

DISSERTATION

POROUS PROTEIN-DNA CO-CRYSTALS FOR NUCLEIC ACID CAPTURE, STORAGE,
AND RELEASE

Submitted by

Caroline K. Slaughter

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2026

Doctoral Committee:

Advisor: Christopher D. Snow

Jeffrey Wilusz

P. Shing Ho

Chaoping Chen

Copyright by Caroline Kathleen Slaughter 2026

All Rights Reserved

ABSTRACT

POROUS PROTEIN-DNA CO-CRYSTALS FOR NUCLEIC ACID CAPTURE, STORAGE, AND RELEASE

Porous biomolecular crystals provide an emerging platform for nanoscale organization, molecular capture, and controlled release, but a system that combines modularity, tunability, and high-resolution structural precision has yet to be elucidated. Biomolecular nanoparticles made of pure protein and pure DNA have been explored for use in biomedical, structural, and nanotechnology applications. By combining the structural rigidity of protein crystals and the modularity of DNA crystals, we have developed a class of biomolecular co-crystals that fills gaps currently present in the biomolecular materials field. We are interested in how protein-DNA co-crystals can be used as scaffolds, hosting guest molecules by soaking them in after the protein-DNA co-crystals have formed. Unique to our platform among crystalline biomaterials capable of spatial organization of guest molecules is the ability to diffract X-rays at a resolution sub 3 Angstroms. This dissertation presents a comprehensive exploration of these high-resolution protein–DNA co-crystals as programmable crystalline nanoparticles engineered for selective nucleic acid sequestration, stabilization, and triggerable release.

We review the landscape of biomolecular crystalline nanoparticles in biomedical applications, highlighting how their periodicity and defined pore environments enable drug loading, biosensing, and molecular templating. To enable our protein-DNA co-crystal to persist in the stringent environments necessary to be of service in these fields, we describe efforts to

enhance the stability of protein–DNA co-crystals by optimizing crosslinking at DNA–DNA junctions using both chemical crosslinkers and enzymatic ligation strategies. These approaches significantly increase crystal robustness while preserving porosity and modular programmability. To demonstrate the utility of having programmable DNA elements that take advantage of scaffold protein rigidity, we next show the chemical and structural diversity of the system by incorporating non-canonical nucleobases, alternative backbone chemistries, and engineered secondary structures. We then leverage these design principles to show nucleic-acid capture from the environment, storage within the crystal pores, and specific release mechanisms using toehold-mediated strand displacement, showcasing highly specific, trigger-dependent unloading of target cargo from protein-DNA co-crystals. For practical use of our platform in biomedical applications, we adapt microcrystallization methods to enable scalable production and consistency in nanoparticle batches. To exemplify quantifiable guest molecule loading within our co-crystals, we present a novel cuvette-based uptake fluorescence assay. Finally, we outline future directions for this work in structural, biomedical, and nanotechnology fields.

Together, this work establishes protein–DNA co-crystals as a modular, engineerable material system capable of selective nucleic acid capture, stable storage, and programmable release. This opens new avenues for responsive biomaterials, biosensing, and therapeutic delivery.

ACKNOWLEDGEMENTS

I want to thank my lab mates in the Snow lab for their mentorship, friendship, and contributions to my growth as a researcher. To my cohort in CMB, these years have flown by, and I couldn't have asked for a better group of folks to do this journey alongside. I know we will never forget the wonderful year we got to spend with our sweet, brilliant Rojina. I want to thank my committee members (Dr. Shing Ho, Dr. Jeff Wilusz, Dr. Chaoping Chen) for their exceptional attention to detail in supporting my work and the thoughtfulness with which they have guided me. To the undergraduates (MURALS, SURE, and independent volunteers) who contributed their efforts and time over the years I am especially grateful, you all have made me a better mentor and leader. To Dr. Phil Lukeman, I say thank you for the beers. To my collaborators nationwide (Jay Nix at ALS, Kay Perry at APS, Dr. Brian Fuglestad and Kelly Petersen at VCU) I thank you for your shared resources and communication. To Dr. Abby Orun, I thank you endlessly for setting me up for success in this project and paving the way.

To the co-crystal team – you are the best teammates I could've asked for, and I can't overstate what your camaraderie and efforts have meant to me. To Ethan and Emma, you two are excellent scientists and even better friends.

To Dr. Christopher Snow, thank you for your endless creativity and believe in this project. Thank you for empowering me to explore my own hypotheses and path in my doctoral work. You are an inspiration to me.

DEDICATION

For Adam, Kathy, and Grady – my everything

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iv
DEDICATION.....	v
LIST OF TABLES.....	ix
LIST OF FIGURES	x
CHAPTER 1: BIOMOLECULAR CRYSTALS AS FUNCTIONAL NANOPARTICLES	1
1.1 Abstract.....	1
1.2 Introduction.....	1
1.3 Engineered Protein Crystal Nanoparticles.....	4
1.4 DNA-based Crystalline Nanoparticles.....	9
1.5 Hybrid Crystalline Biomaterials	11
1.6 Future Directions	13
1.7 Specific Aims.....	14
Chapter One Bibliography	16
CHAPTER 2: CRYSTALLINE BIOMATERIALS FOR SITE SPECIFIC ORGANIZATION OF MODIFIED DNA	24
2.1 Introduction.....	24
2.2 Results and Discussion	28
2.2.1 Engineered Bases in the Expansion	29
2.2.1.1 Common DNA Damage Motifs	29
2.2.1.2 Fluorophore Tags.....	32
2.2.2 Backbone Modification and RNA:DNA Hybrid	35
2.2.3 ssDNA Tail Motifs.....	39
2.3 Conclusions.....	45
2.4 Materials and Methods.....	48
2.4.1 Protein Cloning, Expression, Purification	48
2.4.2 Duplex Annealing and Co-crystal Complexing.....	49
2.4.3 Crystallization.....	49
2.4.4 Crosslinking.....	50
2.4.5 Confocal Microscopy.....	50
2.4.6 X-ray Diffraction Collection and Refinement	51
2.4.7 Quantum Chemistry.....	51
Chapter Two Bibliography	54
CHAPTER 3: ENZYMATIC AND CHEMICAL LIGATION FOR STABILIZATION ACROSS DNA-DNA JUNCTIONS	61
3.1 Introduction.....	61
3.1.1 Motivation for Post-Assembly Ligation in Protein-DNA Co-crystals	61
3.1.2 Chapter Objectives.....	65
3.2 Ligation in CC1+10 Expanded Crystals	66
3.2.1 Experimental Design for EDC Ligation Assays	66
3.2.2 Effect of Junction Architecture on EDC Ligation Yield	68

