-LOC assembly. Bound MBP molecules were modeled by aligning the structures of MBP-DARPin complexes (PDB

code: 1SVX) with unbound DARPins. MBP-bound DARPin rotations were performed as above until we obtained a configuration that fit the same criteria as those above for SOR-LOC. Once the final configuration was chosen, leftover residues from the superimposed ideal helix were discarded as above. Linker sequence was designed as above. Our final model comprised T33-51-A fused to the MBP-binding DARPin by an Ala–Gln linker. T33-51-A was also tagged with an N-terminal (internal) cysteine residue via a flexible polyglycine linker for chemical modification. For nL-T33-51-LOC, component A instead was tagged with an N-terminal NanoLuc complementation peptide (peptide "101"). (54) T33-51-B was tagged with a C-terminal polyhistidine motif for affinity purification.

Recombinant Protein Expression and Purification

All gene fragments encoding SOR-LOC, T33-51-LOC, nL-T33-51-LOC, sfGFP, MBP, and the split NanoLuc component (11S) were purchased from Integrated DNA Technologies. Plasmid sequences were verified via Sanger Sequencing (Azenta Life Sciences/Genewiz). All genes were inserted into a pET-22b+ vector (Novagen) via Gibson assembly with a NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). DNA manipulations and amplification were carried out in Escherichia coli XL2 cells (Agilent).

For T33-51-LOC and nL-T33-51-LOC, both A and B components were generated via polycistronic expression. All of the other constructs were expressed from a single cistron. Expression was performed using Escherichia coli BL21 (DE3) cells (New England Biolabs). Transformed cells were grown in Luria–Bertani (LB) Broth supplemented with 100 μg/mL ampicillin at 37 °C to an OD600 between 0.8 and 1.0, then 100 μM isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to induce protein expression. Following induction,

cultures were grown overnight at 18 °C before harvesting via centrifugation at 3500g for 15 min at 4 °C.

Cells expressing SOR-LOC were resuspended in 100 mM Tris at pH 8, 150 mM NaCl, 5 mM BME, and 20 mM imidazole. Cells expressing T33-51-LOC and nL-T33-51-LOC were resuspended in 50 mM Tris-HCl (pH, 8.0), 2% w/v glycerol, 5 mM TCEP-HCl, 500 mM NaCl, and 20 mM imidazole. Cells expressing MBP and 11S were resuspended in 50 mM Tris-HCl (pH, 8.0), 2% (w/v) glycerol, 5 mM TCEP-HCl, 250 mM NaCl, and 20 mM imidazole. Prior to lysis, resuspension buffers were supplemented with an EDTA-free protease inhibitor tablet (Pierce), 10 U/mL benzonase nuclease, and 0.1 mg/mL lysozyme. Resuspended cells were then pressure-homogenized using an Emulsiflex C3 Homogenizer (Avestin). Lysates were clarified via centrifugation at 20,000g at 4 °C for at least 30 min.

For SOR-LOC, clarified lysates were loaded onto Ni-NTA resin (Thermo Fisher Scientific) that was pre-equilibrated in resuspension buffer, then washed/eluted with purification buffers (100 mM Tris-HCl pH 8.0, 150 mM NaCl, further supplemented with 50, 75, 100, or 500 mM imidazole). SOR-LOC cages were then further purified via an NGC chromatography system (Bio-Rad) over a Superose 6 Increase 10/300 GL resin column (Cytiva). Sample runs were conducted in 50 mM Tris-HCl at pH 8.0, 150 mM NaCl, and 20 mM imidazole at 4 °C with a flow rate of 0.3 mL/min.

For MBP, 11S, T33-51-LOC, and nL-T33-51-LOC, clarified lysates were loaded onto Ni-NTA resin (Thermo Fisher Scientific) that was pre-equilibrated in respective resuspension buffers and then washed/eluted with purification buffers (50 mM Tris-HCl pH 8.0, 2% w/v glycerol, and 250 mM NaCl, further supplemented with 20, 50, 100, or 250 mM imidazole). T33-51-LOC samples were instead eluted at pH 7.5 and additionally supplemented with 5 mM

TCEP-HCl prior to chemical labeling (see below). NL-T33-51-LOC cage proteins were further purified via an NGC chromatography system (Bio-Rad) on a Superose 6 Increase 10/300 GL resin column (Cytiva). Sample runs were conducted in SEC buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2% w/v glycerol) at 4 °C with a flow rate of 0.3 mL/min.

Determination of Protein Cage Concentration

A single monomer of SOR-LOC was treated as a single molecule of cage protein for the purposes of determining the protein cage concentration.

T33-51-LOC and nL-T33-51-LOC components A and B were combined into a single continuous amino acid sequence before the molar mass and absorptivity. A unit comprising one copy of component A and one copy of component B was used as the basis for reporting protein cage concentrations.

Negative Stain Electron Microscopy

A 5 µL aliquot of purified nanocages from size exclusion chromatography samples (ranging from 0.01 to 0.05 mg/mL) was stained with 2% uranyl acetate on glow-discharged 300-mesh copper grids (Ted Pella, Inc.). Nanocages were imaged using an FEI Tecnai T12 transmission electron microscope at 120 kV. Samples mixed with cognate ligand were prepared in a molar ratio of 1:2 cage protein to ligand for SOR-LOC and sfGFP, and a ratio of 1:3 cage protein to ligand for T33-51-LOC and MBP.

SOR-LOC Disassembly Assay

A 25 µM aliquot of Ni-NTA affinity-purified SOR-LOC cage protein was incubated with 50 µM sfGFP or buffer mock for at least 5 min at 25 °C. Samples were then centrifuged at

10,000g for 5 min to eliminate aggregates before they were injected and analyzed via SEC as previously described for SOR-LOC cages.

T33-51-LOC MBP-Binding Assay

A 25 μM aliquot of Ni-NTA affinity-purified T33-51-LOC cage protein was incubated with 75 μM MBP or buffer mock for at least 1 h at 25 °C. Samples were then centrifuged at 10 000g for 5 min to eliminate aggregates before they were injected and analyzed via SEC as previously described for T33-51-LOC cages.

SDS-PAGE

Protein samples were mixed with SDS loading dye (0.004% Bromophenol blue, 6% glycerol, 2% sodium dodecyl sulfate, 50 mM Tris-HCl pH 6.8, 5 mM DTT) before boiling for at least 2 min. Samples were loaded onto a Mini-Protean TGX polyacrylamide gel (BioRad). Electrophoresis was performed at 200 V for 30 min. Gels were stained with Coomassie Blue and destained.

Native-PAGE

Assembled cage and ligand-bound cage fractions collected from SEC were mixed with native sample buffer (30 mM Tris-HCl, pH 6.8, 20% glycerol, 0.005% Bromophenol blue) before loading onto a Mini-Protean TGX polyacrylamide gel (BioRad). Electrophoresis was performed on ice at 170 V for 120 min. Gels were stained with Coomassie Blue and destained.

Dynamic Light Scattering

For SOR-LOC disassembly analysis, 100 nM isolated cages from SEC were mixed with elution buffer (mock) or 200 nM sfGFP and incubated at 25 °C for 5 min. For T33-51-LOC disassembly analysis, 100 nM isolated cages from SEC were mixed with elution buffer (mock) or MBP and incubated at 25 °C for 1 h. All reads were collected at 25 °C on a DynaPro Platereader-II (Wyatt Technology).

Fluorescence Unquenching Assay

Ni-NTA affinity-purified T33-51-LOC cages were labeled with Oregon green 488 maleimide using a commercial thiol-labeling protocol (Thermo Scientific). 10-fold molar excess of Oregon green 488 maleimide dye (predissolved in DMSO) was added (no more than 5% v/v) to 25–50 μM T33-51-LOC and incubated at RT in 50 mM Tris-HCl pH 7.5, 250 mM NaCl, 2% w/v glycerol, 5 mM TCEP-HCl, and 250 mM imidazole for at least 2 h. Oregon green 488-labeled T33-51-LOC (og-T33-51-LOC) cages were then purified over a Superose 6 Increase 10/300 GL resin column (Cytiva) in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, and 2% w/v glycerol.

A 10 nM amount of SEC-purified og-T33-51-LOC was added to an assay solution containing 250 mM NaCl, 5% w/v glycerol, 50 mM Tris pH 8.0, and varying concentrations of MBP (0–3 μM).

Fluorescence intensity readings were taken on a SpectraMax iD3 plate reader (Molecular Devices) at 20 °C for an hour.

Split-NanoLuc Complementation Assay

50 nM nL-T33-51-LOC was added to assay solutions containing 50% v/v SEC buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2% w/v glycerol) and 50% v/v Nanoglo assay buffer (Promega). Assay solutions also contained 20 pM split-NanoLuc component (11S), as well as 0–20 μM MBP or 20 μM lysozyme. Assay mixtures were incubated at room temperature for at least 30 min. Luminescence readings were taken on an Infinite M1000 PRO plate reader (TECAN).

FIGURES

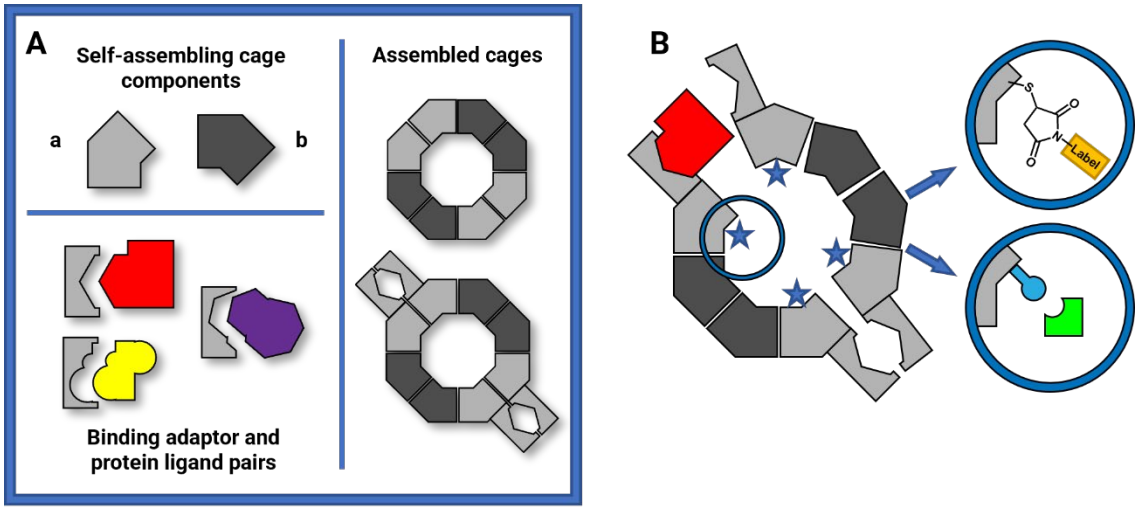

Figure 2.1. Concept of disrupting nanocages by ligand binding. A) Example self-assembling protein cage components "a" (gray) and "b" (dark gray) (top left panel); example pairs of binding adaptors and protein ligands (bottom left panel; adaptors represented in gray, cognate protein ligands in variable colors); self-assembled nanocages, without adaptors (right panel, top) and with fused adaptors (right panel, bottom). B) Bound adaptor–ligand complex is incompatible with the assembled state of the nanocage, driving disruption of the assembly; internalized cargo (star) can be used as a reporter of nanocage opening ─ e.g., based on chemical label (top inset) or on a new protein–protein interaction (bottom inset).

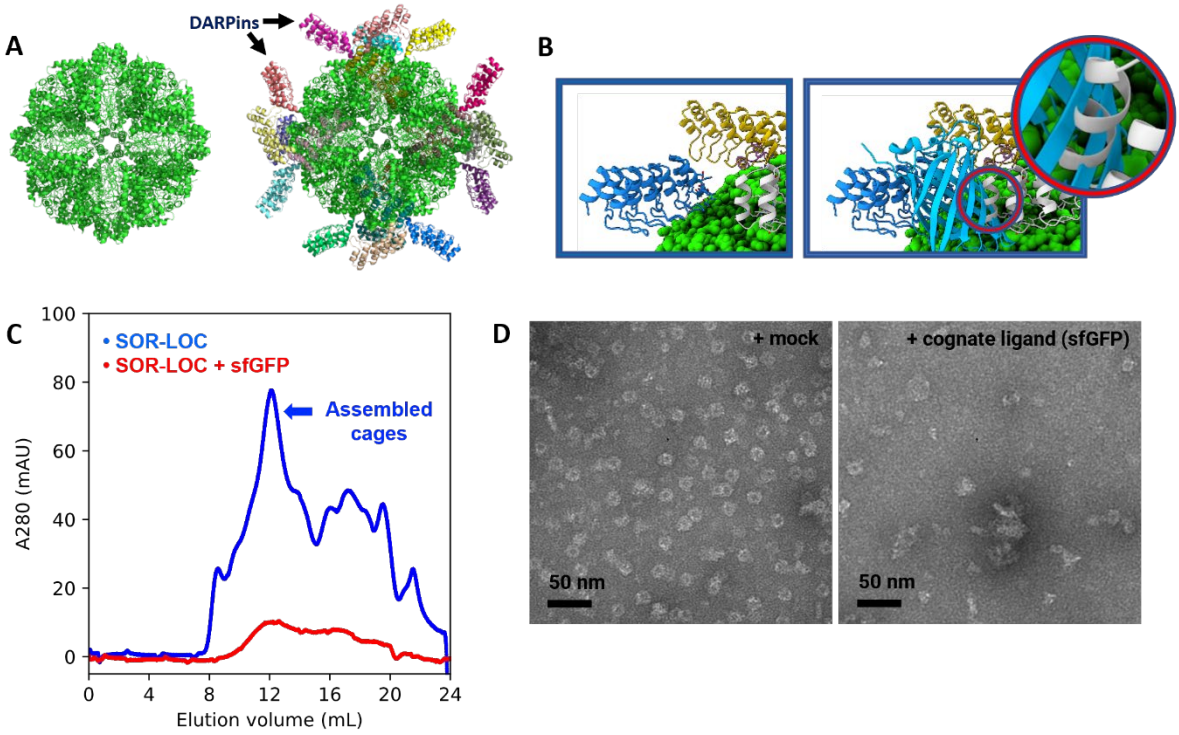

Figure 2.2. Design of a ligand-operable cage (LOC) based on sulfur oxygenase reductase (SOR). A) Biological assembly of sulfur oxygenase reductase (SOR) (PDB: 2YAV) (left); sulfur oxygenase reductase with fused sfGFP-binding DARPins (a.k.a., SOR-LOC; SOR in green, DARPins in various colors). B) Alternate view of select SOR-LOC DARPins (SOR surface in green, DARPins in blue, yellow, white): without cognate protein ligand (sfGFP), steric clashes are absent (left panel); in the sfGFP-bound form (sfGFP in light blue bound to DARPin in blue, right panel), steric clashes are observed between the bound sfGFP and neighboring DARPin (white); alternate, magnified view of clashes (inset). C) SEC analysis of SOR-LOC (blue trace) and SOR-LOC with sfGFP (red trace) following sample centrifugation; A280 (whole protein) for SOR-LOC alone shows a prominent peak at approximately 12 mL elution volume, indicating assembled cages; addition of sfGFP causes significant reduction in overall A280 signal following elimination of aggregates. D) Representative negative-stain electron micrographs of assembled SOR-LOC cage incubated with mock ligand solution (left) and incubated with sfGFP (right).

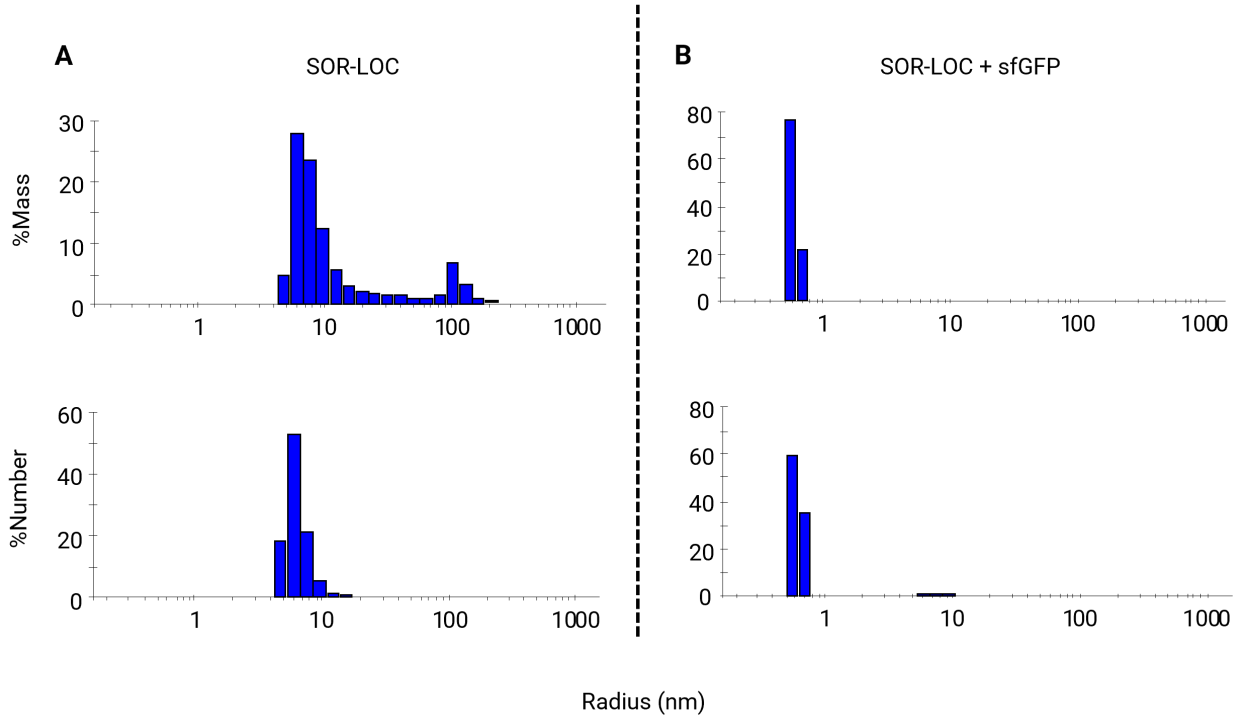

Figure 2.3. DLS analysis of the sfGFP-triggered SOR-LOC disassembly. A) Percent mass (top) and percent number (bottom) particle radius histograms for SOR-LOC. B) Percent mass (top) and percent number (bottom) particle radius histograms for SOR-LOC incubated with sfGFP.