3.2.3 Effect of Magnesium on EDC Ligation in CC1+10	71
3.2.4 Splint Ligation Controls: AmeLig vs EDC Yields	73
3.2.5 Attempted Enzymatic Ligation in CC1+10	75
3.3 Ligation in CC1+21 Crystals	77
3.3.1 Design and Growth of CC1+21 Crystals	77
3.3.2 EDC Chemical Ligation in CC1+21	77
3.3.3 Enzymatic Ligation	80
3.3.4 Restriction Enzyme Digestion	81
3.4 Discussion and Conclusions	84
3.4.1 Chemical Ligation in CC1+10	84
3.4.2 Enzymatic Ligation in CC1+10 and CC1+21	85
3.4.3 Restriction Enzyme Digestion of CC1+21	85
3.4.4 Future Directions	86
3.5 Materials and Methods	88
3.5.1 Protein Cloning, Expression, Purification	88
3.5.2 Duplex Annealing and Co-crystal Complexing	88
3.5.3 Crystallization	89
3.5.4 EDC Crosslinking	89
3.5.5 AmeLig Crosslinking	90
3.5.6 HhaI Digestion	90
3.5.7 DNA Gel Electrophoresis and Densitometry	91
3.5.8 NHS-fluorescein Conjugation	91
3.5.9 Confocal Microscopy	92
3.5.10 Guest Loading	92
3.5.11 X-ray Diffraction Collection and Refinement	93
Chapter Three Bibliography	94
CHAPTER 4: PROTEIN-DNA CO-CRYSTALS FOR CAPTURE, STORAGE, AND RELEASE OF NUCLEIC ACIDS	97
4.1 Introduction	97
4.1.1 Porous Protein-DNA Co-crystals as Programmable Host Materials	97
4.1.2 From DNA Crystal Devices to Functional Guest Handling in Protein-DNA Co-crystals	98
4.1.3 Motivation for Therapeutic RNA Capture and Release	100
4.2 Design of Nucleic Acid Capture Devices in Porous Protein-DNA Co-crystals	101
4.2.1 Architecture of the Capture-Enabled Co-crystal	101
4.2.2 DNA Capture Construct	103
4.2.3 RNA/miR-34a Capture Construct	106
4.3 Structural Analysis of Capture-Enabled Co-crystals by X-ray Diffraction	109
4.3.1 Benefits of XRD Studies	109
4.3.2 Tailed Crystals without Nucleic Acid Guests	109
4.3.3 Tailed Crystals with Nucleic Acid Guests	110
4.4 Discussions and Conclusions	111
4.4.1 Porous Co-crystals as Programmable Nucleic Acid Reservoirs	111
4.4.2 miR-34a Capture and Release as a Step Toward Functional Cargo	112
4.4.3 Limits and Unresolved Questions	112
4.6 Future Directions	113

4.7 Materials and Methods.....	114
4.7.1 Protein Cloning, Expression, Purification	114
4.7.2 Duplex Annealing and Co-crystal Complexing.....	114
4.7.3 Crystallization.....	115
4.7.4 Crosslinking.....	115
4.7.5 Confocal Microscopy.....	116
4.7.6 Guest Loading.....	116
4.7.7 X-ray Diffraction Collection and Refinement	117
Chapter Four Bibliography	118
CHAPTER 5: BATCH MICROCRYSTALLIZATION.....	123
5.1 Introduction.....	123
5.1.1 Motivation for Switching to Microcrystallization	123
5.1.2 Role of Chemical Crosslinking in Microcrystal Stabilization	125
5.2 Development of Crystal Growth and Crosslinking Conditions	126
5.2.1 Batch Microcrystallization Conditions	126
5.2.2 Chemical Crosslinking of Batch-Produced Microcrystals.....	128
5.3 Batch Characterization and Stability Testing	129
5.3.1 Batch Characterization.....	129
5.3.2 Stability of Microcrystals Under Biomimetic Challenge Conditions	133
5.4 Discussion and Conclusions	137
5.4.1 Translating Co-crystals to Materials with Biomedical Potential	137
5.4.2 Barriers Remaining to Wide Adaption of Co-crystals.....	138
Chapter Five Bibliography.....	139
CHAPTER 6: FUTURE DIRECTIONS FOR PROTEIN-DNA CO-CRYSTALS AS FUNCTIONAL NANOPARTICLES	141
APPENDICES	148
Appendix I. Supplemental Information for Chapter 2	148
Appendix II. Supplemental Information for Chapter 3.....	188
Appendix III. Supplemental Information for Chapter 4	213

LIST OF TABLES

Table 3.1: Distribution of DNA block sizes as a function of DNA-DNA junction architecture...	70
Table 3.2: Distribution of DNA block sizes in differing buffer conditions.....	73
Table 3.3: Distribution of DNA block sizes in CC1+21 crystals	80

LIST OF FIGURES

Figure 1.1: Classifications of nanoparticles based on components	2
Figure 1.2: Design schemes in protein-based materials.....	5
Figure 1.3: Methods of protein crosslinking and covalent bond installation.....	6
Figure 1.4: DNA nanostructure designs.....	10
Figure 2.1: Lattice topology of porous CC1+10 and interpenetrating CC1+10 (ipCC1)	27
Figure 2.2: Modifications described herein, shown added to symmetric and asymmetric expansions of CC1	28
Figure 2.3: CC1+10 with fluorescein appended to each building block	32
Figure 2.4: Putative fluorescein stacking.....	34
Figure 2.5 RNA:DNA hybrid crystals	38
Figure 2.6: Fluorescent oligo loading in end-tailed CC1+10	41
Figure 2.7: Post crystallization loading of oligo guest and subsequent structural observation	44
Figure 3.1: Ligation Chemistries	62
Figure 3.2: Constructs of CC1 with varied junction architecture	66
Figure 3.3: TBE-Urea denaturing gel electrophoresis of chemically ligated CC1+10.....	69
Figure 3.4: Buffer condition comparison.....	72
Figure 3.5: Splint ligation assay.....	74
Figure 3.6: Enzymatic ligation on CC1+10 duplexes in free solution and in crystallo	76
Figure 3.7: CC1+21 with 2 sticky base overhangs and high GC content at the junctions	77
Figure 3.8: CC1+21 ligation and cleavage	79
Figure 3.9: Crystal stability in water based cutting buffer.....	83
Figure 4.1: DNA device cycle in protein-DNA co-crystals.....	99
Figure 4.2: Strand displacement mechanisms in the DNA capture device.....	100
Figure 4.3: Hypothetical model of centrally placed 15-nt tail	103
Figure 4.4: TMSD from Strand 1 to Strand 2	105
Figure 4.5: TMSD to complete cycle.....	106
Figure 4.6: Loading and triggered release of dsRNA miR-34a	107
Figure 4.7: Capture and release of mir-34a	108
Figure 5.1: Phase diagram.....	123
Figure 5.2: A field of CC1 microcrystals achieved using microbatch under oil methods.....	127
Figure 5.3: Batch characterization using dynamic light scattering.....	131
Figure 5.4: SEM image of a protein-DNA microcrystal.....	133
Figure 5.5: Fluorescein-tagged microcrystals grown via microbatch under oil	134
Figure 5.6: Crosslinked vs non crosslinked crystals in pH 2 solution	135
Figure 5.7: Crosslinked vs non crosslinked crystals in bovine blood serum	137
Figure 6.1: Poly-T binding cleft of RepE54 binds specifically	144

CHAPTER 1: BIOMOLECULAR CRYSTALS AS FUNCTIONAL NANOPARTICLES

1.1 Abstract

Structural heterogeneity remains a fundamental limitation of many nanoparticle platforms used in nanomedicine, constraining the extent to which structure-function relationships can be rationally defined and engineered. Crystalline biomaterials offer a complementary paradigm, in which periodic molecular organization of chemically homogenous monomers enables direct structural determination, precise control over nanoscale architecture, and predictable functional behavior. This chapter examines crystalline and hybrid biomolecular nanoparticles as a design framework for precision nanomedicine, highlighting how periodicity, programmability, and modularity can be leveraged to improve cargo stabilization, organization, and delivery. By integrating principles from structural biology, biomolecular engineering, and computational design, crystalline protein, DNA, and hybrid protein-DNA assemblies emerge as a versatile class of nanomaterials that bridge atomic-level structural definition with functional adaptability. Collectively, these systems point toward new opportunities for structure-guided nanoparticle design and the development of therapeutic and diagnostic platforms.

1.2 Introduction

Nanomedicine is a rapidly growing field that has a central theme of the delivery of therapeutic agents at the nanoscale (1-3). Most nanoparticle delivery systems consist of assembled components for which size, structure and homogeneity can only be inferred indirectly, *e.g.* via dynamic light scattering (DLS), zeta potential, and encapsulation efficiency assays (4,5). While rational design and bulk characterization assays can provide some control of and information on how these particles are assembled, crystalline biomaterials are uniquely able to be directly observed and defined with atomistic precision at the structural level due to their

periodicity and homogeneity. This in-depth material characterization allows more control when engineering a nanoparticle for biomedical applications, and precision in calculating payload per particle.

Nanoparticles have distinct advantages over free biologics. Nanoparticles can confer protection from degradation within biological solutions, tunability in biodistribution, and the ability to overcome biological barriers for intracellular delivery (6). Nanoparticles in drug delivery are generally sorted into three categories based upon their component building blocks, namely organic (*e.g.*, lipid, protein, or DNA-based materials), carbon, and inorganic (*e.g.* metal-based) particles (Fig. 1.1) (4, 7).