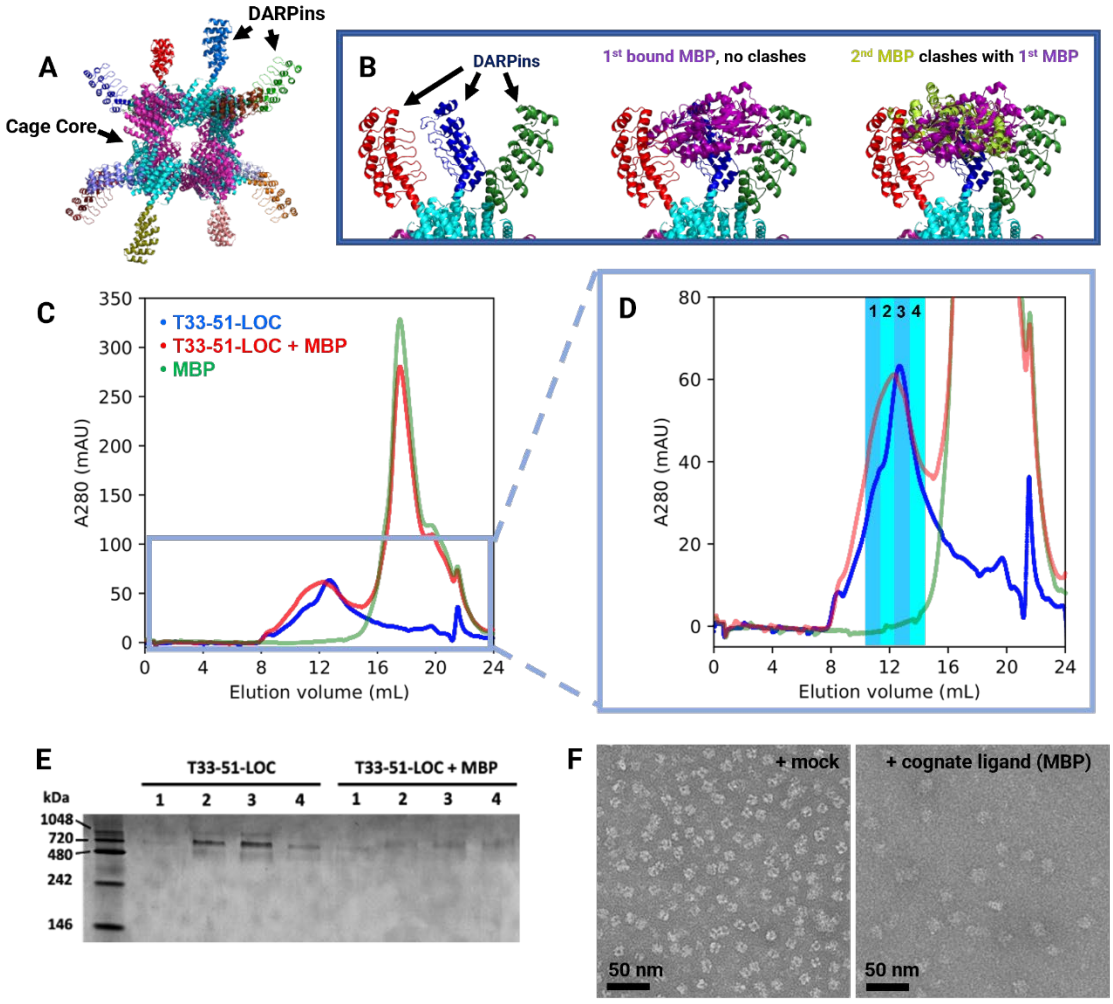

Figure 2.4. Design and synthesis of LOCs using designed nanocage T33-51. A) Design model of T33-51 nanocage core (magenta and cyan, PDB: 5CY5) with surface-fused MBP-binding DARPins (various colors) (whole assembly referred to as T33-51-LOC). B) Magnified view of select T33-51-LOC DARPins: in the absence of cognate protein ligand (MBP), no steric clashes are observed (left); single MBP (purple) binding to DARPin (red) is permitted without steric clashes (middle); second MBP (light green) binding to DARPin (dark green) is sterically incompatible with cage assembly (right). C) SEC chromatogram of T33-51-LOC (blue trace), T33-51-LOC with MBP (red trace), and MBP (green trace); A280 (whole protein) for T33-51-LOC alone shows prominent peak at approximately 12.5 mL elution volume, indicating

assembled cages; addition of MBP causes a leftward shift of assembled cage peak. D) Alternate view of assembled T33-51-LOC cage peak and associated MBP-bound shift. E) Native-PAGE analysis of associated T33-51-LOC and T33-51-LOC + MBP fractions from (D). F) Representative negative-stain electron micrographs of assembled T33-51-LOC cage incubated with mock (left) and incubated with MBP (right).

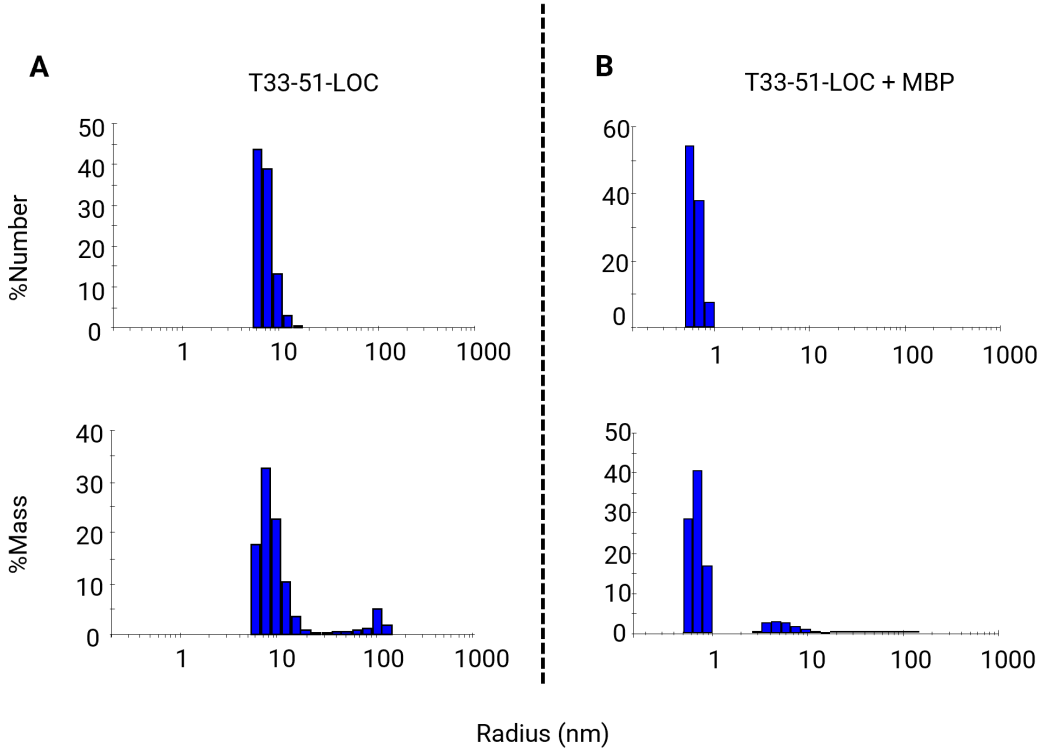

Figure 2.5. DLS analysis of MBP-triggered T33-51-LOC disassembly. A) Percent number (top) and percent mass (bottom) particle radius histograms for T33-51-LOC. B) Percent number (top) and percent mass (bottom) particle radius histograms for T33-51-LOC incubated with MBP.

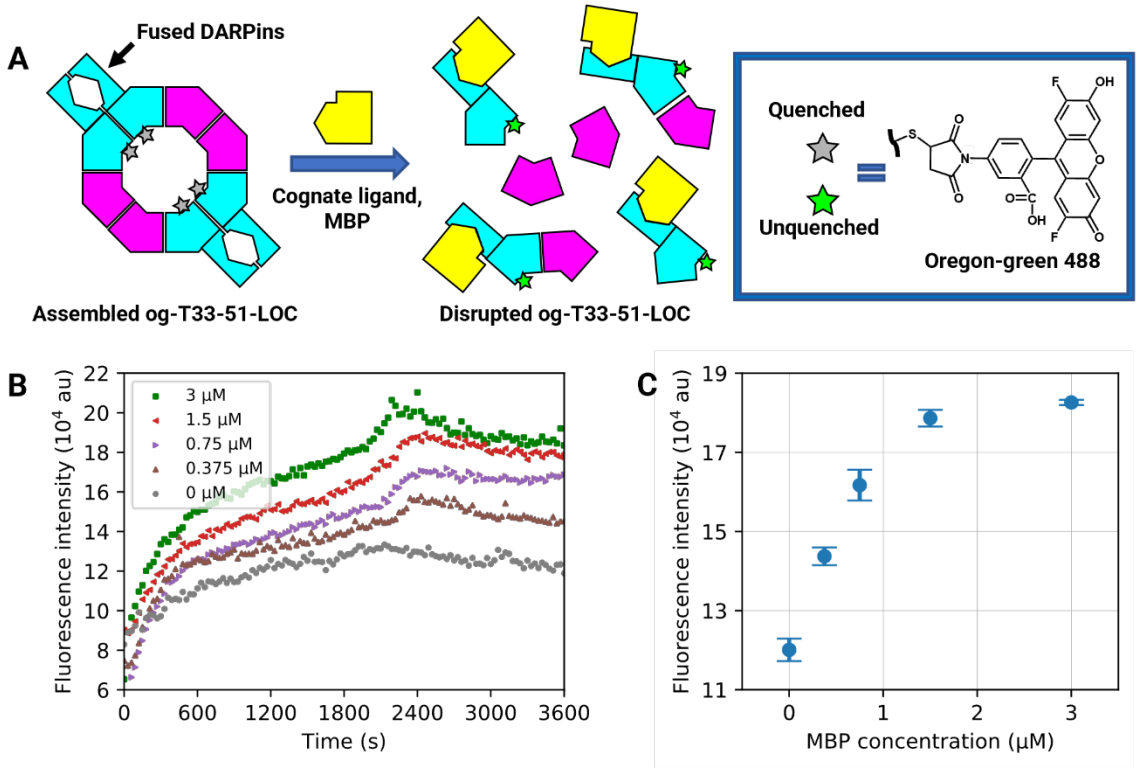

Figure 2.6. Fluorescent reporter readout of the T33-51-LOC disruption. A) Cartoon representation of readout experiment: T33-51-LOC cages (cyan and magenta) with internally conjugated Oregon green 488 dye molecules (stars) exhibit attenuated fluorescence due to self-quenching (left); following addition of cognate ligand, MBP (yellow), nanocages are driven to disassemble, thereby unquenching the fluorophores. B) Plot of fluorescence intensity versus time of Oregon green 488-labeled T33-51-LOC (og-T33-51-LOC): incubation of og-T33-51-LOC with increasing concentrations of MBP leads to increased fluorescence over time. C) Plot of fluorescence intensity of og-T33-51-LOC at signal equilibrium from (B) versus MBP concentration; half-saturation threshold (using unbound signal as baseline) approximately marked by dashed lines.

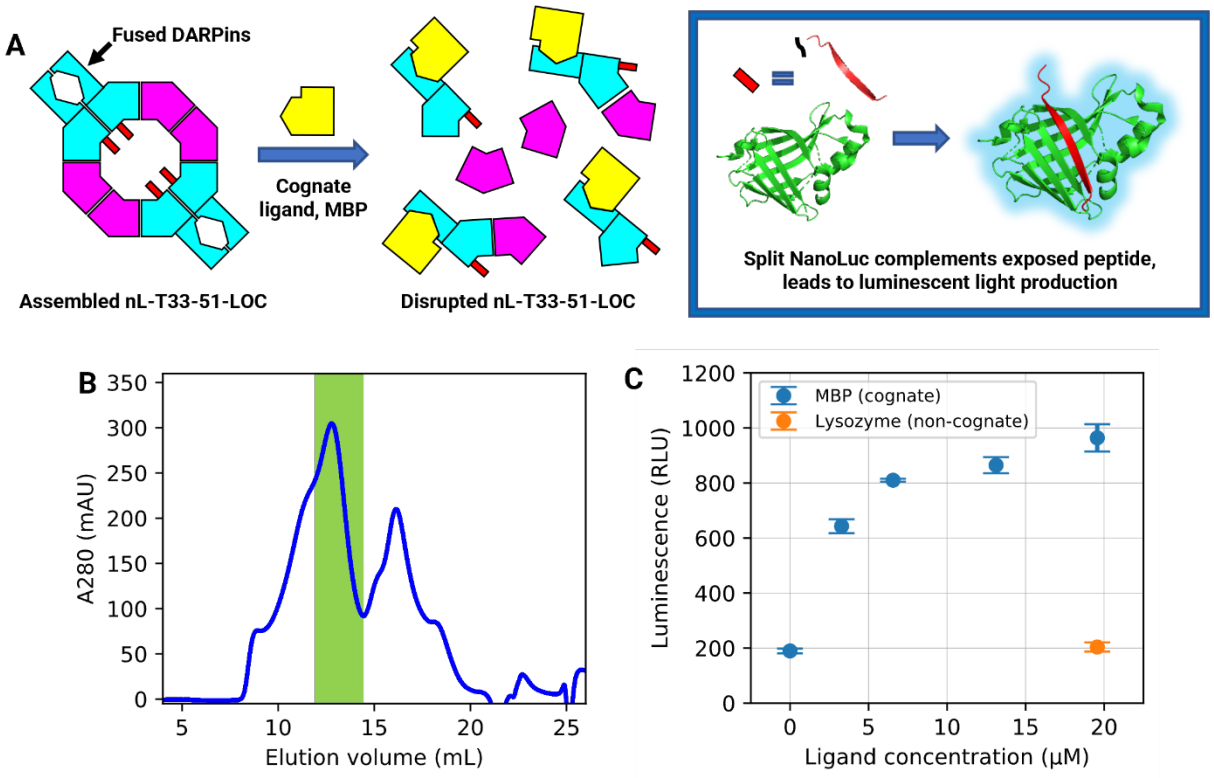

Figure 2.7. Bioluminescent reporter readout of the T33-51-LOC disruption. A) Cartoon representation of the readout experiment: T33-51-LOC with internally fused split-NanoLuc luciferase peptide (nL-T33-51-LOC) (left, cage subunits in cyan and magenta, NanoLuc peptide in red); following addition of cognate ligand, MBP (yellow), nanocages are driven to disassemble, thereby exposing the split-NanoLuc peptide; exposed peptide complements split-NanoLuc component 11S (green), leading to luminescent light production (right inset). B) SEC chromatogram of nL-T33-51-LOC assembled in the absence of MBP; assembled cage fraction highlighted in green. C) Plot of RLU vs ligand concentration following incubation of nL-T33-51-LOC with noncognate (lysozyme) or cognate protein (MBP), in the presence of free 11S and furimazine substrate in solution.

Sequences

Protein	**Amino acid sequence**
SOR-LOC	MPKPYVAINMAELKNEPKTFEMFASVGPKVCMVTARHPGFVGF QNHIQIGILPFGNRYGGAKMDMTKESSTVRVLQYTFWKDWKDH EEMHRQNWSYLFRLCYSCASQMIWGPWEPIYEIIYANMPINTE MTDFTAVVGKKFAEGKPLDIPVISQPYGKRVVAFAEHSVIPGKE KQFEDAIVRTLEMLKKAPGFLGAMVLKEIGVSGIGSMQFGAKGF HQVLENPGSLEPDPNNVMYSVPEAKNTPQQYIVHVEWANTDAL MFGMGRVLLYPELRQVHDEVLDTLVYGPYIRILNPMMEGTFWR EYLNEAA**AQ***GKKLLEAARAGQDDEVRILMANGADVNAADDVGV TPLHLAAQRGHLEIVEVLLKCGADVNAADLWGQTPLHLAATAG HLEIVEVLLKNGADVNARDNIGHTPLHLAAWAGHLEIVEVLLKYG ADVNAQDKFGKTPFDLAIDNGNEDIAEVLQKAAHHHHHH*
T33-51-LOC-A	MCGGGRITTKVGDKGSTRLFGGEEVWKDDPIIEANGTLDELTSF IGEAKHYVDEEMKGILEEIQNDIYKIMGEIGSKGKIEGISEERIKWL AGLIERYSEMNKLSFVLPGGTLESAKLDVCRTIARRAERKVATVL REFGIGTLAAIYLALLSRLLFLLARVIEIEKNKL**AQ***GRKLLEAARA GQDDEVRILMANGADVNAADNTGTTPLHLAAYSGHLEIVEVLLK HGADVDASDVFGYTPLHLAAYWGHLEIVEVLLKNGADVNAMDS DGMTPLHLAAKWGYLEIVEVLLKHGADVNAQDKFGKTAFDISID NGNEDLAEILQKLN*
nL-T33-51-LOC-A	MVTGYRLFEKESGSGSTRLFGGEEVWKDDPIIEANGTLDELTSF IGEAKHYVDEEMKGILEEIQNDIYKIMGEIGSKGKIEGISEERIKWL AGLIERYSEMVNKLSFVLPGGTLESAKLDVCRTIARRAERKVAT VLREFGIGTLAAIYLALLSRLLFLLARVIEIEKNKL**AQ***GRKLLEAAR AGQDDEVRILMANGADVNAADNTGTTPLHLAAYSGHLEIVEVLL KHGADVDASDVFGYTPLHLAAYWGHLEIVEVLLKNGADVNAMD SDGMTPLHLAAKWGYLEIVEVLLKHGADVNAQDKFGKTAFDISI DNGNEDLAEILQKLN*
T33-51-LOC-B (same as nL-T33-51-LOC-B)	MFTRRGDQGETDLANRARVGKDSPVVEVQGTIDELNSFIGYAL VLSRWDDIRNDLFRIQNDLFVLGEDVSTGGKGRTVTMDMIIYLIK RSVEMKAEIGKIELFVVPGGSVESASLHMARAVSRRLERRIKAA SELTEINANVLLYANMLSNILFMHALISNKRLNIPEKIWSIHRVSLE HHHHHH
Split NanoLuc component (11S)	MHHHHHHGSGVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLL QNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVF KVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAVFD GKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINS
SfGFP	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLT LKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSA MPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFK EDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGS VQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHM VLLEFVTAAGITHHHHHH
His-MBP	MHHHHHHLVPRGSGSGSMKTEEGKLVIWINGDKGYNGLAEVG KKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFG GYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEAL SLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFT WPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKN

	KHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGV TVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDE GLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPN IPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTGSGGTPGRP AAKLN

Table 2.S1. Amino acid sequences of all LOC cages, cognate protein ligands, and assay components. Underlined amino acids indicate internally fused residues. Underlined and bolded amino acids indicate helical linker. Underlined and italicized amino acids indicate DARPin.

Protein	Checked	Protein FDR Confidence: Sequest HT	Exp. q-value: Sequest HT	Coverage [%]	# Peptides	# Unique Peptides	# AAs	MW [kDa]	Score Sequest HT: Sequest HT	Found in Sample
SOR-LOC	TRUE	High	0	90	179	174	473	52.7	429.1	High
T33-51-LOC-A	TRUE	High	0	70	42	38	327	35.7	133.42	High
T33-51-LOC-B	TRUE	High	0	95	35	35	185	21	97.74	High

Table 2.S2. Mass spectrometry verification of designed cages.