Classification of Nanoparticles

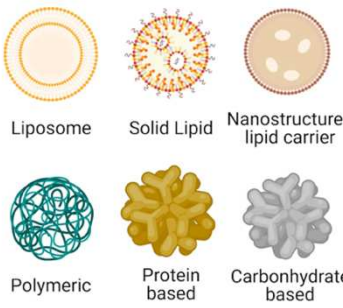
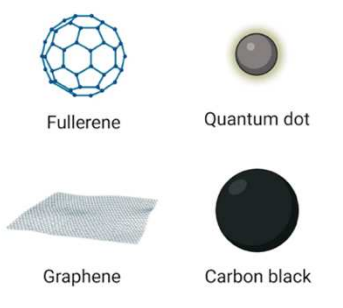
Organic Nanoparticles	Inorganic Nanoparticles	Carbon-Based Nanoparticles
 Liposome Solid Lipid Nanostructured lipid carrier Polymeric Protein based Carbohydrate based	 Silica Titanium Oxide Zinc Oxide / Zirconium dioxide Iron oxide Gold Silver	 Fullerene Quantum dot Graphene Carbon black
Advantages • Biocompatibility and biodegradability • Flexibility of design • Surface modification • Functionalization • Controlled release • Response behaviour Disadvantages • Complex synthesis • Limited Thermal and Chemical Resistance • Regulatory and Safety Concerns	Advantages • Unique electrical, magnetic, optical properties • Variability in size, structure, geometry • Superior Mechanical Properties • Catalytic Activity Disadvantages • Agglomeration and Stability Issues • Complex Synthesis and High Cost • Environmental Concerns • Toxicity problems	Advantages • Exceptional Mechanical Properties • Thermal Conductivity • Chemical Stability • Surface Functionalization • Biocompatibility and environmental remediation Disadvantages • Potential Toxicity and Biocompatibility Concerns • Aggregation and Dispersion Challenges • Complex and Costly Production

Figure 1.1: Classifications of nanoparticles based on components. Organic nanoparticles are made of biomolecules, inorganic nanoparticles are inorganic atoms covalently bound, and carbon-based nanoparticles are purely made of carbon. From “A Comprehensive Review of Nanoparticles: From Classification to Application and Toxicity.” Eker, F., *et al.* 2024, *Molecules*, 29(15), Figure 1 (<https://doi.org/10.3390/molecules29153482>) CC BY (7).

As the most analogous to our protein-DNA co-crystal-based platform, this review chapter will focus on organic nanoparticles, with some mention of organic components functionalizing

metal-based inorganic nanoparticles (metal organic frameworks, or MOFs) as their porous crystalline frameworks are similar to our porous protein-DNA platform. Of the organic nanoparticles, lipid nanoparticles (LNPs), exosomes (cell-secreted extracellular vesicles), and related lipid-based nanocarriers are generally not amenable to direct atomistic structure determination as intact particles by diffraction-based or EM-based methods, because they are typically heterogeneous, dynamic, and non-periodic. Instead, these platforms more often provide ensemble-averaged structural information or nanoscale morphological features (8). While these nanoparticles are highly effective and are the predominately prevalent nanoparticle in biomedical applications, their amorphous nature limits precise structural definition at the single particle level.

In contrast, crystalline biomaterials exhibit periodicity in their structures. Crystals with biomedical functions are found in living systems, keeping biomolecules stable inside and outside of cells (9). Most notably, insulin forms crystalline granules stabilized by Zn^{2+} within pancreatic β -cells (10). These granules protect against proteolytic degradation, allowing long term storage in cells. Similarly, eosinophils are white blood cells that store eosinophil major basic protein (EMBP) in crystalloid assemblies within secretory granules (11). In bacterial systems, a prominent example of *in vivo* crystallization is the Cry3A protein, that forms parasporal crystals during sporulation and functions as an insecticidal agent (12). Notably, these intracellular crystals can achieve a degree of order sufficient for structural interrogation, and Cry3A crystals grown directly within bacterial cells have been resolved to 2.8 Å using X-ray free-electron laser methods (13). This example illustrates that biological systems can exploit crystallization not merely for storage, but also to stabilize functional protein cargo in a densely packed and highly protected state. A related principle is seen in viral protein. Virus capsids are highly ordered, self-

assembled protein architectures whose central biological functions are to package, protect, and deliver nucleic acid cargo, illustrating how repetitive protein organization can stabilize fragile payloads (14, 15). In some viral systems, this protective strategy is taken even further through crystalline occlusion bodies, such as baculoviral polyhedra, in which polyhedrin proteins assemble into exceptionally robust lattices that shield embedded virions in the external environment while preserving infectivity until transmission to a new host (16, 17). Ferritin provides a complementary example of a highly ordered proteinaceous assembly that packages and protects a biologically important cargo, in this case iron, within a symmetric nanoscale shell capable of regulated disassembly and reassembly (18–20). While ferritin is not a crystal in the classical extended-lattice sense, it demonstrates that precise, repetitive supramolecular organization can endow biomolecular assemblies with stability, encapsulation capacity, and environmentally responsive behavior. Taken together, these natural systems underscore that ordered biomolecular assembly can be leveraged to protect fragile cargo, preserve function, and regulate access to encapsulated components. We now turn to how this principle of crystallinity, and more broadly periodic nanoscale organization, has been adapted for the engineering of organic nanomaterials.

1.3 Engineered Protein Crystal Nanoparticles

Proteins are attractive building blocks for engineered crystalline nanoparticles because they can present chemically diverse surfaces, form specific intermolecular interfaces, and be modified genetically or chemically to install new functions (21). In addition, many protein-based materials can be produced using mature recombinant expression and purification workflows, which makes them more tractable than some alternative nanomaterial classes for scale-up, purity control, and process development, although manufacturability remains strongly protein- and platform-

dependent (22, 23). Protein-based biomaterials have therefore been widely explored for therapeutic and diagnostic applications, and recent advances in computational and machine-learning-guided protein design have expanded access to de novo protein nanoparticles and other ordered protein assemblies (Fig. 1.2) (24, 25).

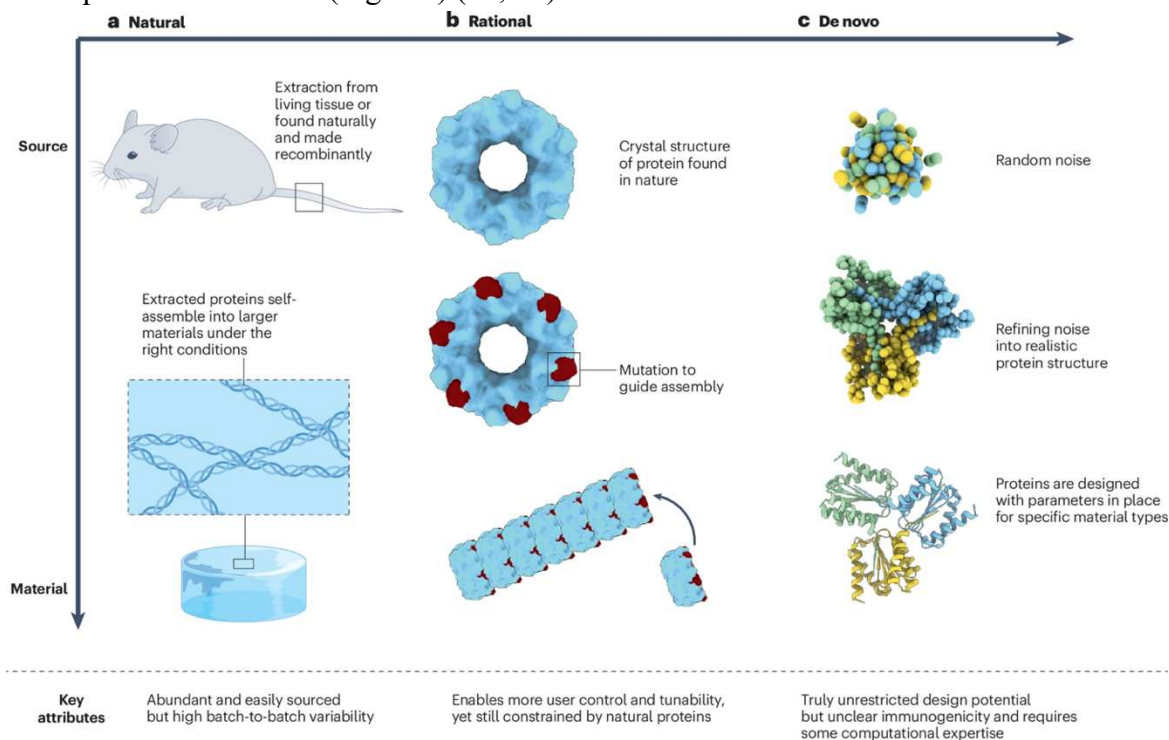


Figure 1.2: Design schemes in protein-based materials. **A.** shows adaptation of natural tissues/recombinant proteins for assembly in heterogenous materials such as hydrogels. **B.** depicts rational design principles for engineering biomaterials while **C.** shows one of the approaches to de novo protein design, gaussian noise is used to predict structures for new materials. From “Engineering complexity into protein-based biomaterials for biomedical applications.” Gregorio, N. E., *et al.* 2025, Nature Reviews Materials, 1–15. Figure 1 (<https://doi.org/10.1038/s41578-025-00861-8>) Reprinted with permission from Springer Nature. (24)

The Cry3A protein provides a well-studied natural example of biological crystallinity that has been repurposed as a delivery and immobilization scaffold (26, 27). One major functionalization strategy is genetic fusion, in which the cargo is encoded directly in the amino acid sequence of the crystal-forming polypeptide (27-30). Using this approach, the Chan lab has shown that engineered Cry3A-derived crystals can deliver fused protein cargo intracellularly, and that a positively charged Cry3A variant, Pos3A, improves cellular uptake and cytosolic delivery in these systems. (31). Reported cargoes include proteins such as p53 and antimicrobial peptides (AMPs), illustrating the capacity of this system to transport diverse functional biomolecules (31, 32). In parallel, Cry3A crystals have also been adapted for post-expression functionalization (26).

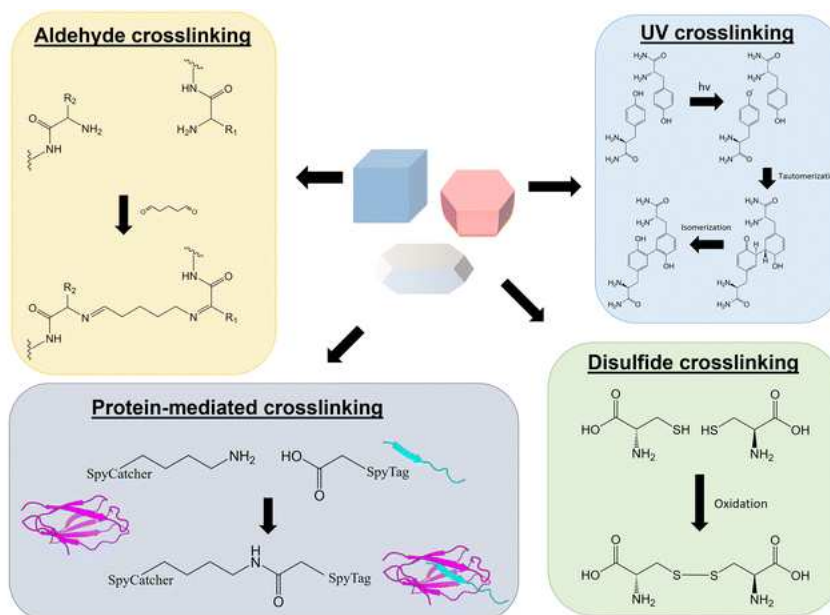


Figure 1.3 Methods of protein crosslinking and covalent bond installation. From “Porous protein crystals: synthesis and applications” A. A. Jones and C. D. Snow, Chem. Commun., 2024, 60, 5790 DOI: 10.1039/D4CC00183D. CC BY (39)

This underscores an advantageous ability of protein nanoparticles to be functionalized post-expression through crosslinking or covalent bond introduction (27, 33, 35). Examples of common post-expression crosslinking schema used for protein functionalization are split-

peptides (SpyCatcher-SpyTag systems), disulfide reduction, aldehyde mediated, and UV-catalysis (Fig. 1.3) (35-39).