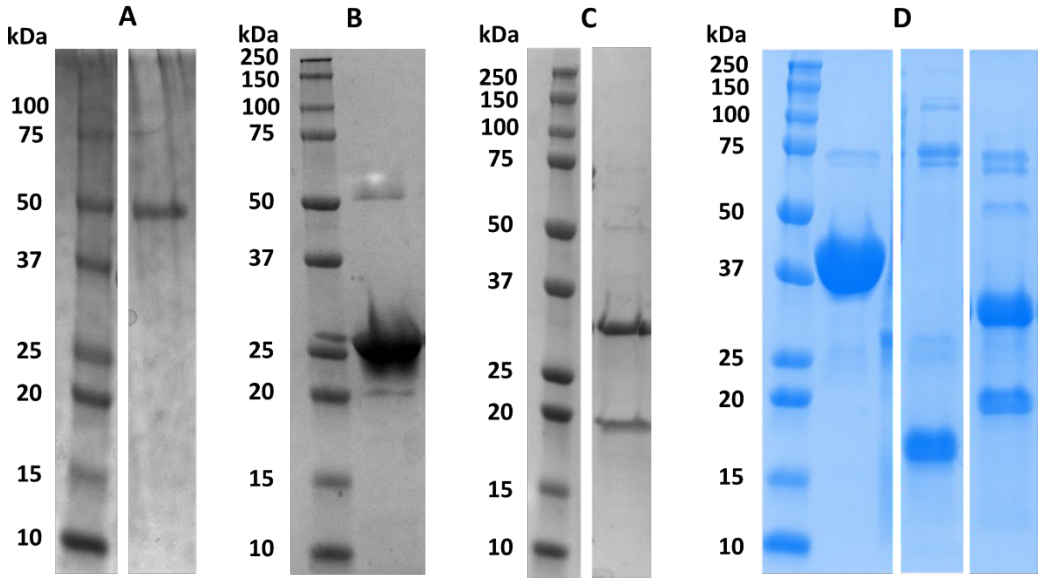

Figure 2.S1. SDS-PAGE analysis of purified LOC components following Ni-NTA gravity purification. A) SOR-LOC. B) sfGFP. C) T33-51-LOC-A (top band), T33-51-LOC-B (bottom

band). D) From left to right: His-MBP; His-11S; nL-T33-51-LOC-A (top), T33-51-LOC-B (bottom).

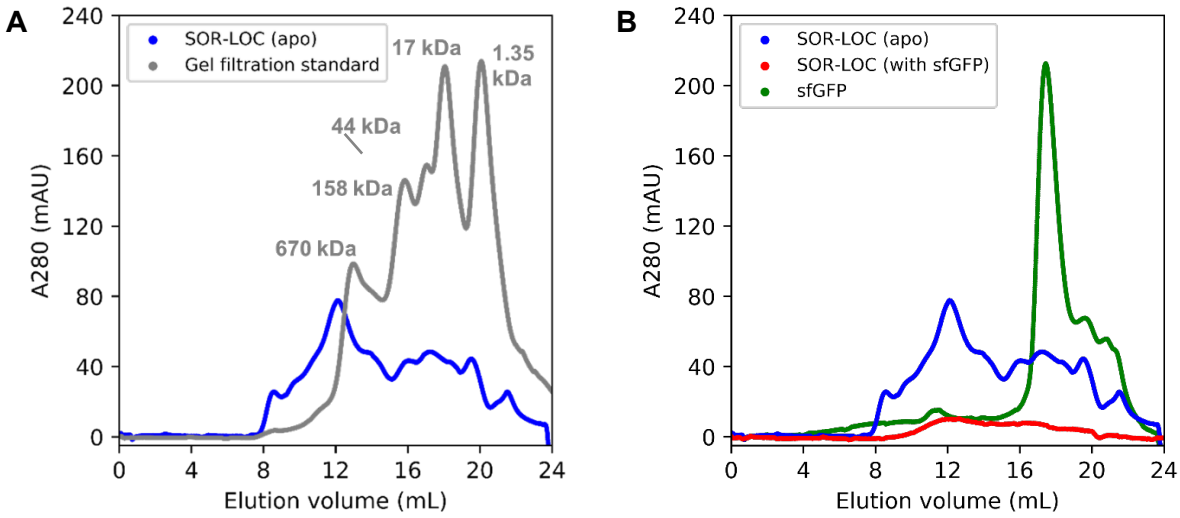

Figure 2.S2. Size exclusion chromatography of SOR-LOC and sfGFP-induced disassembly. A) SOR-LOC elution profile (blue trace) superimposed with molecular weight standard (grey trace). B) SOR-LOC elution profile (blue trace) superimposed with SOR-LOC with sfGFP (red trace), as well as free sfGFP monomers (green trace).

REFERENCES

(1) Keefe, A. D.; Szostak, J. W. Functional Proteins from a Random-Sequence Library. Nature 2001, 410 (6829), 715–718. https://doi.org/10.1038/35070613.

(2) Joh, N. H.; Wang, T.; Bhate, M. P.; Acharya, R.; Wu, Y.; Grabe, M.; Hong, M.; Grigoryan, G.; DeGrado, W. F. De Novo Design of a Transmembrane Zn 2+ -Transporting Four-Helix Bundle.

Science 2014, 346 (6216), 1520–1524. https://doi.org/10.1126/science.1261172.

(3) Huang, P.-S.; Feldmeier, K.; Parmeggiani, F.; Fernandez Velasco, D. A.; Höcker, B.; Baker, D. De Novo Design of a Four-Fold Symmetric TIM-Barrel Protein with Atomic-Level Accuracy. Nat. Chem. Biol. 2016, 12 (1), 29–34. https://doi.org/10.1038/nchembio.1966.

(4) Boyken, S. E.; Chen, Z.; Groves, B.; Langan, R. A.; Oberdorfer, G.; Ford, A.; Gilmore, J. M.; Xu, C.; DiMaio, F.; Pereira, J. H.; Sankaran, B.; Seelig, G.; Zwart, P. H.; Baker, D. De Novo Design of Protein Homo-Oligomers with Modular Hydrogen-Bond Network–Mediated Specificity.

Science 2016, 352 (6286), 680–687. https://doi.org/10.1126/science.aad8865.

(5) Kries, H.; Blomberg, R.; Hilvert, D. De Novo Enzymes by Computational Design. Curr. Opin. Chem. Biol. 2013, 17 (2), 221–228. https://doi.org/10.1016/j.cbpa.2013.02.012.

(6) 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. 2001, 98 (5), 2217–2221. https://doi.org/10.1073/pnas.041614998.

(7) 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. https://doi.org/10.1038/nature13404.

(8) Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C.; Cascio, D.; Yeates, T. O.; Gonen, T.; King, N. P.; Baker, D. Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes. Science 2016, 353 (6297), 389–394. https://doi.org/10.1126/science.aaf8818.

(9) Yeates, T. O. Geometric Principles for Designing Highly Symmetric Self-Assembling Protein Nanomaterials. Annu. Rev. Biophys. 2017, 46, 23–42. https://doi.org/10.1146/annurev-biophys-070816-033928.

(10) Laniado, J.; Yeates, T. O. A Complete Rule Set for Designing Symmetry Combination Materials from Protein Molecules. Proc. Natl. Acad. Sci. 2020, 117 (50), 31817–31823. https://doi.org/10.1073/pnas.2015183117.

(11) Laniado, J.; Cannon, K. A.; Miller, J. E.; Sawaya, M. R.; McNamara, D. E.; Yeates, T. O. Geometric Lessons and Design Strategies for Nanoscale Protein Cages. ACS Nano 2021, 15 (3), 4277–4286. https://doi.org/10.1021/acsnano.0c07167.

(12) Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N. Self-Assembling Cages from Coiled-Coil Peptide Modules. Science 2013, 340 (6132), 595–599. https://doi.org/10.1126/science.1233936.

(13) Ni, T. W.; Tezcan, F. A. Structural Characterization of a Microperoxidase inside a Metal-Directed Protein Cage. Angew. Chem. Int. Ed Engl. 2010, 49 (39), 7014–7018. https://doi.org/10.1002/anie.201001487.

(14) Sciore, A.; Su, M.; Koldewey, P.; Eschweiler, J. D.; Diffley, K. A.; Linhares, B. M.; Ruotolo, B. T.; Bardwell, J. C. A.; Skiniotis, G.; Marsh, E. N. G. Flexible, Symmetry-Directed

Approach to Assembling Protein Cages. Proc. Natl. Acad. Sci. U. S. A. 2016, 113 (31), 8681–8686. https://doi.org/10.1073/pnas.1606013113.

(15) Lai, Y.-T.; Reading, E.; Hura, G. L.; Tsai, K.-L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O. Structure of a Designed Protein Cage That Self-Assembles into a Highly Porous Cube. Nat. Chem. 2014, 6 (12), 1065–1071. https://doi.org/10.1038/nchem.2107.

(16) Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C.; Huang, P.-S.; Ravichandran, R.; Yi, S.; Davis, T. N.; Gonen, T.; King, N. P.; Baker, D. Design of a Hyperstable 60-Subunit Protein Dodecahedron. [Corrected]. Nature 2016, 535 (7610), 136–139. https://doi.org/10.1038/nature18010.

(17) Li, H.; DeRosier, D. J.; Nicholson, W. V.; Nogales, E.; Downing, K. H. Microtubule Structure at 8 Å Resolution. Structure 2002, 10 (10), 1317–1328. https://doi.org/10.1016/S0969-2126(02)00827-4.

(18) Caspar, D. L. D. Structure of Bushy Stunt Virus. Nature 1956, 177 (4506), 475–476. https://doi.org/10.1038/177475a0.

(19) Crick, F. H. C.; Watson, J. D. Structure of Small Viruses. Nature 1956, 177 (4506), 473–475. https://doi.org/10.1038/177473a0.

(20) Sevvana, M.; Klose, T.; Rossmann, M. G. Principles of Virus Structure. In Encyclopedia of Virology; Elsevier, 2021; pp 257–277. https://doi.org/10.1016/B978-0-12-814515-9.00033-3.

(21) Sára, M.; Sleytr, U. B. S-Layer Proteins. J. Bacteriol. 2000, 182 (4), 859–868. https://doi.org/10.1128/JB.182.4.859-868.2000.

(22) Cannon, K. A.; Ochoa, J. M.; Yeates, T. O. High-Symmetry Protein Assemblies: Patterns and Emerging Applications. Curr. Opin. Struct. Biol. 2019, 55, 77–84.

https://doi.org/10.1016/j.sbi.2019.03.008.

(23) Marcandalli, J.; Fiala, B.; Ols, S.; Perotti, M.; de van der Schueren, W.; Snijder, J.; Hodge, E.; Benhaim, M.; Ravichandran, R.; Carter, L.; Sheffler, W.; Brunner, L.; Lawrenz, M.; Dubois, P.; Lanzavecchia, A.; Sallusto, F.; Lee, K. K.; Veesler, D.; Correnti, C. E.; Stewart, L. J.; Baker, D.; Loré, K.; Perez, L.; King, N. P. Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. Cell 2019, 176 (6), 1420-1431.e17. https://doi.org/10.1016/j.cell.2019.01.046.

(24) McConnell, S. A.; Cannon, K. A.; Morgan, C.; McAllister, R.; Amer, B. R.; Clubb, R. T.; Yeates, T. O. Designed Protein Cages as Scaffolds for Building Multienzyme Materials. ACS Synth. Biol. 2020, 9 (2), 381–391. https://doi.org/10.1021/acssynbio.9b00407.

(25) Liu, Y.; Gonen, S.; Gonen, T.; Yeates, T. O. Near-Atomic Cryo-EM Imaging of a Small Protein Displayed on a Designed Scaffolding System. Proc. Natl. Acad. Sci. 2018, 115 (13), 3362–3367. https://doi.org/10.1073/pnas.1718825115.

(26) Liu, Y.; Huynh, D. T.; Yeates, T. O. A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold. Nat. Commun. 2019, 10 (1), 1864. https://doi.org/10.1038/s41467-019-09836-0.

(27) 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. https://doi.org/10.1021/acschembio.1c00007.

(28) Hills, R. A.; Tan, T. K.; Cohen, A. A.; Keeffe, J. R.; Keeble, A. H.; Gnanapragasam, P. N. P.; Storm, K. N.; Hill, M. L.; Liu, S.; Gilbert-Jaramillo, J.; Afzal, M.; Napier, A.; James, W. S.; Bjorkman, P. J.; Townsend, A. R.; Howarth, M. Multiviral Quartet Nanocages Elicit Broad Anti-Coronavirus Responses for Proactive Vaccinology. BioRxiv Prepr. Serv. Biol. 2023,

2023.02.24.529520. https://doi.org/10.1101/2023.02.24.529520.

(29) Lilavivat, S.; Sardar, D.; Jana, S.; Thomas, G. C.; Woycechowsky, K. J. In Vivo Encapsulation of Nucleic Acids Using an Engineered Nonviral Protein Capsid. J. Am. Chem. Soc. 2012, 134 (32), 13152–13155. https://doi.org/10.1021/ja302743g.

(30) Zhang, T.; Lv, C.; Chen, L.; Bai, G.; Zhao, G.; Xu, C. Encapsulation of Anthocyanin Molecules within a Ferritin Nanocage Increases Their Stability and Cell Uptake Efficiency. Food Res. Int. 2014, 62, 183–192. https://doi.org/10.1016/j.foodres.2014.02.041.

(31) Schoonen, L.; Nolte, R. J. M.; van Hest, J. C. M. Highly Efficient Enzyme Encapsulation in a Protein Nanocage: Towards Enzyme Catalysis in a Cellular Nanocompartment Mimic. Nanoscale 2016, 8 (30), 14467–14472. https://doi.org/10.1039/C6NR04181G.

(32) Yang, Z.; Wang, X.; Diao, H.; Zhang, J.; Li, H.; Sun, H.; Guo, Z. Encapsulation of Platinum Anticancer Drugs by Apoferritin. Chem. Commun. 2007, No. 33, 3453. https://doi.org/10.1039/b705326f.

(33) Dalmau, M.; Lim, S.; Wang, S.-W. PH-Triggered Disassembly in a Caged Protein Complex. Biomacromolecules 2009, 10 (12), 3199–3206. https://doi.org/10.1021/bm900674v.

(34) Ren, D.; Kratz, F.; Wang, S.-W. Protein Nanocapsules Containing Doxorubicin as a PH-Responsive Delivery System. Small 2011, 7 (8), 1051–1060. https://doi.org/10.1002/smll.201002242.

(35) Peng, T.; Lim, S. Trimer-Based Design of PH-Responsive Protein Cage Results in Soluble Disassembled Structures. Biomacromolecules 2011, 12 (9), 3131–3138. https://doi.org/10.1021/bm2005438.

(36) Chen, H.; Zhang, S.; Xu, C.; Zhao, G. Engineering Protein Interfaces Yields Ferritin

Disassembly and Reassembly under Benign Experimental Conditions. Chem. Commun. 2016, 52 (46), 7402–7405. https://doi.org/10.1039/C6CC03108K.

(37) Le Vay, K.; Carter, B. M.; Watkins, D. W.; Dora Tang, T.-Y.; Ting, V. P.; Cölfen, H.; Rambo, R. P.; Smith, A. J.; Ross Anderson, J. L.; Perriman, A. W. Controlling Protein Nanocage Assembly with Hydrostatic Pressure. J. Am. Chem. Soc. 2020, 142 (49), 20640–20650. https://doi.org/10.1021/jacs.0c07285.

(38) 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.; Wu, D.; Wrobel, T. P.; Fineberg, A.; Kushwah, M. S.; Kelemen, M.; Vavpetič, P.; Pelicon, P.; Kukura, P.; Benesch, J. L. P.; Iwasaki, K.; Heddle, J. G. An Ultra-Stable Gold-Coordinated Protein Cage Displaying Reversible Assembly. Nature 2019, 569 (7756), 438–442. https://doi.org/10.1038/s41586-019-1185-4.

(39) Xu, M.; Zeng, R.; Xiang, J.; Yan, Q. Self-Assembly of Switchable Protein Nanocages via Allosteric Effect. CCS Chem. 2021, 3 (8), 2223–2232. https://doi.org/10.31635/ccschem.020.202000437.

(40) Zang, J.; Chen, H.; Zhang, X.; Zhang, C.; Guo, J.; Du, M.; Zhao, G. Disulfide-Mediated Conversion of 8-Mer Bowl-like Protein Architecture into Three Different Nanocages. Nat. Commun. 2019, 10 (1), 778. https://doi.org/10.1038/s41467-019-08788-9.

(41) Hoersch, D.; Roh, S.-H.; Chiu, W.; Kortemme, T. Reprogramming an ATP-Driven Protein Machine into a Light-Gated Nanocage. Nat. Nanotechnol. 2013, 8 (12), 928–932. https://doi.org/10.1038/nnano.2013.242.

(42) Miller, J. E.; Srinivasan, Y.; Dharmaraj, N. P.; Liu, A.; Nguyen, P. L.; Taylor, S. D.; Yeates, T. O. Designing Protease-Triggered Protein Cages. J. Am. Chem. Soc. 2022, 144 (28),

12681–12689. https://doi.org/10.1021/jacs.2c02165.

(43) Kang, Y. J.; Park, D. C.; Shin, H.-H.; Park, J.; Kang, S. Incorporation of Thrombin Cleavage Peptide into a Protein Cage for Constructing a Protease-Responsive Multifunctional Delivery Nanoplatform. Biomacromolecules 2012, 13 (12), 4057–4064. https://doi.org/10.1021/bm301339s.

(44) Veith, A.; Urich, T.; Seyfarth, K.; Protze, J.; Frazão, C.; Kletzin, A. Substrate Pathways and Mechanisms of Inhibition in the Sulfur Oxygenase Reductase of Acidianus Ambivalens. Front. Microbiol. 2011, 2. https://doi.org/10.3389/fmicb.2011.00037.

(45) Cannon, K. A.; Park, R. U.; Boyken, S. E.; Nattermann, U.; Yi, S.; Baker, D.; King, N. P.; Yeates, T. O. Design and Structure of Two New Protein Cages Illustrate Successes and Ongoing Challenges in Protein Engineering. Protein Sci. 2020, 29 (4), 919–929. https://doi.org/10.1002/pro.3802.

(46) Blois, T. M.; Hong, H.; Kim, T. H.; Bowie, J. U. Protein Unfolding with a Steric Trap. J. Am. Chem. Soc. 2009, 131 (39), 13914–13915. https://doi.org/10.1021/ja905725n.

(47) Douglas, S. M.; Bachelet, I.; Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. Science 2012, 335 (6070), 831–834. https://doi.org/10.1126/science.1214081.

(48) 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.; Mulligan, V. K.; Dueber, J. E.; Novak, W. R. P.; El-Samad, H.; Baker, D. De Novo Design of Bioactive Protein Switches. Nature 2019, 572 (7768), 205–210. https://doi.org/10.1038/s41586-019-1432-8.

(49) Steiner, D.; Forrer, P.; Plückthun, A. Efficient Selection of DARPins with Sub-Nanomolar Affinities Using SRP Phage Display. J. Mol. Biol. 2008, 382 (5), 1211–1227.

https://doi.org/10.1016/j.jmb.2008.07.085.

(50) Seeger, M. A.; Zbinden, R.; Flütsch, A.; Gutte, P. G. M.; Engeler, S.; Roschitzki-Voser, H.; Grütter, M. G. Design, Construction, and Characterization of a Second-Generation DARPin Library with Reduced Hydrophobicity: Second-Generation DARPin Library. Protein Sci. 2013, 22 (9), 1239–1257. https://doi.org/10.1002/pro.2312.

(51) Luo, R.; Liu, H.; Cheng, Z. Protein Scaffolds: Antibody Alternatives for Cancer Diagnosis and Therapy. RSC Chem. Biol. 2022, 3 (7), 830–847. https://doi.org/10.1039/D2CB00094F.

(52) Wu, Y.; Batyuk, A.; Honegger, A.; Brandl, F.; Mittl, P. R. E.; Plückthun, A. Rigidly Connected Multispecific Artificial Binders with Adjustable Geometries. Sci. Rep. 2017, 7 (1), 11217. https://doi.org/10.1038/s41598-017-11472-x.

(53) Yao, Q.; Weaver, S. J.; Mock, J.-Y.; Jensen, G. J. Fusion of DARPin to Aldolase Enables Visualization of Small Protein by Cryo-EM. Structure 2019, 27 (7), 1148-1155.e3. https://doi.org/10.1016/j.str.2019.04.003.