Beyond genetic fusion and covalent bond stabilized cargo loading, the Cry3A platform exemplifies another favorable feature of protein crystals- the presence of solvent channel pores that permit diffusion of macromolecules throughout the crystalline lattice. In AMP-delivery studies, cargo loading was attributed, in part, to electrostatic interactions between cationic peptide cargo and negatively charged regions within the crystal channels. Increasing the positive charge of the solvent-exposed patch reduced peptide uptake, whereas introducing additional negative charge enhanced loading efficiency (40). These findings underscore how the understanding of defined crystal architectures, where regions not directly involved in subunit assembly, such as channel walls enable rational tuning of cargo interactions while maintaining other key materials properties. Correspondingly, Cry3A-based formulations demonstrated peptide stabilization and delivery in both cell-based and murine models, including antimicrobial and anticancer applications, with preclinical studies reporting minimal cytotoxicity and no observable liver or kidney toxicity under the conditions tested (40).

Although batch-to-batch characterization of Cry3A nanoparticles is routinely performed using ensemble measurements such as zeta potential, larger crystalline assemblies have been directly characterized using X-ray-based structural methods (13). Structural information obtained from these ordered crystals can be extrapolated to the nanoparticle scale, enabling hypothesis-driven design of cargo loading strategies. The AMP-Cry3A delivery system thus illustrates how atomic-resolution structural information can inform functional nanoparticle engineering.

Cry3A highlights the versatility of protein-based crystalline nanoparticles but represents only one example within a broader class of ordered protein assemblies adapted for biomedical applications. Building on the systems from living cells described earlier, engineered protein nanoparticles often take natural biomolecular assemblies and expand their utility. Ferritin cages, for instance, have been extensively engineered beyond their natural formation as highly ordered protein nanoparticles for cargo encapsulation, surface functionalization, and targeted delivery (41, 42). While ferritin is not an extended crystal, its repetitive symmetry confers substantial structural definition, and its assembly behavior has been characterized in solution by small angle X-ray scattering (SAXS) (20). Similarly, engineered polyhedrin-based particles have emerged as a distinct class of protein delivery materials (43). Derived from baculoviral polyhedrin, which form dense protective crystalline matrices around occlusion-derived virions, these assemblies leverage self-assembly and environmental stability for engineered cargo incorporation (44). Collectively, these systems demonstrate that periodic or highly repetitive organization is a powerful design principle for protein-based nanoparticles, enabling the creation of structurally defined, stable, and functionally adaptable biomaterials.

Beyond direct engineering of naturally occurring proteins and their crystal contacts, advances in *de novo* protein design now allow periodic protein assemblies to be designed computationally (45-47). This capability is rooted in the fact that modern structure-prediction and sequence-design algorithms, including AlphaFold, ProteinMPNN, and RFdiffusion, were trained on experimentally solved protein structures and assemblies archived in the PDB (48-50). Consequently, these models can replicate the structural regularities of folded proteins and multimeric interfaces, including the same stereochemical and packing constraints that underlie many crystallographic assemblies (47). These tools have enabled the forward design of new

protein nanomaterials whose architectures can be realized experimentally and validated structurally (51, 52). For example, King and colleagues reported a generalizable workflow for designing self-assembling protein nanoparticles from AlphaFold2-predicted oligomeric building blocks, generating 60-subunit icosahedral nanoparticles whose structures were validated by single-particle cryo-EM with atomic-level agreement to the design models (53). When one such scaffold was converted into an influenza immunogen, it elicited potent receptor-blocking and neutralizing antibody responses in mice, highlighting that these designed assemblies can also serve as functional vaccine platforms. The King lab's work demonstrates how protein structure and sequence prediction algorithms are expanding the field of designed protein-based nanomaterials.

1.4 DNA-based Crystalline Nanoparticles

DNA-based materials leverage the predictability of Watson–Crick base pairing to program complex secondary and tertiary architectures with high precision. This programmability has enabled a broad family of ordered assemblies, ranging from classical motif-based DNA crystals such as 13-mer lattices and tensegrity triangle frameworks to larger DNA-origami lattices and DNA-directed colloidal superlattices (Fig. 1.3) (54-57). Of these, the DNA-based crystalline assemblies are periodic lattices with defined geometries, tunable pore dimensions, and high structural precision (58). These properties have been leveraged to study enzyme immobilization, molecular sieving, guest encapsulation, and ordered molecular display, collectively illustrating DNA's capacity to form porous crystalline scaffolds capable of organizing molecular cargo (59-62).