(54) Kwon, N.-Y.; Kim, Y.; Lee, J.-O. The Application of Helix Fusion Methods in Structural Biology. Curr. Opin. Struct. Biol. 2020, 60, 110–116. https://doi.org/10.1016/j.sbi.2019.12.007.

(55) Mittl, P. R.; Ernst, P.; Plückthun, A. Chaperone-Assisted Structure Elucidation with DARPins. Curr. Opin. Struct. Biol. 2020, 60, 93–100. https://doi.org/10.1016/j.sbi.2019.12.009.

(56) Vulovic, I.; Yao, Q.; Park, Y.-J.; Courbet, A.; Norris, A.; Busch, F.; Sahasrabuddhe, A.; Merten, H.; Sahtoe, D. D.; Ueda, G.; Fallas, J. A.; Weaver, S. J.; Hsia, Y.; Langan, R. A.; Plückthun, A.; Wysocki, V. H.; Veesler, D.; Jensen, G. J.; Baker, D. Generation of Ordered Protein Assemblies Using Rigid Three-Body Fusion. Proc. Natl. Acad. Sci. 2021, 118 (23), e2015037118. https://doi.org/10.1073/pnas.2015037118.

(57) Kramer, M. A.; Wetzel, S. K.; Plückthun, A.; Mittl, P. R. E.; Grütter, M. G. Structural

Determinants for Improved Stability of Designed Ankyrin Repeat Proteins with a Redesigned C-Capping Module. J. Mol. Biol. 2010, 404 (3), 381–391. https://doi.org/10.1016/j.jmb.2010.09.023.

(58) Brauchle, M.; Hansen, S.; Caussinus, E.; Lenard, A.; Ochoa-Espinosa, A.; Scholz, O.; Sprecher, S. G.; Plückthun, A.; Affolter, M. Protein Interference Applications in Cellular and Developmental Biology Using DARPins That Recognize GFP and MCherry. Biol. Open 2014, 3 (12), 1252–1261. https://doi.org/10.1242/bio.201410041.

(59) Binz, H. K.; Amstutz, P.; Kohl, A.; Stumpp, M. T.; Briand, C.; Forrer, P.; Grütter, M. G.; Plückthun, A. High-Affinity Binders Selected from Designed Ankyrin Repeat Protein Libraries. Nat. Biotechnol. 2004, 22 (5), 575–582. https://doi.org/10.1038/nbt962.

(60) Zhuang, X.; Ha, T.; Kim, H. D.; Centner, T.; Labeit, S.; Chu, S. Fluorescence Quenching: A Tool for Single-Molecule Protein-Folding Study. Proc. Natl. Acad. Sci. 2000, 97 (26), 14241–14244. https://doi.org/10.1073/pnas.97.26.14241.

(61) Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.; Wood, K. V. NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. ACS Chem. Biol. 2016, 11 (2), 400–408. https://doi.org/10.1021/acschembio.5b00753.

Chapter III: The design and characterization of cancer protein-responsive cages

Eric J. Lee[1]

[1] UCLA Department of Chemistry and Biochemistry, Los Angeles, CA 90095

ABSTRACT

Recent advances in nanoparticle (NP)-based drug delivery have demonstrated exciting potential in cancer therapy. These systems range from pro-apoptotic metal NPs to small molecule-encapsulating protein assemblies. In general, a nanomaterials-based approach to cancer therapy is attractive due to biocompatibility and low invasiveness of treatment. Despite these advances, however, many drug delivery systems are either susceptible to off-target effects or require laborious engineering to achieve high specificity against a single target. Therefore, a facile platform that can achieve both modularity and specificity for a wide range of targets is desirable. Previously, our group demonstrated a new type of protein nanocage, the ligand-operable cage (LOC), which has the potential to release an encapsulated cargo upon binding a target protein. In this work, we aimed to explore the modularity of the LOC platform and their ability to target different cancer proteins. We designed new variants of the original T33-51-LOC nanocages that bind KRAS, BARD1, and BRCA1. The aberrant expression and/or mutagenesis of these targets are implicated in multiple clinical cancers. Here we discuss the design, synthesis, and characterization of these new LOC designs, as well as challenges and future directions.

INTRODUCTION

Cancer treatment research has been a global imperative for over a century[1]. Current methods primarily comprise surgery, radiotherapy, chemotherapy, and immunotherapy[2–5]. While these remain the standard of care for many patients, their limitations are apparent in more extreme cases, such as late-stage metastatic cancers[6,7]. Thus, the field of nanomaterials-based

drug delivery has emerged to meet this challenge. Already, many nanoparticle-based treatments have demonstrated clinical potential, with a few approved for commercial use[8].

Among the emergent approaches are protein nanocage-based drug delivery systems, mostly involving drug-loaded ferritin[9–11]. Protein nanocages generally possess favorable characteristics for engineering, such as an exterior surface for surface modification and an interior cavity for cargo storage[12]. Previous work has demonstrated the ability to leverage these features for a diverse range of applications, including materials, therapeutics, encapsulation, and delivery[13–18]. However, prior studies for protein nanocage-based drug delivery mainly rely on spontaneous or environmentally triggered disassembly of the cage to release drug cargo, which could lead to off-target effects.

We previously demonstrated a new type of nanocage - the ligand-operable cage (LOC) - which possesses the ability to disassemble upon specifically binding a target protein[19]. This was achieved by fusing target-specific designed ankyrin repeat proteins (DARPins) to the surface of the nanocage[20,21]. Furthermore, these DARPins were arranged in a specific orientation, such that binding of many cognate ligands (i.e., the target protein) to the surface DARPins was sterically incompatible with the intact cage assembly. Therefore, high ligand concentrations drove disassembly of the nanocage.

In this work, we aimed to reprogram one of our previously designed LOCs to bind alternative, cancer-implicated proteins. To this end, we generated new designs derived from the original T33-51-LOC that bind Kirsten rat sarcoma viral oncogene homologue (KRAS), breast cancer gene 1 (BRCA1), and BRCA1-associated ring domain 1 (BARD1) proteins. KRAS is one of the most extensively characterized oncogenes, associated with multiple fatal cancers and possessing the highest mutation rate[22]. BRCA1-BARD1 is a tumor suppressor complex,

mutations of which lead to various cancers including breast and ovarian cancer[23–26]. We chose these targets not only due to their therapeutic significance, but also the availability of DARPin sequences that bind them. Designs were recombinantly expressed in *E. coli*, purified, and characterized by size exclusion chromatography (SEC), dynamic light scattering (DLS), and negative-stain electron microscopy (nsEM). We discuss preliminary findings, experimental challenges, and future directions.

RESULTS AND DISCUSSION

Design of αKRAS-, αBARD1-, and αBRCA1-LOCs

We based our cancer-targeting LOC designs on the original T33-51-LOC described in previous work (Figure 3.1A)[19]. The LOC variants discussed in this work follow the same general design and assembly principles. All variants were designed with the T33-51 nanocage core, which had been previously characterized[27]. Surface adaptors used in this study also comprised DARPins, which were recombinantly fused to the surface residues of T33-51 via helical linker. Furthermore, the BARD1 and the BRCA1 constructs used in this study comprised the C-terminal domain of each protein[28–30].

We generated models of αKRAS-LOC, αBARD1-LOC, and αBRCA1-LOCs via AlphaFold[31]. To do this, we replaced the MBP-binding DARPin domain of the original T33-51-LOC with the respective KRAS, BARD1, or BRCA1-binding DARPin sequence and used these new sequences as our structure prediction query on ColabFold[32]. All input sequences returned a virtually identical secondary structure prediction as the original T33-51-LOC, indicating that the new DARPins did not affect the overall orientation of the adaptor binding sites[19]. We

demonstrated that, using a previous KRAS-bound DARPin structure, the mode of ligand-based disruption was nearly identical to that of the original T33-51-LOC with MBP (Figure 3.2A-C)[19,33]. That is, in the unbound state, no steric clashes were observed in the modeled assembly. However, upon binding of multiple KRAS molecules, we observed a steric incompatibility between the bound ligand molecules and the intact cage assembly. Therefore, we hypothesized that higher concentrations of KRAS would drive disassembly of the nanocage. Lacking structural data on BRCA1- and BARD1-bound DARPin, we continued with the assumption that, due to the orientation of DARPin binding sites, if BRCA1 and BARD1 could bind they would drive disassembly of their respective LOCs in a similar manner.

Expression and purification of αKRAS-, αBARD1-, and αBRCA1-LOCs

All ligands and LOC variants appeared to pass initial small-scale expression screening (Figure 3.3A)**.** SDS-PAGE analysis was performed following scaled-up expression and affinity purification of proteins (Figure 3.3B). Following scaled-up expression, purified eluate fractions containing αKRAS-LOC and αBRCA1-LOC demonstrated only faint band intensity at their expected molecular weights. These observations were consistent across different expression conditions, including autoinduction, which did not noticeably improve expression levels. Due to the lack of adequate BARD1 and αBARD1-LOC expression following upscaled production, we did not pursue additional experimentation with these designs, and instead further scaled up αKRAS-LOC and αBRCA1-LOC production (Figure 3.3C-D).

Size-exclusion chromatography (SEC) was performed on concentrated elution fractions of αKRAS and αBRCA1-LOC cages. Resulting chromatographs revealed that most of the

protein for both designs did not elute at the volume corresponding to assembled cages, which is approximately 13 mL (Figure 3.3E-F). This indicated that most of the cage proteins in both samples were already disassembled or favored smaller subassemblies in the absence of their respective ligand.

Characterization of αKRAS- and αBRCA1-LOC assembly and ligand binding

To characterize αKRAS- and αBRCA1-LOC assembly, we first examined the SEC-purified cage proteins via NsEM. SEC fractions collected between 13-14 mL demonstrated the most cage-like assemblies, with particles closest to expected size and symmetry (Figure 3.4). Following concentration of purified cage proteins, αBRCA1-LOC yields were extremely low, as demonstrated by little to no particle density apparent under nsEM (Figure 3.4A). Furthermore, the few αBRCA1-LOC assemblies present in samples were polydisperse with no clear symmetry.

αKRAS-LOC, on the other hand, demonstrated stronger particle density with cage-like assemblies of expected pseudo four-fold symmetry and approximately 10 nm diameter core (the surface DARPins, which would contribute additional 5 nm diameter, could not be resolved) (Figure 3.4B). We therefore pursued further imaging experiments using the same purified αKRAS-LOC proteins and incubated them with threefold molar excess of purified KRAS. NsEM images were then collected following approximately 30 minutes post-incubation; however, no significant changes in cage particle density or polydispersity were clearly observed (Figure 3.4C).

We sought a more quantitative method to characterize the nanocages' assembly and ligand-binding capability. To this end, we used dynamic light scattering (DLS) to examine the

assembly states of αKRAS-LOC and αBRCA1-LOC when incubated with mock buffer or cognate ligand (Figure 3.5). For αBRCA1-LOC, in the absence of BRCA1, we observed that most particles in solution lied between 6 and 7 nanometers in radius, which is slightly smaller than the expected, approximately 7.5 nm radius (Figure 3.5A). Notably, these results contrasted with our observations from nsEM. After incubating αBRCA1-LOC with BRCA1, however, we observed a shift in the radius towards approximately 1 nm radius in some samples, which suggests that αBRCA1-LOC was disassembling upon binding BRCA1. For αKRAS-LOC, samples containing only cages yielded a majority particle radius of 7 to 8 nm (Figure 3.5B). Upon addition of KRAS, we observed that some samples demonstrated a decrease in particle radius, though these results suggested the appearance of small, polydisperse subassemblies, rather than completely disassembled particles. Notably, these observations for the ligand-triggered disassembly of αKRAS-LOC and αBRCA1-LOC could not be consistently replicated, and further experimental optimization is required for conclusive analysis.

CONCLUSIONS

In this work, we designed new LOCs with cancer protein-binding adaptors. These designs were based on our original T33-51-LOC core, reprogrammed to bind different targets. This was achieved by substituting alternative DARPin sequences for the original T33-51-LOC MBP-binding DARPin sequence. AlphaFold predictions of these designs yielded virtually identical secondary structure models to the original T33-51-LOC. This was within our expectations, as rigid, helical-based platforms should share the same overall structure provided only the sequences of the variable binding regions are mutated.

While the backbone structures of the new designs were supported by AlphaFold, the expression levels of these proteins remained extremely low compared to the original T33-51-LOC. Since the DARPins are the only source of sequence deviation from the original T33-51-LOC, it follows that the new DARPin sequences are the probable cause of low expression. Further codon optimization could be performed for αBARD-, αKRAS-, and αBRCA1-LOC, where silent mutations are incorporated into respective DARPin repeats. However, it is also possible that, regardless of codon optimization, certain DARPin amino acid sequences may have unforeseen folding consequences during recombinant expression. Since the original T33-51-LOC core is nearly entirely helical, as are the DARPins themselves, it is possible that non-specific affinity between helical regions interact during the folding process, leading to undesired intermediates or subassemblies.

Furthermore, our findings did not clearly indicate that αKRAS-LOC and αBRCA1-LOC bound their respective target ligand and subsequently disassembled. Although DLS experiments demonstrated that some cage samples experienced a significant decrease in average apparent particle radius upon incubation with their cognate ligand, these observations were not readily replicated (Figure 3.5A-B). This is likely because neither αKRAS-LOC nor αBRCA1-LOC were readily purified as monodisperse particles, causing inconsistencies between sample preparations (Figure 3.4A-B). However, if additional optimization was performed such that αKRAS-LOC and αBRCA1-LOC could be isolated as monodisperse cages, the ensuing binding experiments may be more informative.

In theory, the modularity of a designed LOC spans as wide as the sequence diversity of its fused adaptor (in this case, the diversity of available DARPin sequences raised against therapeutic targets). For example, studies using other exemplary cancer targets such as HER2,

VEGF, and various interleukins (all possessing an associated DARPin binder) offer avenues for experimentation. Furthermore, investigations into the relative viability of certain LOC targets versus others based on various characteristics (size, shape, affinity, etc.) would be highly informative, considering the diversity of known adaptor targets. Other helical binders, such as affibodies, may prove useful as alternative adaptors to DARPins[34]. Affibodies are also significantly shorter than DARPins, which may affect overall cage stability and/or expression levels. Nanobodies, which are not helical, would also be interesting to test as the choice adaptor of cancer-targeting LOCs[35]. Since many cancer targets possess high molecular weight (often large extracellular domains of transmembrane proteins), it may not be as necessary to have a LOC adaptor that is rigidly fused via helical connection to drive disassembly of the nanocage. Future experiments may shed light on these open questions.

MATERIALS AND METHODS

Sequence redesign of LOC variants

Sequences for BARD1 and BRCA1 BRCT domains, and their associated DARPins, were graciously provided by Dr. Mark Arbing from UCLA-DOE.

All LOC variants were initially based on the original T33-51-LOC core, the creation of which was discussed previously[19]. The T33-51-LOC amino acid sequence was modified only in the DARPin-coding region, where the sequence of MBP-binding DARPin was replaced with that of either KRAS-, BARD1-, or BRCA1-binding DARPin. The rest of the amino acid sequence was kept identical. This produced variants where the overall cage architectures were virtually identical, except for the variable binding sites of the surface-fused DARPins. All recombinant

LOCs comprised two genetic components: component A, which comprised one of the trimeric proteins, a helical linker, and the fused DARPin; and component B, which comprised the other trimeric protein and a polyhistidine tag.

AlphaFold-supported modeling of LOC variants

The component A trimer structure of each LOC variant was predicted using AlphaFold-based Collab fold[31,32]. The amino acid sequence for each LOC-A variant was input as a trimer.

For αKRAS-LOC component A, the predicted trimer structure of was superimposed with 1-3 copies of KRAS-bound DARPin (PDB: 6H46)[33]. Curative modeling was performed on PyMOL[36]. This ensured that the binding of multiple KRAS molecules would be sterically incompatible with the assembly of the trimer, and therefore the overall nanocage. For αBARD1-LOC and αBRCA1-LOC, no structural information was available for their ligand-bound DARPins.

Recombinant Protein Expression and Purification

All gene fragments encoding BARD1, BRCA1, KRAS, αBARD-, αKRAS-, and αBRCA1-LOC were purchased from Twist Bioscience. Plasmid sequences were verified via Sanger Sequencing (Azenta Life Sciences/Genewiz). All genes were inserted into a pET-22b+ vector (Novagen) via Gibson assembly with a NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). DNA manipulations and amplification were carried out in Escherichia coli XL2 cells (Agilent). For all LOC variants, both A and B components were generated via polycistronic expression. All other constructs were generated as a single cistron.

Expression was performed using Escherichia coli BL21 (DE3) cells (New England Biolabs). Transformed cells were grown in Luria–Bertani (LB) Broth supplemented with 100 μg/mL ampicillin at 37 °C to an OD600 between 0.8 and 1.0 for LOC variants, and between 0.6 and 0.8 for all other proteins. 100 μM isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to induce protein expression. Following induction, cultures were grown overnight at 18 °C before harvesting via centrifugation at 3500g for 15 min at 4 °C.

Harvested cells were resuspended in 50 mM Tris-HCl (pH, 8.0), 2% w/v glycerol, 5 mM TCEP-HCl, 250 mM NaCl, and 20 mM imidazole. Prior to lysis, resuspension buffers were supplemented with an EDTA-free protease inhibitor tablet (Pierce), 10 U/mL benzonase nuclease, and 0.1 mg/mL lysozyme. Resuspended cells were then pressure-homogenized using an Emulsiflex C3 Homogenizer (Avestin). Lysates were clarified via centrifugation at 20,000g at 4 °C for at least 30 min.

Clarified lysates were loaded onto Ni-NTA resin (Thermo Fisher Scientific) that was pre-equilibrated in resuspension buffer, then washed/eluted with purification buffers (50 mM Tris-HCl pH 8.0, 2% w/v glycerol, and 250 mM NaCl, further supplemented with 20, 50, 100, or 250 mM imidazole). LOC variants were further purified via an NGC chromatography system (Bio-Rad) over a Superose 6 Increase 10/300 GL resin column (Cytiva). Sample runs were conducted in 50 mM Tris-HCl at pH 8.0, 150 mM NaCl, and 2% w/v glycerol at 4 °C with a flow rate of 0.3 mL/min.

Negative-stain electron microscopy

A 5 μL aliquot of purified nanocages from size exclusion chromatography samples (ranging from 0.01 to 0.05 mg/mL) was stained with 2% uranyl acetate on glow-discharged 300-

mesh copper grids (Ted Pella, Inc.). Nanocages were imaged using an FEI Tecnai T12 transmission electron microscope at 120 kV. Samples mixed with cognate ligand were prepared in a molar ratio of 1:3 cage protein to ligand.

Dynamic light scattering

For LOC assembly and disassembly analysis, 100 nM isolated cages from SEC were mixed with elution buffer (mock) or 300 nM cognate ligand and incubated at 25 °C for at least 30 min. All reads were collected at 25 °C on a DynaPro Platereader-II (Wyatt Technology).

SDS-PAGE

Protein samples were mixed with SDS loading dye (0.004% Bromophenol blue, 6% glycerol, 2% sodium dodecyl sulfate, 50 mM Tris-HCl pH 6.8, 5 mM DTT) before boiling for at least 2 min. Samples were loaded onto a Mini-Protean TGX polyacrylamide gel (BioRad). Electrophoresis was performed at 200 V for 30 min. Gels were stained with Coomassie Blue and destained.

FIGURES

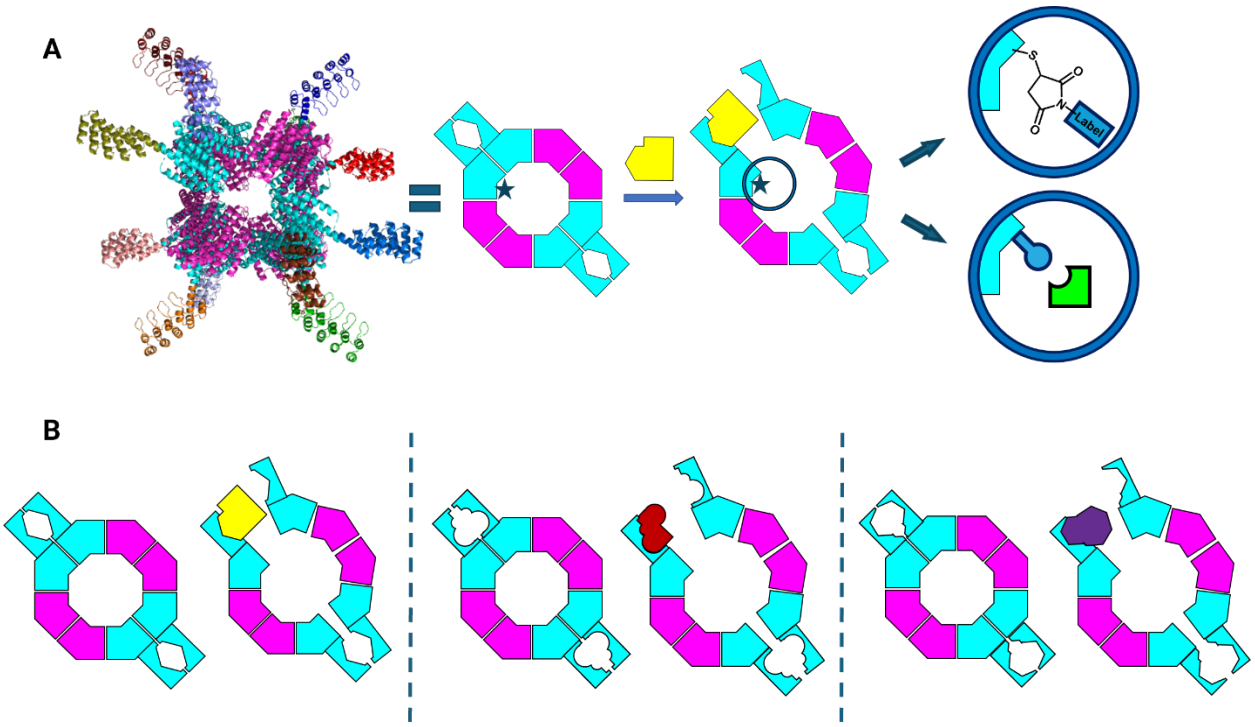

Figure 3.1. Reprogramming T33-51-LOC. (A) Previously designed T33-51-LOC with MBP-specific surface adaptors (far left); binding of ligand (yellow) to adaptors (cyan) leads to disassembly and exposure of cargo element (star), which can be coupled to a fluorescent (top inset) or protein-protein interaction (bottom inset) readout. (B) Original T33-51-LOC platform (left) can be readily modified to bind different proteins of interest (middle, red; and right, purple) by substituting adaptor sequences.

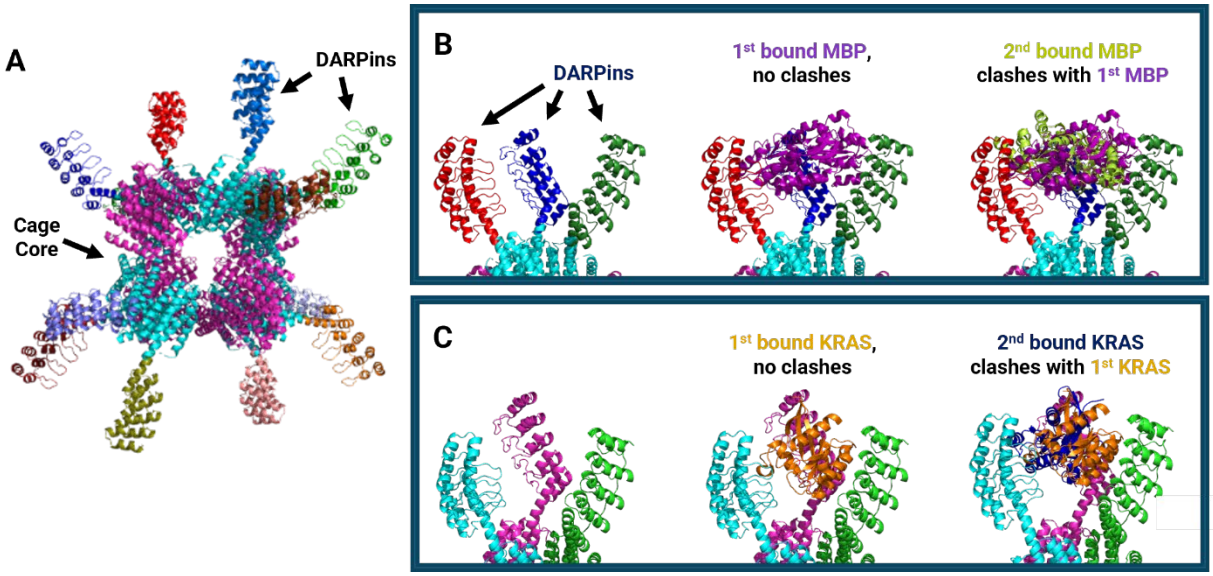

Figure 3.2. Modeling of T33-51-LOC variants. A) Design model of T33-51 nanocage core (magenta and cyan, PDB: 5CY5) with surface-fused MBP-binding DARPins (various colors) (whole assembly referred to as T33-51-LOC). B) Magnified view of select T33-51-LOC DARPins: in absence of cognate protein ligand (MBP), no steric clashes observed (left); single MBP (purple) binding to DARPin (red) is permitted without steric clashes (middle); second MBP (light green) binding to DARPin (dark green) is sterically incompatible with cage assembly (right). (C) AlphaFold-generated model of T33-51-LOC with surface-fused KRAS-binding DARPins: in absence of cognate protein ligant (KRAS), no steric clashes observed (left); single KRAS (orange) binding to DARPin (pink) is permitted without steric clashes (middle); second MBP (dark blue) binding to DARPin (green) is sterically incompatible with cage assembly (right).

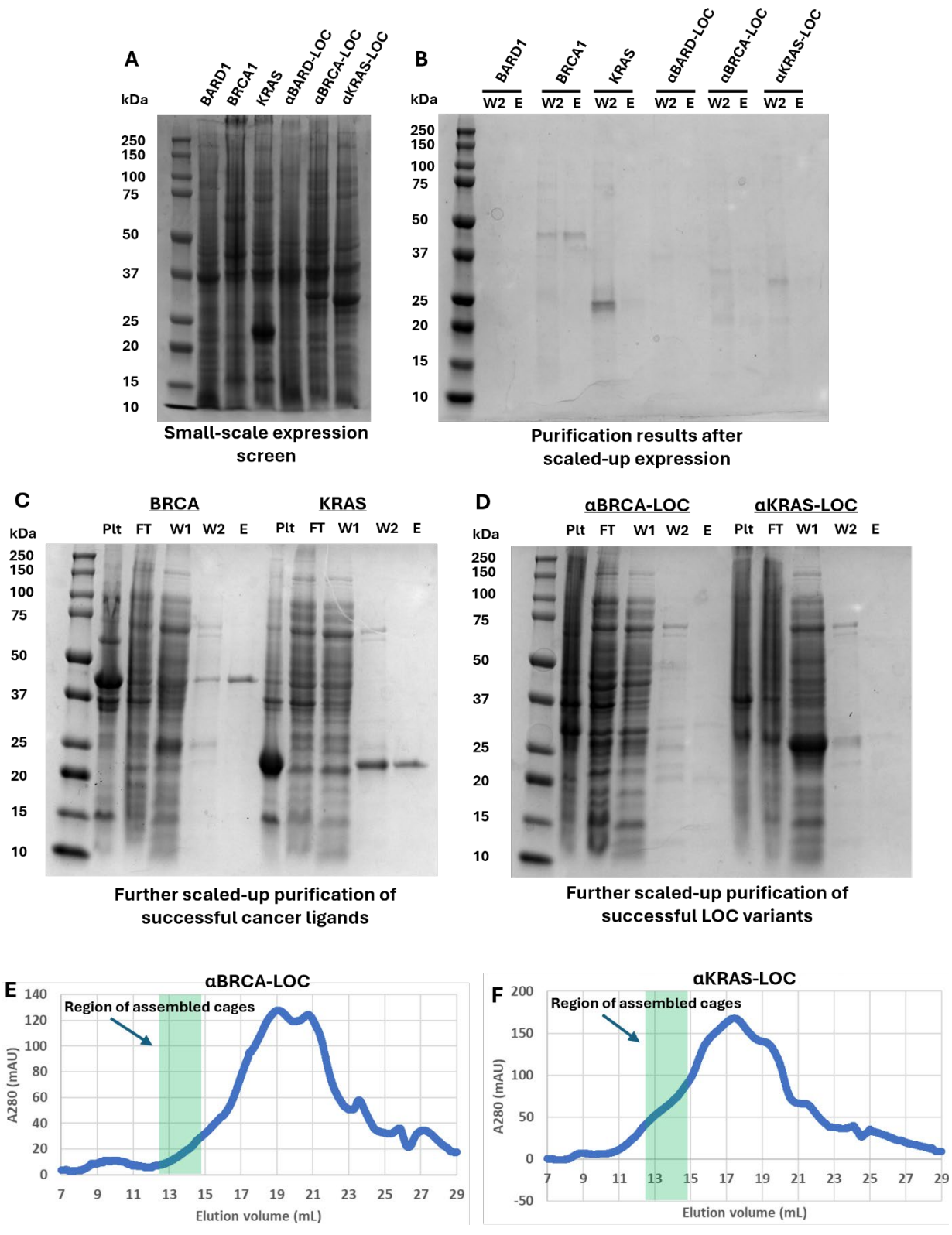

Figure 3.3. Expression and purification of T33-51-LOC variants. (A) SDS-PAGE analysis of T33-51-LOC variants and target ligands following small-scale expression test. (B) SDS-PAGE

analysis of affinity chromatography-purified T33-51-LOC variants and target ligands following scaled-up expression. (C) SDS-PAGE analysis of further scaled-up purification fractions of BRCA and KRAS. (D) SDS-PAGE analysis of further scaled-up purification of αBRCA-LOC and αKRAS-LOC. (E) SEC chromatograph of αBRCA-LOC following affinity chromatography. (F) SEC chromatograph of αKRAS-LOC following affinity chromatography.

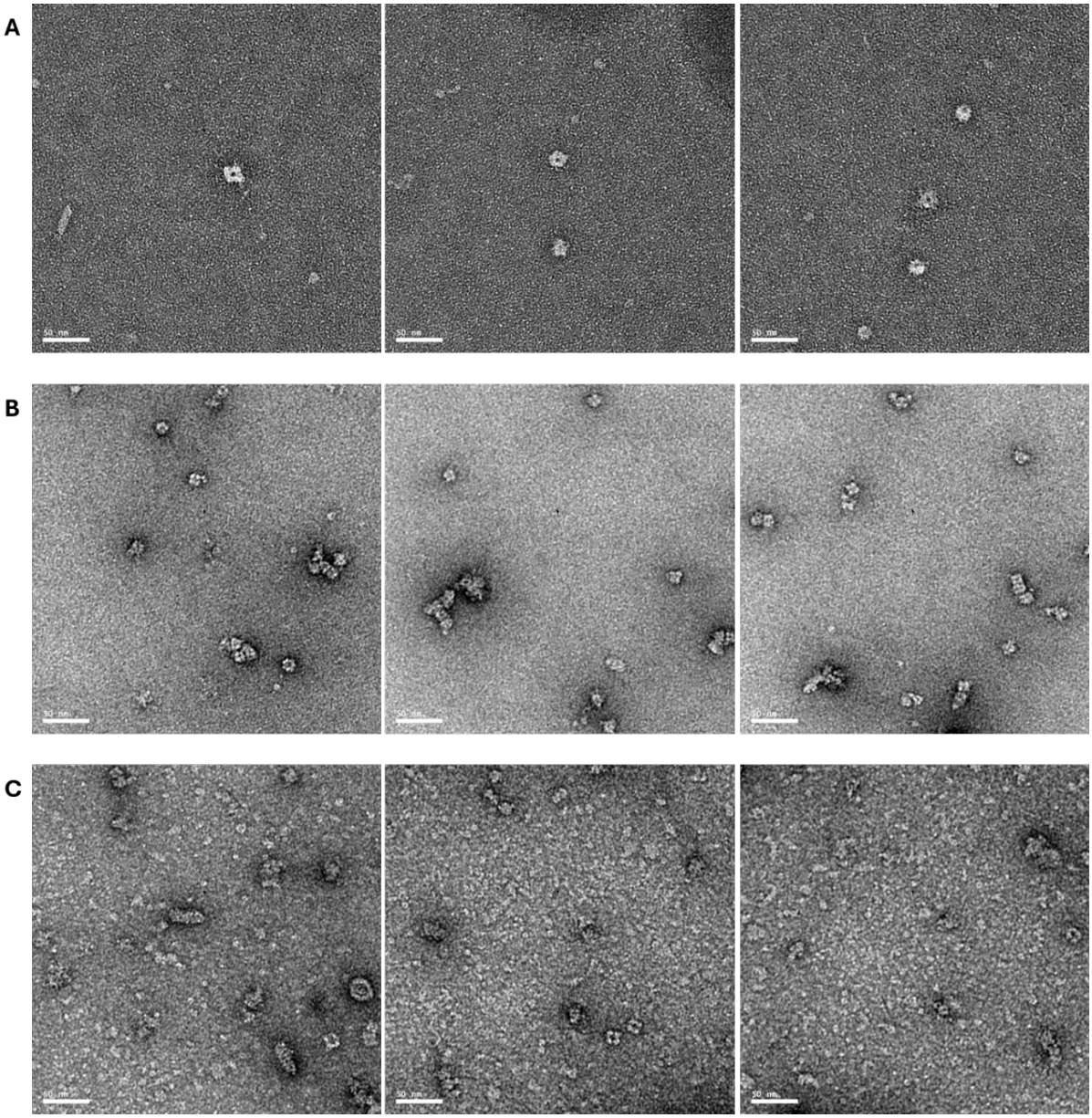

Figure 3.4. Representative negative-stain electron micrographs of T33-51-LOC variants following SEC purification. (A) αBRCA1-LOC protein cages. (B) αKRAS-LOC protein cages incubated with mock and (C) with KRAS. 50 nm scale bars.

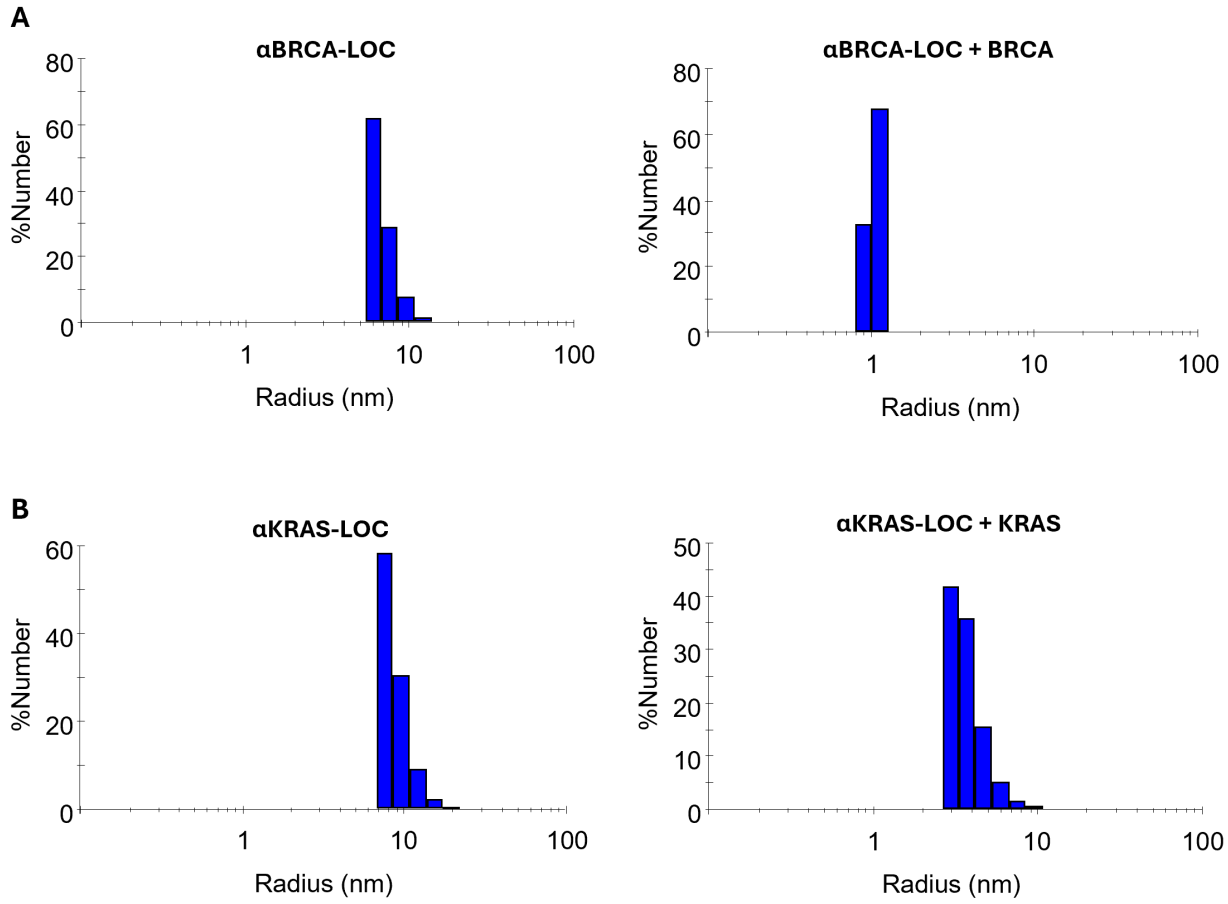

Figure 3.5. Example DLS analysis of T33-51-LOC variants. Percent particle number versus apparent particle radius of (A) αBRCA1-LOC incubated without (left) and with ligand (right) in some samples; (B) αKRAS-LOC incubated without (left) and with ligand (right) in some samples. Results from both (A) and (B) could not be consistently replicated between experiments.