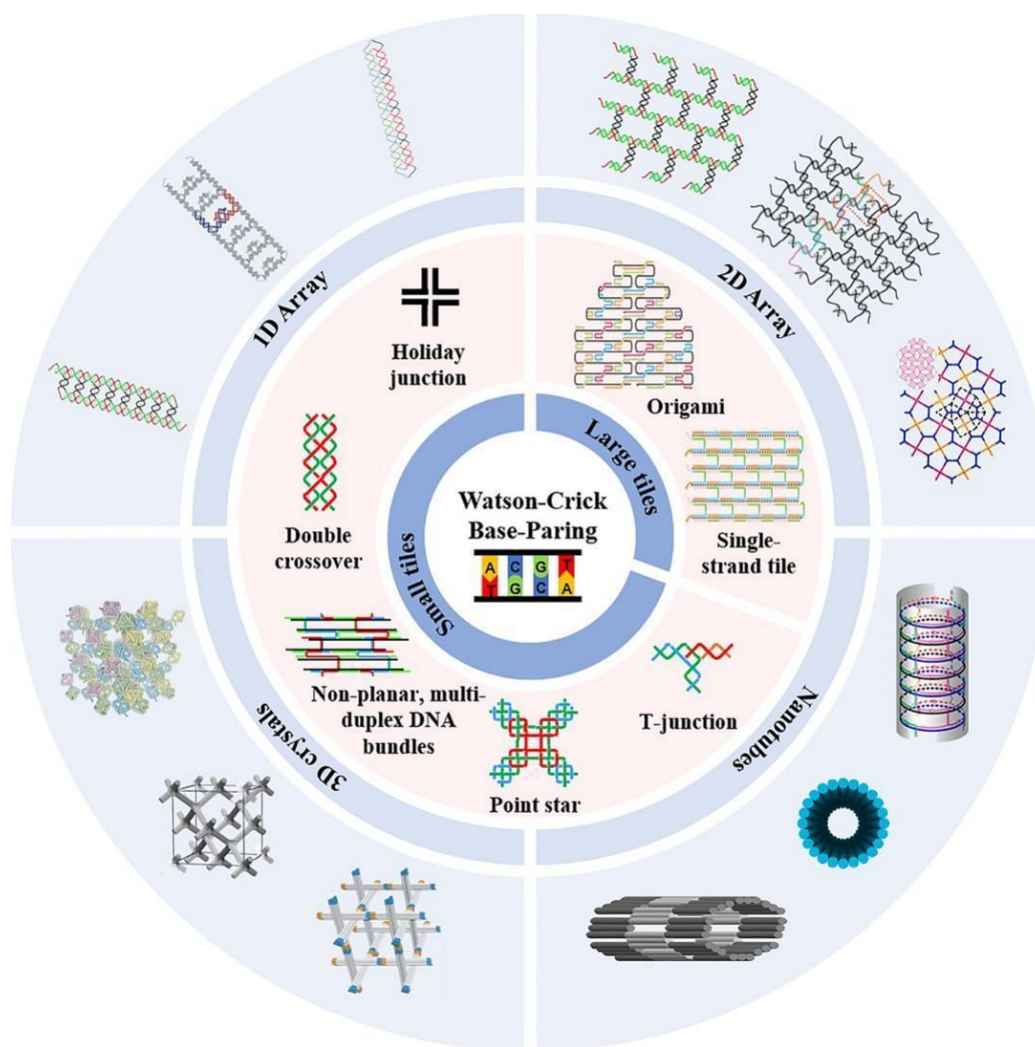


Figure 1.4: DNA nanostructure designs, ranging from 1D arrays (such as double crossover tiles), 2D arrays (such as DNA origami), nanotubes, to 3D crystals (such as tensegrity triangles). From “Design principles for construction of DNA-based nanostructures.” Chen, C., *et al*, 2025, *Advanced Drug Delivery Reviews*, 226, 115679 Figure 1 (<https://doi.org/10.1016/j.addr.2025.115679>) Reprinted with permission from Elsevier. (58)

Despite these strengths, extended DNA crystals remain far less developed for biomedical applications than discrete DNA nanostructures such as tetrahedral DNA nanoparticles and DNA origami-based carriers (63-65), as well as DNA-programmed nanoparticle superlattices that impose order on colloidal substrates (66). This divergence reflects fundamental engineering challenges associated with translating porous, lattice-based DNA materials into nanoparticle systems. These challenges include limited mechanical rigidity of non-covalent DNA junctions,

susceptibility to nuclease degradation, sensitivity to ionic strength and solution conditions, non-specific protein corona accumulation, and difficulty preserving long-range crystalline order under physiological environments (67-70). These constraints complicate the production of monodisperse crystalline particles with defined morphology and consistent functional performance. In addition, the cost and synthetic burden associated with production of modified DNA strands have slowed adoption of porous DNA crystals outside of the structural biology and nanotechnology communities (69). While recent chemical and enzymatic ligation strategies have improved lattice robustness and expanded accessible architectures (55, 58, 68, 69), scalable manufacturing and interdisciplinary work to demonstrate DNA crystal utility in biomedical applications remain unaddressed challenges.

Even so, DNA crystalline systems retain distinctive advantages that justify continued investigation. Their periodicity, porosity, and atomically defined organization enable spatial control over guest positioning that is difficult to achieve in amorphous or dynamically assembled nanomaterials. Recent advances in biomolecular structure prediction further expand this design space. Notably, AlphaFold 3 enables joint modeling of complexes containing nucleic acids, ligands, ions, and chemically modified residues, facilitating increasingly sophisticated, structure-guided design of pure and hybrid DNA-containing biomaterials (50, 70-72). Taken together, these advances position DNA-based crystalline systems as an underutilized class of materials whose greatest impact may emerge through hybrid architectures that combine crystalline order with the functional versatility of other biomolecular scaffolds.

1.5 Hybrid Crystalline Biomaterials

Hybrid materials integrate distinct classes of building blocks within a single nanoparticle platform, enabling complementary properties to be combined. In these systems, one component

may contribute structural order, rigidity, or environmental stability, while another provides biocompatibility, molecular recognition, or programmable cargo interactions. By uniting these attributes, hybrid nanoparticles can achieve functional combinations that are difficult or impossible to realize in single-material systems (73).

One prominent class of hybrid materials that are defined by their periodicity are biomolecule-integrated metal-organic frameworks (MOFs). MOFs are porous crystalline materials composed of metal ions and organic ligands that connect via metal-ligand coordination (74). In hybrid MOF nanoparticles, the crystalline order and porosity of the MOF act as a scaffold for molecular encapsulation and organization, while incorporated biomolecules impart biological specificity and function (75). Protein-MOF hybrids, for example, have been developed as platforms for enzyme immobilization, catalysis, biosensing, and therapeutic delivery, where the framework can stabilize proteins while preserving their activity (76-78). Similarly, DNA-MOF hybrids combine the structural regularity of the framework with the programmability of nucleic acids, enabling sequence-specific recognition, stimuli-responsive behavior, and guided assembly (79, 80). Together, these modified MOFs illustrate how crystalline order can be paired with biological functionality to produce structurally defined yet adaptable hybrid materials.