Protein	Amino acid sequence
αBARD1-T33-51-LOC-A	MCGGGRITTKVGDKGSTRLFGGEEVWKDDPIIEANGTLDELTSFIGEA KHYVDEEMKGILEEIQNDIYKIMGEIGSKGKIEGISEERIKWLAGLIERYS EMNKLSFVLPGGTLESAKLDVCRTIARRAERKVATVLREFGIGTLAAIYL ALLSRLLFLLARVIEIEKNKL**AQ***GKKLLEAARAGQDDEVRILMANGADV NASNWWGWTPLHAAATNGHLEIVDVLLANGADVNASDDDGETPLHS AAYDGHLEIVDVLLAYGADVNASDEWGQTPLHSAAHQGHLEIVDVLLA YGADVNAQDKFGKTPFDLAIDNGNEDIAEVLQKAA*
αBRCA1-T33-51-LOC-A	MCGGGRITTKVGDKGSTRLFGGEEVWKDDPIIEANGTLDELTSFIGEA KHYVDEEMKGILEEIQNDIYKIMGEIGSKGKIEGISEERIKWLAGLIERYS EMNKLSFVLPGGTLESAKLDVCRTIARRAERKVATVLREFGIGTLAAIYL ALLSRLLFLLARVIEIEKNKL**AQ***GKKLLEAARAGQDDEVRILMANGADV NASDHAGTTPLHAAAEWGHLEIVDVLLAHGADVNASDYWGYTPLHAA AEHGHLEIVDVLLAHGADVNASDWYGWTPLHLAADSGHLEIVDVLLAY GADVNAQ*
αKRAS-T33-51-LOC-A	MCGGGRITTKVGDKGSTRLFGGEEVWKDDPIIEANGTLDELTSFIGEA KHYVDEEMKGILEEIQNDIYKIMGEIGSKGKIEGISEERIKWLAGLIERYS EMNKLSFVLPGGTLESAKLDVCRTIARRAERKVATVLREFGIGTLAAIYL ALLSRLLFLLARVIEIEKNKL**AQ***GKKLLEAARAGQDDEVRILMANGADV NASDRWGWTPLHLAAWWGHLEIVEVLLKHGADVNAADLHGQTPLHL AAMVGHLEIVEVLLKYGADVNAKDTMGATPLHLAAQSGHLEIVEVLLK NGADVNAQDKFGKTAFDISIDNGNEDLAEILQK*
T33-51-LOC-B	MFTRRGDQGETDLANRARVGKDSPVVEVQGTIDELNSFIGYALVLSR WDDIRNDLFRIQNDLFVLGEDVSTGGKGRTVTMDMIIYLIKRSVEMKAE IGKIELFVVPGGSVESASLHMARAVSRRLERRIKAASELTEINANVLLYA NMLSNILFMHALISNKRLNIPEKIWSIHRVSLEHHHHHH
His-KRAS (G12C)	MGSSHHHHHHSSGENLYFQS**MTEYKLVVVGAGGVGKSALTIQLIQN HFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAMRDQYMRT GEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPS RTVDTKQAQDLARSYGIPFIETSAKTRQGVDDAFYTLVREIRKHKEK**
His-SUMO-BARD1 BRCT	MGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETH INLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQ ADQTPEDLDMEDNDIIEAHREQIGGSLEVLFQGPGLNDIFEAQKIEWHE ENLYFQS**MKRNHRGETLLHIASIKGDIPSVEYLLQNGSDPNVKDHAG WTPLHEACNHGHLKVVELLLQHKALVNTTGYQNDSPLHDAAKNGHV DIVKLLLSYGASRNAVNIFGLRPVDYTDDESMKSLLLLPEKNESSSAS HCSVMNTGQRRDGPLVLIGSGLSSEQQKMLSELAVILKAKKYTEFDS TVTHVVVPGDAVQSTLKCMLGILNGCWILKFEWVKACLRRKVCEQE EKYEIPEGPRRSRLNREQLLPKLFDGCYFYLWGTFKHHPKDNLIKLV TAGGGQILSRKPKPDSDVTQTINTVAYHARPDSDQRFCTQYIIYEDLC NYHPERVRQGKVWKAPSSWFIDCVMSFELLPLDS**
His-SUMO-BRCA1 BRCT	MGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETH INLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQ ADQTPEDLDMEDNDIIEAHREQIGGSSLEVLFQGPGLNDIFEAQKIEWH EENLYFQS**STERVNKRMSMVVSGLTPEEFMLVYKFARKHHITLTNLIT EETTHVVMKTDAEFVCERTLKYFLGIAGGKWVVSYFWVTQSIKERKM LNEHDFEVRGDVVNGRNHQGPKRARESQDRKIFRGLEICCYGPFTN MPTDQLEWMVQLCGASVVKELSSFTLGTGVHPIVVVQPDAWTEDNG FHAIGQMCEAPVVTREWVLDSVALYQCQELDTYLIPQIPSHY**

Table 3.S1. Amino acid sequences of LOC variants and cancer targets. Underlined and bolded amino acids indicate helical linker. Underlined and italicized amino acids indicate DARPin. Bolded regions indicate cancer protein.

REFERENCES

(1) Arruebo, M.; Vilaboa, N.; Sáez-Gutierrez, B.; Lambea, J.; Tres, A.; Valladares, M.; González-Fernández, Á. Assessment of the Evolution of Cancer Treatment Therapies. *Cancers* **2011**, *3* (3), 3279–3330. https://doi.org/10.3390/cancers3033279.

(2) Mun, E. J.; Babiker, H. M.; Weinberg, U.; Kirson, E. D.; Von Hoff, D. D. Tumor-Treating Fields: A Fourth Modality in Cancer Treatment. *Clin. Cancer Res.* **2018**, *24* (2), 266–275. https://doi.org/10.1158/1078-0432.CCR-17-1117.

(3) Schaue, D.; Micewicz, E. D.; Ratikan, J. A.; Xie, M. W.; Cheng, G.; McBride, W. H. Radiation and Inflammation. *Semin. Radiat. Oncol.* **2015**, *25* (1), 4–10. https://doi.org/10.1016/j.semradonc.2014.07.007.

(4) Zhang, Y.; Zhang, Z. The History and Advances in Cancer Immunotherapy: Understanding the Characteristics of Tumor-Infiltrating Immune Cells and Their Therapeutic Implications. *Cell. Mol. Immunol.* **2020**, *17* (8), 807–821. https://doi.org/10.1038/s41423-020-0488-6.

(5) Hellmann, M. D.; Li, B. T.; Chaft, J. E.; Kris, M. G. Chemotherapy Remains an Essential Element of Personalized Care for Persons with Lung Cancers. *Ann. Oncol.* **2016**, *27* (10), 1829–1835. https://doi.org/10.1093/annonc/mdw271.

(6) Bishop, A. J.; Zagars, G. K.; Torres, K. E.; Bird, J. E.; Feig, B. W.; Guadagnolo, B. A. Malignant Peripheral Nerve Sheath Tumors: A Single Institution's Experience Using Combined Surgery and Radiation Therapy. *Am. J. Clin. Oncol.* **2018**, *41* (5), 465–470. https://doi.org/10.1097/COC.0000000000000303.

(7) Tian, L.; Wang, S.; Jiang, S.; Liu, Z.; Wan, X.; Yang, C.; Zhang, L.; Zheng, Z.; Wang, B.; Li, L. Luteolin as an Adjuvant Effectively Enhances CTL Anti-Tumor Response in B16F10

Mouse Model. *Int. Immunopharmacol.* **2021**, *94*, 107441. https://doi.org/10.1016/j.intimp.2021.107441.

(8) Liu, G.; Yang, L.; Chen, G.; Xu, F.; Yang, F.; Yu, H.; Li, L.; Dong, X.; Han, J.; Cao, C.; Qi, J.; Su, J.; Xu, X.; Li, X.; Li, B. A Review on Drug Delivery System for Tumor Therapy. *Front. Pharmacol.* **2021**, *12*, 735446. https://doi.org/10.3389/fphar.2021.735446.

(9) Chen, Z.; Zhai, M.; Xie, X.; Zhang, Y.; Ma, S.; Li, Z.; Yu, F.; Zhao, B.; Zhang, M.; Yang, Y.; Mei, X. Apoferritin Nanocage for Brain Targeted Doxorubicin Delivery. *Mol. Pharm.* **2017**, *14* (9), 3087–3097. https://doi.org/10.1021/acs.molpharmaceut.7b00341.

(10) Yang, Z.; Wang, X.; Diao, H.; Zhang, J.; Li, H.; Sun, H.; Guo, Z. Encapsulation of Platinum Anticancer Drugs by Apoferritin. *Chem. Commun.* **2007**, No. 33, 3453. https://doi.org/10.1039/b705326f.

(11) Deng, G.; Li, Y.; Liang, N.; Hu, P.; Zhang, Y.; Qiao, L.; Zhang, Y.; Xie, J.; Luo, H.; Wang, F.; Chen, F.; Liu, F.; Xu, D.; Zhang, J. Ferritin in Cancer Therapy: A Pleiotropic Tumoraffin Nanocage-based Transport. *Cancer Med.* **2023**, *12* (10), 11570–11588. https://doi.org/10.1002/cam4.5778.

(12) Cannon, K. A.; Ochoa, J. M.; Yeates, T. O. High-Symmetry Protein Assemblies: Patterns and Emerging Applications. *Curr. Opin. Struct. Biol.* **2019**, *55*, 77–84. https://doi.org/10.1016/j.sbi.2019.03.008.

(13) Marcandalli, J.; Fiala, B.; Ols, S.; Perotti, M.; de van der Schueren, W.; Snijder, J.; Hodge, E.; Benhaim, M.; Ravichandran, R.; Carter, L.; Sheffler, W.; Brunner, L.; Lawrenz, M.; Dubois, P.; Lanzavecchia, A.; Sallusto, F.; Lee, K. K.; Veesler, D.; Correnti, C. E.; Stewart, L. J.; Baker, D.; Loré, K.; Perez, L.; King, N. P. Induction of Potent Neutralizing

Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **2019**, *176* (6), 1420-1431.e17. https://doi.org/10.1016/j.cell.2019.01.046.

(14) McConnell, S. A.; Cannon, K. A.; Morgan, C.; McAllister, R.; Amer, B. R.; Clubb, R. T.; Yeates, T. O. Designed Protein Cages as Scaffolds for Building Multienzyme Materials. *ACS Synth. Biol.* **2020**, *9* (2), 381–391. https://doi.org/10.1021/acssynbio.9b00407.

(15) Liu, Y.; Huynh, D. T.; Yeates, T. O. A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold. *Nat. Commun.* **2019**, *10* (1), 1864. https://doi.org/10.1038/s41467-019-09836-0.

(16) 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. https://doi.org/10.1021/acschembio.1c00007.

(17) Hills, R. A.; Tan, T. K.; Cohen, A. A.; Keeffe, J. R.; Keeble, A. H.; Gnanapragasam, P. N. P.; Storm, K. N.; Hill, M. L.; Liu, S.; Gilbert-Jaramillo, J.; Afzal, M.; Napier, A.; James, W. S.; Bjorkman, P. J.; Townsend, A. R.; Howarth, M. Multiviral Quartet Nanocages Elicit Broad Anti-Coronavirus Responses for Proactive Vaccinology. *BioRxiv Prepr. Serv. Biol.* **2023**, 2023.02.24.529520. https://doi.org/10.1101/2023.02.24.529520.

(18) Zhang, T.; Lv, C.; Chen, L.; Bai, G.; Zhao, G.; Xu, C. Encapsulation of Anthocyanin Molecules within a Ferritin Nanocage Increases Their Stability and Cell Uptake Efficiency. *Food Res. Int.* **2014**, *62*, 183–192. https://doi.org/10.1016/j.foodres.2014.02.041.

(19) Lee, E. J.; Gladkov, N.; Miller, J. E.; Yeates, T. O. Design of Ligand-Operable Protein-Cages That Open Upon Specific Protein Binding. *ACS Synth. Biol.* **2024**, *13* (1), 157–167. https://doi.org/10.1021/acssynbio.3c00383.

(20) Steiner, D.; Forrer, P.; Plückthun, A. Efficient Selection of DARPins with Sub-Nanomolar Affinities Using SRP Phage Display. *J. Mol. Biol.* **2008**, *382* (5), 1211–1227. https://doi.org/10.1016/j.jmb.2008.07.085.

(21) Seeger, M. A.; Zbinden, R.; Flütsch, A.; Gutte, P. G. M.; Engeler, S.; Roschitzki-Voser, H.; Grütter, M. G. Design, Construction, and Characterization of a Second-Generation DARPin Library with Reduced Hydrophobicity: Second-Generation DARPin Library. *Protein Sci.* **2013**, *22* (9), 1239–1257. https://doi.org/10.1002/pro.2312.

(22) Huang, L.; Guo, Z.; Wang, F.; Fu, L. KRAS Mutation: From Undruggable to Druggable in Cancer. *Signal Transduct. Target. Ther.* **2021**, *6* (1), 386. https://doi.org/10.1038/s41392-021-00780-4.

(23) Gorodetska, I.; Kozeretska, I.; Dubrovska, A. *BRCA* Genes: The Role in Genome Stability, Cancer Stemness and Therapy Resistance. *J. Cancer* **2019**, *10* (9), 2109–2127. https://doi.org/10.7150/jca.30410.

(24) Cavanagh, H.; Rogers, K. M. A. The Role of BRCA1 and BRCA2 Mutations in Prostate, Pancreatic and Stomach Cancers. *Hered. Cancer Clin. Pract.* **2015**, *13* (1), 16. https://doi.org/10.1186/s13053-015-0038-x.

(25) Mersch, J.; Jackson, M. A.; Park, M.; Nebgen, D.; Peterson, S. K.; Singletary, C.; Arun, B. K.; Litton, J. K. Cancers Associated with *BRCA 1* and *BRCA 2* Mutations Other than Breast and Ovarian. *Cancer* **2015**, *121* (2), 269–275. https://doi.org/10.1002/cncr.29041.

(26) Noh, J. M.; Choi, D. H.; Baek, H.; Nam, S. J.; Lee, J. E.; Kim, J. W.; Ki, C.-S.; Park, W.; Huh, S. J. Associations between *BRCA* Mutations in High-Risk Breast Cancer Patients and Familial Cancers Other than Breast or Ovary. *J. Breast Cancer* **2012**, *15* (3), 283. https://doi.org/10.4048/jbc.2012.15.3.283.

(27) Cannon, K. A.; Park, R. U.; Boyken, S. E.; Nattermann, U.; Yi, S.; Baker, D.; King, N. P.; Yeates, T. O. Design and Structure of Two New Protein Cages Illustrate Successes and Ongoing Challenges in Protein Engineering. *Protein Sci.* **2020**, *29* (4), 919–929. https://doi.org/10.1002/pro.3802.

(28) Futreal, P. A.; Liu, Q.; Shattuck-Eidens, D.; Cochran, C.; Harshman, K.; Tavtigian, S.; Bennett, L. M.; Haugen-Strano, A.; Swensen, J.; Miki, Y.; Eddington, K.; McClure, M.; Frye, C.; Weaver-Feldhaus, J.; Ding, W.; Gholami, Z.; Söderkvist, P.; Terry, L.; Jhanwar, S.; Berchuck, A.; Iglehart, J. D.; Marks, J.; Ballinger, D. G.; Barrett, J. C.; Skolnick, M. H.; Kamb, A.; Wiseman, R. *BRCA1* Mutations in Primary Breast and Ovarian Carcinomas. *Science* **1994**, *266* (5182), 120–122. https://doi.org/10.1126/science.7939630.

(29) Miki, Y.; Swensen, J.; Shattuck-Eidens, D.; Futreal, P. A.; Harshman, K.; Tavtigian, S.; Liu, Q.; Cochran, C.; Bennett, L. M.; Ding, W.; Bell, R.; Rosenthal, J.; Hussey, C.; Tran, T.; McClure, M.; Frye, C.; Hattier, T.; Phelps, R.; Haugen-Strano, A.; Katcher, H.; Yakumo, K.; Gholami, Z.; Shaffer, D.; Stone, S.; Bayer, S.; Wray, C.; Bogden, R.; Dayananth, P.; Ward, J.; Tonin, P.; Narod, S.; Bristow, P. K.; Norris, F. H.; Helvering, L.; Morrison, P.; Rosteck, P.; Lai, M.; Barrett, J. C.; Lewis, C.; Neuhausen, S.; Cannon-Albright, L.; Goldgar, D.; Wiseman, R.; Kamb, A.; Skolnick, M. H. A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene *BRCA1*. *Science* **1994**, *266* (5182), 66–71. https://doi.org/10.1126/science.7545954.

(30) Edwards, R. A.; Lee, M. S.; Tsutakawa, S. E.; Williams, R. S.; Tainer, J. A.; Glover, J. N. M. The BARD1 C-Terminal Domain Structure and Interactions with Polyadenylation Factor CstF-50. *Biochemistry* **2008**, *47* (44), 11446–11456. https://doi.org/10.1021/bi801115g.

(31) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(32) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19* (6), 679–682. https://doi.org/10.1038/s41592-022-01488-1.

(33) Bery, N.; Legg, S.; Debreczeni, J.; Breed, J.; Embrey, K.; Stubbs, C.; Kolasinska-Zwierz, P.; Barrett, N.; Marwood, R.; Watson, J.; Tart, J.; Overman, R.; Miller, A.; Phillips, C.; Minter, R.; Rabbitts, T. H. KRAS-Specific Inhibition Using a DARPin Binding to a Site in the Allosteric Lobe. *Nat. Commun.* **2019**, *10* (1), 2607. https://doi.org/10.1038/s41467-019-10419-2.

(34) Lindborg, M.; Dubnovitsky, A.; Olesen, K.; Björkman, T.; Abrahmsén, L.; Feldwisch, J.; Härd, T. High-Affinity Binding to Staphylococcal Protein A by an Engineered Dimeric Affibody Molecule. *Protein Eng. Des. Sel. PEDS* **2013**, *26* (10), 635–644. https://doi.org/10.1093/protein/gzt038.

(35) Khodabakhsh, F.; Behdani, M.; Rami, A.; Kazemi-Lomedasht, F. Single-Domain Antibodies or Nanobodies: A Class of Next-Generation Antibodies. *Int. Rev. Immunol.* **2018**, *37* (6), 316–322. https://doi.org/10.1080/08830185.2018.1526932.

(36) Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8, 2015.

Chapter IV: Developing a combinatorial DNA-triggered bioswitch incorporating a protein nanocage

Eric J. Lee[1]

[1] UCLA Department of Chemistry and Biochemistry, Los Angeles, CA 90095

ABSTRACT

The development of novel bioactive switches is a prominent goal of protein engineering. These systems critically comprise an engineered "input domain" that is designed to respond to a specific stimulus to turn "on" or "off." One ongoing design challenge in the field is the lack of modularity in many protein-based switch systems. However, recent work has demonstrated that the integration of nucleic acids in the input domains may permit greater modularity, since the input domain can be readily redesigned to bind different targets via complementary base substitutions. An even more advanced switch system, however, would be one that integrates multiple switches, where one modular switch acts as the input for another modular switch. To this end, we incorporated a ligand-operable cage (LOC) into an existing DNA-triggered protein switch. In this work, we discuss the design, initial results, and challenges of developing such a system.

INTRODUCTION

A major goal of protein engineering is the creation of bioactive switches, a.k.a. "bioswitches," that can turn protein activity "on" or "off." To achieve this, a designed "input" component interacts with a target protein to achieve a specific response. Current methodologies for developing bioswitches generally involve incorporating an input "domain" via fusion to the target protein[1]. The input domain then effectively functions as a signal "receptor," recognizing a specific stimulus and thereby causing a desired change in the protein's activity, typically in the form of target binding, enzymatic activity, fluorescence, and/or luminescence[2–9]. However, such protein-based bioswitches require specific conformational changes and/or structural features to function, which may not be transferrable between different trigger-response systems[1]. In such

cases, readily adapting a particular protein-based biosensor to adequately respond to different stimuli may be impractical or laborious, thereby limiting modularity of protein-based switches.

Consequently, an emergent class of engineered switches is nucleotide-based biosensors. Oligonucleotides have a distinct advantage over protein-based switches because they rely solely on base-pairing for complementation. Modifying such a switch requires only altering the base-pairing nucleotides to recognize a different target, allowing for greater modularity when compared to purely protein-based bioswitches.

One recent example of a nucleotide-based bioswitch is the oligodeoxyribonucleotide (ODN)-conjugated NanoLuc BRET-beacon created by Engelen et al.[9]. In this system, a NanoLuc luciferase is conjugated to a single-stranded ODN "handle" component, which complements another ODN that possesses a stem loop region and a terminal Cy3 fluorophore (Figure 4.1). The purpose of this BRET-beacon is to turn "on" or "off" depending on the stem loop recognition of a "target" ODN trigger. In the target-unbound state, the stem loop retains its structure and BRET is permitted due to proximity of the stem loop Cy3 to the NanoLuc. In the target-bound state, the Cy3 fluorophore is displaced due to stem loop unfolding, turning BRET signal "off."

In this work, we aimed to further expand this system by coupling the bioswitch output to the "open" or "closed" state of a protein nanocage carrier. By encapsulating the target ODN trigger, we can introduce an additional condition of bioswitch activation (i.e., the opening of the nanocage). To achieve this, we used our previously designed ligand-operable cage (LOC) as the carrier. Because LOCs themselves are also theoretically modular and depend on specific ligand "input" to open, this would result in a complex, combinatorial bioswitch. That is, by coupling the output of the DNA-triggered bioswitch, which is modular, to the opening of a LOC, which itself is also modular, we can theoretically generate *combinations* of inputs that lead to a desired

output. In this work, we discuss the design and characterization of ODN-containing LOCs (ODN-LOCs), as well as the experimental challenges of developing a modular, combinatorial bioswitch.

RESULTS AND DISCUSSION

Experimental scheme of BRET-beacon reporter incorporating a protein nanocage

To create a cage-integrated oligonucleotide reporter system, we based our experiment on work by Engelen et al.[9]. This system uses bioluminescent resonance energy transfer (BRET), where differences in intermolecular distance between NanoLuc luciferase and Cy3 fluorophore drive the readout for the "on" and "off" states of the bioswitch. The interaction is based on a "handle" ODN that is conjugated to the NanoLuc and a stem loop-containing "antihandle" ODN that is conjugated to a Cy3 fluorophore (Figure 4.1A). The final component of this system is a "target" ODN trigger, which complements the stem loop region of the antihandle, causing it to "unravel" and displace the Cy3 molecule (Figure 4.1B). Therefore, the system can exist in two states: the target-unbound form, where the Cy3 is in BRET proximity to the NanoLuc, and the target-bound form, where the Cy3 is displaced due to antihandle unraveling, causing a loss of BRET. In this manner, the assay reports loss of BRET (i.e. relative increase in NanoLuc luminescence and/or relative decrease in Cy3 fluorescence) as a readout for target ODN binding.

We sought to couple this bioswitch readout assay to the opening of a ligand-operable cage. To do this, we encapsulated the target ODN by conjugating it to interior residues of T33-51-LOC (conjugated form henceforth titled "ODN-LOC") (Figure 4.1A)[10]. We had previously demonstrated that T33-51-LOC "opens" (i.e., disassembles) upon binding maltose-binding protein (MBP) via surface-fused designed ankyrin repeat proteins (DARPins). Consequently, we

hypothesized that ODN-LOC, in the absence of MBP, would remain intact and occlude the encapsulated target ODN and prevent its complementation to the BRET-beacon stem loop, thereby maintaining Cy3 fluorescence (Figure 4.2B). However, upon addition of MBP and subsequent MBP-DARPin binding, we hypothesized that ODN-LOC would disassemble and expose the encapsulated target ODN (Figure 4.2C). As a result, target ODN would then bind the BRET-beacon stem loop, causing displacement of Cy3 fluorophore, and therefore a relative decrease in Cy3 fluorescence and a relative increase in NanoLuc luminescence.

Synthesis of ODN-LOCs

LOC synthesis was performed as previously established[10]. Following expression in *E. coli*, cell lysates were processed via affinity chromatography to purify recombinant cage proteins (Figure 4.3A). Oligonucleotides were conjugated to residues in the cage interior via thiol-reactive labeling. Post-labeling reaction samples were further purified via size-exclusion chromatography (SEC). ODN-labeled NanoLuc eluted at the expected volume at approximately 17-18 mL (Figure 4.3B), and a small yield of ODN-T33-51-LOC eluted at the expected volume at approximately 13 mL (Figure 4.3C). These isolated peak fractions were used for further experimentation.

Combinatorial ODN-LOC and BRET-beacon switch assay

We then investigated whether an ODN-LOC opening event could be coupled to BRET-beacon activity. To achieve this, we performed a wavelength emission scan of wells containing NanoLuc BRET-beacon and ODN-LOCs, comparing signal changes between samples incubated with either mock or cognate ligand (MBP). Our initial findings suggested that NanoLuc

luminescence (emission maximum 460 nm) increased in wells containing MBP relative to wells containing mock, but we observed little to no fluorescence of Cy3 (emission maximum 560 nm) in any of the samples (Figure 4.4A). One possible explanation for the lack of Cy3 fluorescence is that at extremely low BRET-beacon concentrations (the assays were conducted in the picomolar range of BRET-beacon), Cy3 fluorescence may be too dim to resolve. On the other hand, NanoLuc is known to provide strong signal at such concentrations. Furthermore, these results could not be consistently replicated, as the relative increase in NanoLuc luminescence due to LOC-ligand binding was only observed in some experiments.

Despite these limitations, our findings suggested that the ligand-triggered opening of ODN-LOC and subsequent target ODN release may have affected the BRET-beacon at low levels. We supposed that in the absence of MBP, ODN-LOC would remain closed, preventing the encapsulated target ODN from interfacing with the BRET-beacon stem loop, permitting NanoLuc-Cy3 BRET (Figure 4.4B). In this case, it is possible that, in the initial "on" state, the dim Cy3 fluorophore is still acting as a NanoLuc quencher, due to molecular proximity. Conversely, in the presence of MBP, we hypothesized that the target ODN would be exposed due to cage opening, which would permit the binding of target ODN to the BRET beacon stem loop, causing loss of BRET (Figure 4.4C). Therefore, this could still explain the relative increase in NanoLuc luminescence in wells containing ODN-LOC and BRET-beacon following incubation with MBP. However, further experimental optimization is required for clearer analysis.

CONCLUSIONS

We sought to integrate a trigger-responsive protein nanocage with an existing DNA sensor to create a combinatorial, bioactive switch. This design aimed to further push the modularity and complexity of a previously characterized BRET-beacon by encapsulating its input trigger inside a protein nanocage. The nanocage, based on the previously designed T33-51-LOC, can only open in response to a specific ligand, thereby adding a conditional layer for achieving a desired switch output.

To achieve this, we modified the original T33-51-LOC by covalently labeling the interior cavity residues with the BRET-beacon trigger (i.e., the target ODN). We hypothesized that the target ODN would then only affect the BRET-beacon if ODN-LOC first "opened" in response to its own specific stimulus (in this case, MBP).

Our findings did not clearly indicate that ODN-LOC successfully delivered its target ODN cargo to the BRET-beacon. While initial results from the combinatorial ODN-LOC BRET-beacon switch assay suggested the possibility of a relative increase in NanoLuc luminescence post-MBP binding, these results were ultimately difficult to reproduce. Furthermore, the apparent lack of Cy3 fluorescence in the target ODN-free, "on" state of the BRET-beacon calls into question the cause of increase NanoLuc luminescence.

Furthermore, there are multiple experimental components that may have adversely affected the experiment. A major goal of future experiments is the optimization of the thiol-reactive labeling of LOC with target ODN cargo. Under the current experimental setup, it is difficult to quantify the exact labeling efficiency, and therefore the working concentration, of target ODN to LOC. This was because overall absorbance values at A280 (protein) versus A260 (DNA) obtained for proteins both pre- and post-labeling had very low signal over noise. Future

experiments where the target ODN is labeled with a fluorophore that does not interfere with NanoLuc-Cy3 BRET may be helpful for sample preparation. This would allow the tracking of target ODN throughout the purification process, which would be especially informative during SEC. Another method may also be to analyze the UV-fluorescence of ODN-labeled components following gel electrophoresis. This way, the resulting fluorescent signal of ODN-labeled components can be compared to a standard curve of UV-fluorescence vs known DNA concentrations, permitting the deduction of ODN-labeling efficiency.

Other potential sources of error are the concentrations of each component in the assay, such as ligand, cage, etc. Originally, the K_D between the stem loop of the active beacon and the target oligonucleotide was reported to be in the nanomolar range. However, the affinity of the cargo oligonucleotide to the stem loop could be affected by its conjugation to the protein cage cavity. Further experiments testing various concentrations of each component may be informative.

For our experiments, we only investigated one type of cage, T33-51-LOC, and one type of bioswitch. Alternative approaches, utilizing different combinations of cages and reporter systems, are theoretically possible. Options such as a protease-triggered cages may provide different avenues for design and may have different levels of successful cargo release compared to LOCs.[11,12] Other DNA-based reporter systems, such as the previously established GCN4-nLuc biosensor offers an interesting alternative to the NanoLuc BRET-beacon presented in this work.[13] In this system, the interaction between the trigger and the bioswitch is based on DNA-induced conformational changes of a protein, rather than DNA-DNA complementation. Most famously, Cas enzymes also offer exciting opportunities for synthetic biology work integrating protein nanocages.[14–17] For example, a gRNA-carrying nanocage that can open in response to a

cancer metalloprotease, is one of many potential ideas. Other useful experiments would be to investigate alternative methods for the encapsulation of target ODN molecules in designed nanocages. Strategies involving spontaneous coassembly of the ODN cargo within the cage (e.g., via charge-charge interactions) instead of covalent labeling may yield better results. However, these approaches may require the usage of larger cages, so that repulsive between negative charges of packed ODN cargo do not interfere with assembly or storage. Additional designs should be explored to answer these questions.

MATERIALS AND METHODS

Cloning, expression, and purification

The LOC and MBP sequences used in this study were described previously[10]. The sequence design of NanoLuc luciferase was previously described[9]. All gene fragments were purchased from Twist Bioscience. Plasmid sequences were verified via Sanger Sequencing (Azenta Life Sciences/Genewiz). All genes were inserted into a pET-22b+ vector (Novagen) via Gibson assembly with a NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). DNA manipulations and amplification were carried out in Escherichia coli XL2 cells (Agilent). Expression was performed using Escherichia coli BL21 (DE3) cells (New England Biolabs). Induction, harvesting, lysis, and purification procedures were performed as previously described[10].

Thiol-reactive ODN labeling of proteins

Proteins for labeling were eluted from Ni-NTA affinity columns into labeling buffer (50 mM Tris-HCl pH 7.0, 250 mM NaCl, 2% w/v glycerol, 5 mM TCEP-HCl). Proteins were labeled

with associated ODN-maleimide molecules (BioSyn) using a commercial thiol-labeling protocol (Thermo Scientific). 10-fold molar excess of ODN-maleimide (resuspended in 90% v/v DMSO/water) was added (no more than 5% v/v) to 5–15 μM protein and incubated at RT in 50 mM Tris-HCl pH 7.0, 250 mM NaCl, 2% w/v glycerol, 5 mM TCEP-HCl, and 250 mM imidazole for at least 2 h. ODN-labeled proteins were then purified over a Superose 6 Increase 10/300 GL resin column (Cytiva) in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, and 2% w/v glycerol at 4 °C, flow rate 0.3 mL/min.

ODN-LOC BRET-beacon switch assay

Prior to the assay, 5-15 μM ODN handle-labeled NanoLuc was incubated with threefold molar excess of ODN-stem-Cy3 in labeling buffer for at least 30 minutes at 25 °C.

50 nM ODN-LOC, 20 pM NanoLuc-Cy3, and either mock (labeling buffer) or 20 μM MBP (all pre-equilibrated to 25 °C) were added to assay solutions containing 50% v/v SEC buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2% w/v glycerol) and 50% v/v Nanoglo assay reagent (Promega). Assay samples were incubated at room temperature for at least 5 min. Wavelength emission scans were taken at 25 °C on a SpectraMax iD3 (Molecular Devices).

SDS-PAGE

Protein samples were mixed with SDS loading dye (0.004% Bromophenol blue, 6% glycerol, 2% sodium dodecyl sulfate, 50 mM Tris-HCl pH 6.8, 5 mM DTT) before boiling for at least 2 min. Samples were loaded onto a Mini-Protean TGX polyacrylamide gel (BioRad). Electrophoresis was performed at 200 V for 30 min. Gels were stained with Coomassie Blue and destained.

FIGURES

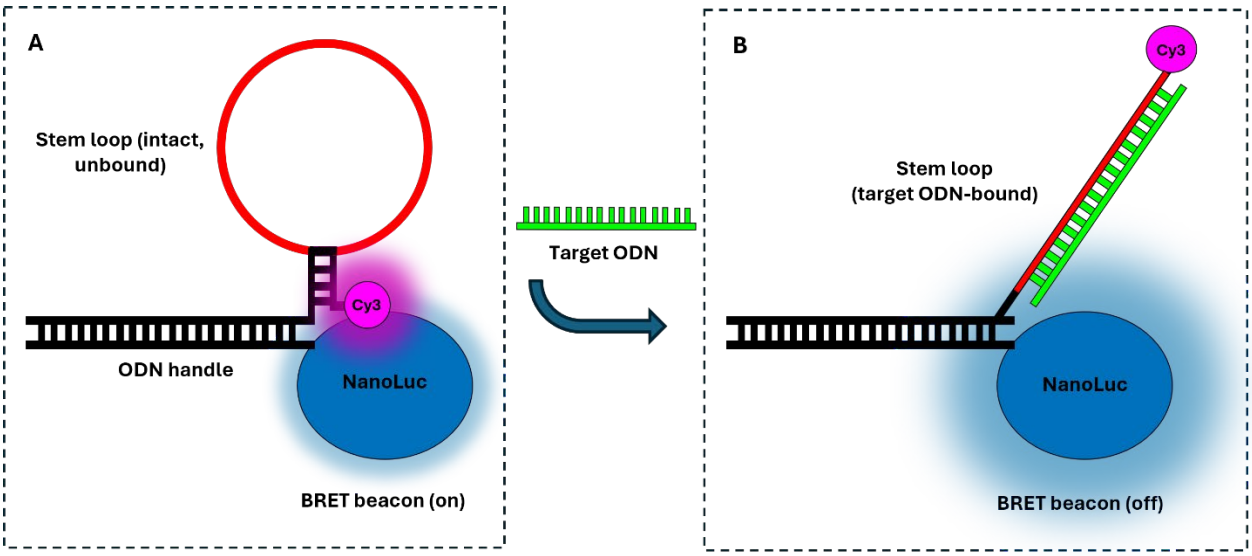

Figure 4.1. Diagram of ODN-triggered NanoLuc BRET beacon. (A) In the intact BRET-beacon state, stem loop-Cy3 (stem loop region in red, non-stem region in black, Cy3 in magenta) is base-paired with ODN handle (black), permitting BRET between NanoLuc (blue) and Cy3. (B) In the target ODN (green)-bound state, the stem loop region is unraveled, prohibiting BRET interaction between Cy3 and NanoLuc due to increased intermolecular distance. Figure redrawn from Engelen et al., 2017[9].

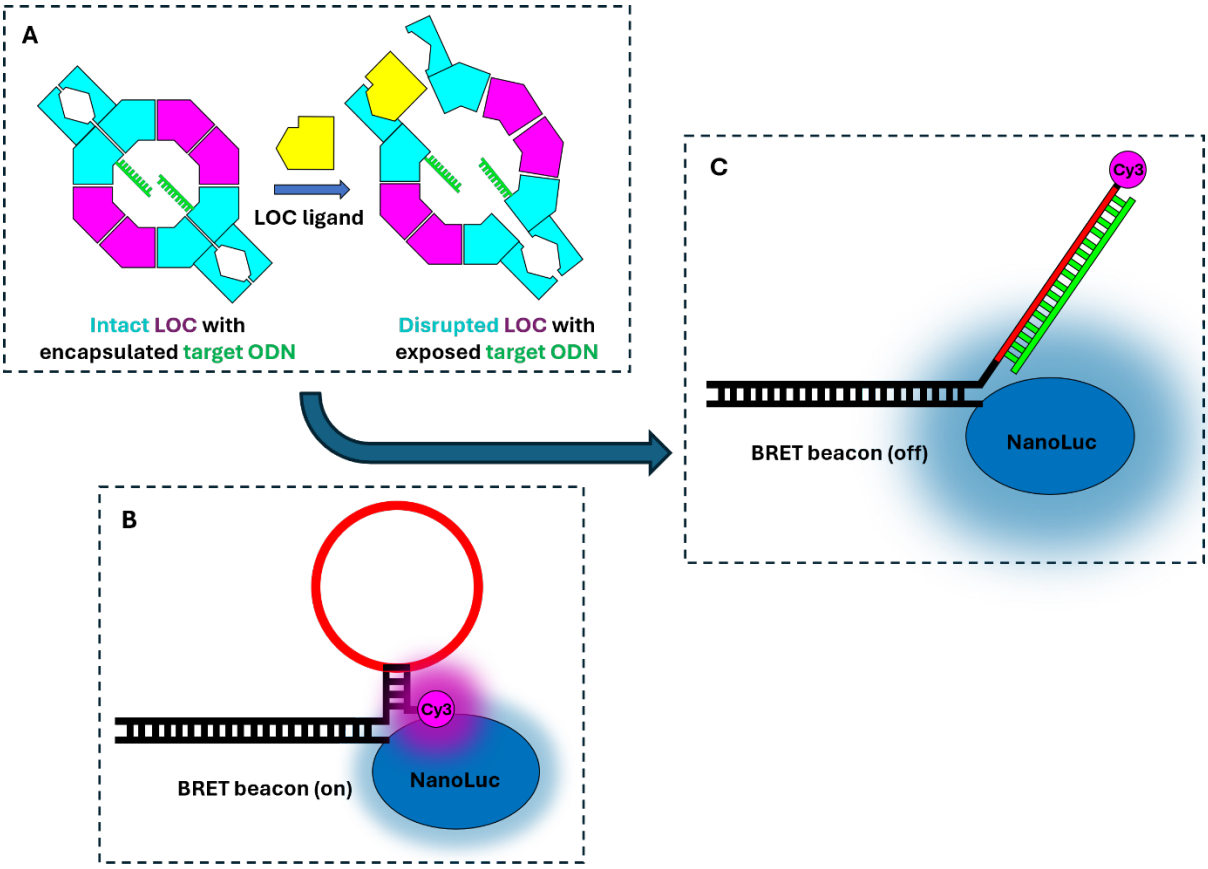

Figure 4.2. Diagram of LOC-switch assay. (A) Previously designed T33-51-LOC, modified to encapsulate target ODN (green): intact form occludes target ODN (left); ligand-bound, disrupted state exposes target ODN (right). (B) BRET-beacon is "on" in absence of disrupted ODN-LOCs. (C) BRET-beacon is "off" in presence of disrupted ODN-LOCs.

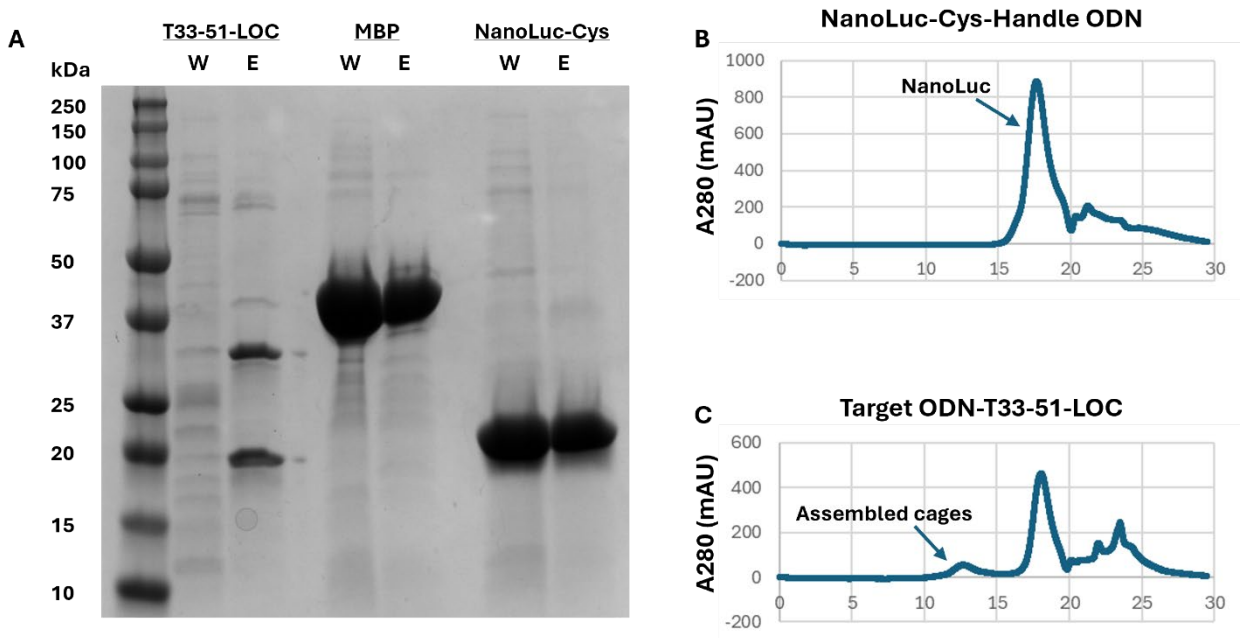

Figure 4.3. Expression and purification of LOC-switch assay components. (A) SDS-PAGE of assay proteins following affinity chromatography; T33-51-LOC, MBP, and NanoLuc wash (W) and eluate (E) fractions. (B) SEC chromatograph of NanoLuc following thiol-reactive labeling with Handle ODN. (C) SEC chromatograph of T33-51-LOC following thiol-reactive labeling with Target ODN.

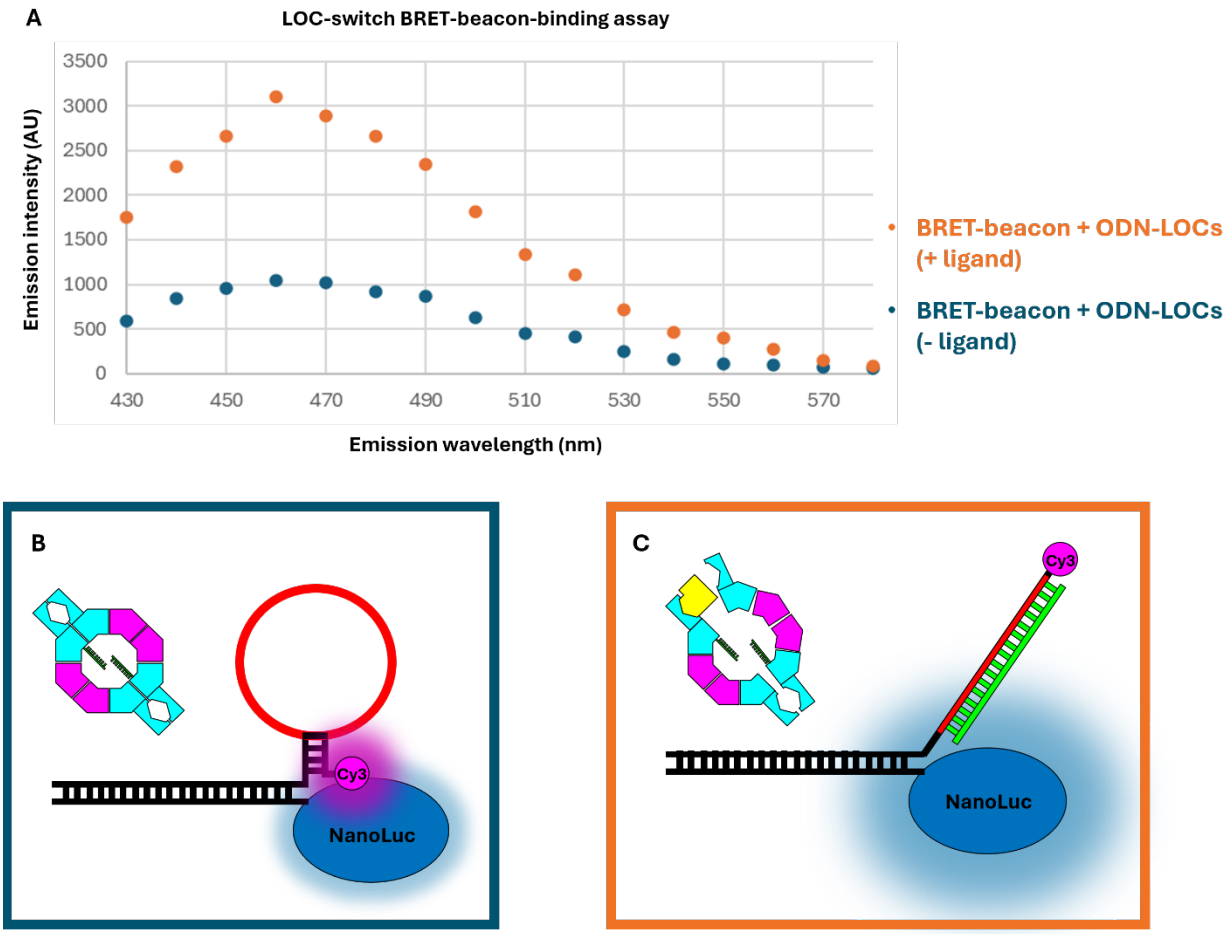

Figure 4.4. LOC-switch BRET beacon-binding assay. (A) Wavelength emission spectrum versus emission intensity: odn-T33-51-LOC + mock + NanoLuc BRET-beacon (blue trace); odn-T33-51-LOC + ligand (MBP) + NanoLuc BRET-beacon (orange trace). (B) Diagrammatic hypothesis of blue trace (no ligand, ODN-LOCs intact, target ODN cargo occluded). (C) Diagrammatic hypothesis of orange trace (ligand-incubated, ODN-LOCs disrupted, target ODN cargo exposed).

Protein	Amino acid sequence
T33-51-LOC-A	MCGGGRITTKVGDKGSTRLFGGEEVWKDDPIIEANGTLDELTSFIGE AKHYVDEEMKGILEEIQNDIYKIMGEIGSKGKIEGISEERIKWLAGLIER YSEMNKLSFVLPGGTLESAKLDVCRTIARRAERKVATVLREFGIGTLA AIYLALLSRLLFLLARVIEIEKNKL**AQ***GRKLLEAARAGQDDEVRILMAN GADVNAADNTGTTPLHLAAYSGHLEIVEVLLKHGADVDASDVFGYTP LHLAAYWGHLEIVEVLLKNGADVNAMDSDGMTPLHLAAKWGYLEIV EVLLKHGADVNAQDKFGKTAFDISIDNGNEDLAEILQKLN*
T33-51-LOC-B	MFTRRGDQGETDLANRARVGKDSPVVEVQGTIDELNSFIGYALVLS RWDDIRNDLFRIQNDLFVLGEDVSTGGKGRTVTMDMIIYLIKRSVEM KAEIGKIELFVVPGGSVESASLHMARAVSRRLERRIKAASELTEINAN VLLYANMLSNILFMHALISNKRLNIPEKIWSIHRVSLEHHHHHH
NanoLuc luciferase (C164S, G180C)	MWSHPQFEKMVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQN LGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVYPV DDHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTL WNGNKIIDERLINPDGSLLFRVTINGVTGWRLSERILAAAALELPETG CHHHHHH
His-MBP	MHHHHHHLVPRGSGSGSMKTEEGKLVIWINGDKGYNGLAEVGKKF EKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQS GLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLL PNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAF KYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEE ELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQT VDEALKDAQTGSGGTPGRPAAKLN

Table 4.S1. Amino acid sequences of LOC and BRET-beacon assay proteins. Underlined amino acids indicate terminal cysteine for thiol-reactive labeling. Underlined and bolded amino acids indicate helical linker. Underlined and italicized amino acids indicate DARPin.

ODN	Nucleotide sequence
Target ODN	[maleimide]**CAGCACATTCAA**
NanoLuc handle	[maleimide]*GTGATGTAGGTGGTAGAGGAA*
Antihandle stem loop	*TTCCTCTACCACCTACATCAC*ACAGC**TGGAGACGTAGGGTATTG AATGT**GCTGT[Cy3 fluorophore]

Table 4.S2. Nucleotide sequences of BRET-beacon assay ODN components. All sequences read 5' to 3' end. Bracketed regions indicate chemical modifications. Underlined and bolded

bases indicate stem loop or stem loop-complementary region. Underlined and italicized bases indicate handle or handle-complementary region. Sequences taken from Engelen et al., 2017.[9]

REFERENCES

(1) Sekhon, H.; Ha, J.-H.; Loh, S. N. Engineering Protein and DNA Tools for Creating DNA-Dependent Protein Switches. In *Methods in Enzymology*; Elsevier, 2022; Vol. 675, pp 1–32. https://doi.org/10.1016/bs.mie.2022.07.002.

(2) Karchin, J. M.; Ha, J.-H.; Namitz, K. E.; Cosgrove, M. S.; Loh, S. N. Small Molecule-Induced Domain Swapping as a Mechanism for Controlling Protein Function and Assembly. *Sci. Rep.* **2017**, *7* (1), 44388. https://doi.org/10.1038/srep44388.

(3) Nicholes, N.; Date, A.; Beaujean, P.; Hauk, P.; Kanwar, M.; Ostermeier, M. Modular Protein Switches Derived from Antibody Mimetic Proteins. *Protein Eng. Des. Sel.* **2016**, *29* (2), 77–85. https://doi.org/10.1093/protein/gzv062.

(4) Wu, Y. I.; Frey, D.; Lungu, O. I.; Jaehrig, A.; Schlichting, I.; Kuhlman, B.; Hahn, K. M. A Genetically Encoded Photoactivatable Rac Controls the Motility of Living Cells. *Nature* **2009**, *461* (7260), 104–108. https://doi.org/10.1038/nature08241.

(5) Carrasco-López, C.; Zhao, E. M.; Gil, A. A.; Alam, N.; Toettcher, J. E.; Avalos, J. L. Development of Light-Responsive Protein Binding in the Monobody Non-Immunoglobulin Scaffold. *Nat. Commun.* **2020**, *11* (1), 4045. https://doi.org/10.1038/s41467-020-17837-7.

(6) Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. Circular Permutation and Receptor Insertion within Green Fluorescent Proteins. *Proc. Natl. Acad. Sci.* **1999**, *96* (20), 11241–11246. https://doi.org/10.1073/pnas.96.20.11241.

(7) Nasu, Y.; Shen, Y.; Kramer, L.; Campbell, R. E. Structure- and Mechanism-Guided Design of Single Fluorescent Protein-Based Biosensors. *Nat. Chem. Biol.* **2021**, *17* (5), 509–518. https://doi.org/10.1038/s41589-020-00718-x.

(8) Chang, D.; Kim, K. T.; Lindberg, E.; Winssinger, N. Smartphone DNA or RNA Sensing Using Semisynthetic Luciferase-Based Logic Device. *ACS Sens.* **2020**, *5* (3), 807–813. https://doi.org/10.1021/acssensors.9b02454.

(9) Engelen, W.; Van De Wiel, K. M.; Meijer, L. H. H.; Saha, B.; Merkx, M. Nucleic Acid Detection Using BRET-Beacons Based on Bioluminescent Protein–DNA Hybrids. *Chem. Commun.* **2017**, *53* (19), 2862–2865. https://doi.org/10.1039/C6CC10032E.

(10) Lee, E. J.; Gladkov, N.; Miller, J. E.; Yeates, T. O. Design of Ligand-Operable Protein-Cages That Open Upon Specific Protein Binding. *ACS Synth. Biol.* **2024**, *13* (1), 157–167. https://doi.org/10.1021/acssynbio.3c00383.

(11) Kang, Y. J.; Park, D. C.; Shin, H.-H.; Park, J.; Kang, S. Incorporation of Thrombin Cleavage Peptide into a Protein Cage for Constructing a Protease-Responsive Multifunctional Delivery Nanoplatform. *Biomacromolecules* **2012**, *13* (12), 4057–4064. https://doi.org/10.1021/bm301339s.

(12) Miller, J. E.; Srinivasan, Y.; Dharmaraj, N. P.; Liu, A.; Nguyen, P. L.; Taylor, S. D.; Yeates, T. O. Designing Protease-Triggered Protein Cages. *J. Am. Chem. Soc.* **2022**, *144* (28), 12681–12689. https://doi.org/10.1021/jacs.2c02165.

(13) Sekhon, H.; Loh, S. N. Engineering Protein Activity into Off-the-Shelf DNA Devices. *Cell Rep. Methods* **2022**, *2* (4), 100202. https://doi.org/10.1016/j.crmeth.2022.100202.

(14) Nalefski, E. A.; Patel, N.; Leung, P. J. Y.; Islam, Z.; Kooistra, R. M.; Parikh, I.; Marion, E.; Knott, G. J.; Doudna, J. A.; Le Ny, A.-L. M.; Madan, D. Kinetic Analysis of Cas12a and Cas13a RNA-Guided Nucleases for Development of Improved CRISPR-Based Diagnostics. *iScience* **2021**, *24* (9), 102996. https://doi.org/10.1016/j.isci.2021.102996.

(15) O'Connell, M. R. Molecular Mechanisms of RNA Targeting by Cas13-Containing Type VI CRISPR–Cas Systems. *J. Mol. Biol.* **2019**, *431* (1), 66–87. https://doi.org/10.1016/j.jmb.2018.06.029.

(16) Das, A.; Goswami, H. N.; Whyms, C. T.; Sridhara, S.; Li, H. Structural Principles of CRISPR-Cas Enzymes Used in Nucleic Acid Detection. *J. Struct. Biol.* **2022**, *214* (1), 107838. https://doi.org/10.1016/j.jsb.2022.107838.

(17) Kaminski, M. M.; Abudayyeh, O. O.; Gootenberg, J. S.; Zhang, F.; Collins, J. J. CRISPR-Based Diagnostics. *Nat. Biomed. Eng.* **2021**, *5* (7), 643–656. https://doi.org/10.1038/s41551-021-00760-7.

Chapter V: Summary and conclusions

Eric J. Lee[1]

[1] UCLA Department of Chemistry and Biochemistry, Los Angeles, CA 90095

Designed protein cages are an emergent technology that has exciting applications for materials, therapeutics, and synthetic biology. These particles possess several features that make them attractive as a platform for synthetic biology applications. Firstly, they are generally monodisperse, stable, and biocompatible. Furthermore, they comprise an external surface that can be decorated with adaptors and/or antigens, and an interior volume that can store moieties of interest. Nanocages are also readily modified post-assembly, or genetically fused with various tags or biomarkers per intended function.

Already, protein nanocages have demonstrated their potential as molecular containers and delivery vehicles. Prior works involving ferritin, VLPs, and other notable nonviral cages have demonstrated the ability for these cages to not only encapsulate non-native cargo, but also disassemble and release their cargo under certain environmental conditions. Advent synthetic biology approaches have led to further engineering of these assemblies, allowing them to respond to various, unique stimuli such as light, pressure, etc.

Despite these achievements, there has been a general lack of specific, target-based systems of nanocage disassembly and cargo delivery. Some recent works have integrated proteolysis sites into a designed nanocage, which enabled more specific conditions of cage opening. Overall, a nanocage-based delivery platform that can be triggered to open with a target *molecule*-specific mechanism is highly desirable. Such a platform, if also made modular, would greatly impact the field.

To address this challenge, we created a new class of designed nanocages, the ligand-operable cage (LOC). LOCs comprise a self-assembling nanocage core and surface-fused, target-specific adaptors. These adaptors are further oriented such that binding of their specific target is sterically incompatible with the retention of the overall cage assembly. Therefore, we

hypothesized that high concentrations of target ligand would drive disassembly of the cage. We demonstrated two test cases using a natural and a designed cage as a core and fused DARPins as adaptors. The first was based on sulfur oxygenase-reductase (SOR) fused to designed ankyrin repeat proteins (DARPins) that bound sfGFP. The second was based on the previously designed T33-51 two-component cage, fused to DARPins that bound MBP. The helical termini of these cage proteins and DARPins permitted facile fusion and linker modeling, resulting in rigid assemblies. Our findings showed that the engineered SOR-LOC and T33-51-LOC proteins successfully assembled as cages following expression and purification. We demonstrated successful ligand-triggered disassembly using SEC, nsEM, native-PAGE, and DLS. Furthermore, we developed two readout assays to detect successful ligand-triggered exposure of a cargo molecule: one based on the fluorescence unquenching of an internal fluorophore, and another based on complementation of a cage-encapsulated peptide to an external luciferase.

Next, we sought to test the viability and modularity of LOCs in targeting cancer proteins. New LOC variants were generated by modifying the original T33-51-LOC by replacing the MBP-binding DARPin with one that binds KRAS, BARD1, or BRCA1. We modeled these new designs using AlphaFold and characterized their assembly and ligand-binding via SDS-PAGE, SEC, nsEM, and DLS. Our preliminary findings suggested that αKRAS-LOC and αBRCA1-LOC may be able to bind their target and disassemble. Overall, further investigations, especially with improved LOC variant expression and assembly yield, may be highly informative.

Finally, we attempted to design and characterize a combinatorial DNA-triggered bioswitch incorporating a protein nanocage. This system comprised the previously designed T33-51-LOC and a previously published DNA-triggered BRET-beacon. We aimed to integrate these systems by encapsulating the BRET-beacon trigger inside the LOC, which itself can be triggered

to open. This approach permits the combination of two otherwise independent, target-specific systems, creating a complex, multiconditional bioswitch that is also highly modular. From our initial wavelength emission scan experiments, we found that the binding of ODN-LOC to its triggering ligand may have caused successful target ODN delivery. We observed a shift in the BRET-beacon reporter readout, suggesting that the released target ODN was able to successfully bind the BRET-beacon following the ligand-triggered disassembly of ODN-LOC. Further experimentation, in conjunction with improved yield of assay components, would greatly benefit this work.

In review, these findings demonstrate the potential of designed protein nanocages as a platform for highly specific and modular synthetic biology applications. Despite ongoing experimental challenges, this technology may have significant implications for future protein engineering studies. Furthermore, the work discussed in the dissertation ultimately encompasses a miniscule number of designs relative to the total, vast design space for these assemblies. Myriad alternatives await exploration, and, considering the recent pace of upcoming protein design tools, we may see novel synthetic cages emerge as a new standard for biotechnology in the near future.