A second major class of hybrid nanoparticles is represented by protein-DNA assemblies. In these systems, proteins contribute folded structural motifs, catalytic surfaces, or recognition interfaces, while DNA provides programmability and predictable, tunable spacing (81, 82). Depending on the design strategy, protein-DNA hybrids can form discrete nanoparticles, extended lattices, or porous co-crystals, making them particularly attractive for molecular capture, scaffolded assembly, and therapeutic cargo organization (82-87). Our lab has demonstrated that hybrid protein-DNA co-crystals can be chemically ligated to overcome

structural integrity issues in biomimetic solutions, a main barrier to date in pure DNA crystals in biomedical applications (55, 88). Additionally, our lab has shown that protein-DNA co-crystals are capable of site-specific small molecule accumulation (87). These examples demonstrate that hybrid assemblies are a powerful route for merging crystallinity, programmability, and biomolecular function within a single nanomaterial platform.

1.6 Future Directions

Precision nanomedicine continues to be a focus of drug delivery advancement, and the structural precision imparted by periodicity is a distinct advantage of crystalline nanoparticles in this space. *De novo* design of biomolecules and their assembly, as well as reliable sequence and structure prediction expansion to more complex assemblies is giving rise to increasingly sophisticated, predictably assembled nanomaterials with atomically defined architectures and programmable function.

Here we expand on the functionalization of a protein-DNA co-crystal system with potential as a hybrid nanoparticle. This system is made up of a prokaryotic transcription factor (RepE54) bound to its cognate 19 bp iteron plus two flanking bases, for a core 21 bp sequence. This protein-DNA complex can be expanded by adding full turns of B-form DNA (10 bp, 21 bp, etc.). Our previous reports on protein-DNA co-crystals characterized our system, Co-Crystal 1 (CC1) utilizing the same building blocks described in this dissertation, but with crystallization conditions that yielded a less porous, doubled lattice referred to as interpenetrating CC1 (ipCC1) (55, 87, 89). Our recent paper reports our discovery of crystallization conditions which yield porous protein-DNA co-crystals (90). In this report we demonstrate intracrystalline diffusion of guest proteins from three families, their site-specific organization, and structural observation of these guests via XRD. Porous CC1 combines the programmability of the DNA components,

interfacial structural rigidity of the protein components, large solvent channel pores, and high-resolution structural characterization. These advantages position CC1 as an ideal platform for the future of biomolecular, crystalline nanoparticles for biomedical applications.

1.7 Specific Aims

With the advent of porous CC1 and the demonstrated modularity with which its crystals can be grown, further development of the design principles that govern this class of novel protein-DNA co-crystals remains to be explored. This dissertation looks at expanding the utility of CC1 through design of novel co-crystal variants (Chapter II), optimization of enzymatic and chemical ligation (Chapter III), demonstration of an intracrystalline device operation (Chapter IV), micro batch crystallization (Chapter V), and enumeration of future directions for this project (Chapter VI).

In Chapter II, we demonstrate the robust modularity of the CC1 platform in accommodating perturbations to the nucleic acid components of the lattice. We show that while crystallization conditions favoring porous CC1 formation are generally reliable, they are not infallible. Chemical modifications of the building blocks can bias lattice assembly toward the interpenetrating CC1 (ipCC1) habit even under these conditions. Additionally, we demonstrate that favorable symmetry contacts within the lattice can enable the structural observation of co- and post-crystallization guest molecules that would otherwise be too labile for detection by XRD.

In Chapter III we show optimization of ligation yields across the DNA-DNA junctions of CC1 via enzymatic and chemical ligation. Previous work with ipCC1 showed that chemical ligation was sufficient to provide remarkable stability in biomimetic solvent mixtures. Here we show that chemical ligation in porous CC1 can similarly enable crystal survival outside of

mother liquor, with discussion of limitations in reaching the upper limits of ligation efficiency. The solvent channel pores of CC1+10 are sufficient to allow diffusion of small ligase AmeLig. This chapter shows how enzymatic ligation yields differ from chemical ligation, as well as how restriction enzyme diffusion can perform controlled demolition of CC1 building blocks.

In Chapter IV we demonstrate how CC1 grown with unpaired ssDNA tails can be adapted to capture, store, and release nucleic acids from solution with sequence specificity. Using base complementarity, single stranded DNA and RNA are specifically accumulated in CC1 by post-crystallization soaking. Wash steps show how these bound guest species are stabilized within the lattice. Then, using strand displacement and toehold switches we show release of the guest species in both single-stranded and double-stranded forms.

In Chapter V we show batch micro crystallization of CC1. Moving through crystallization space from macro, vapor diffusion conditions to micro, under oil conditions allows for scaled up production of crystals that can then be size selected via filtration and batch characterized.

This dissertation provides exploration of the limits of modularity that the CC1 platform can tolerate, with a focus on modalities that enhance the potential of this platform to be adapted into nanomedicine. We also explore site specific organization of biomolecules for structural observation, enzymatic ligation in a novel platform, and dynamic processes within the intracrystalline environment. The developments presented herein takes the demonstrated utility of CC1 from structural biology into the nanotechnology and nanomedicine space.

CHAPTER 1 BIBLIOGRAPHY

1. Diaz, M. J., Natarelli, N., Aflatooni, S., Aleman, S. J., Neelam, S., Tran, J. T., Taneja, K., Lucke-Wold, B., & Forouzandeh, M. (2023). Nanoparticle-Based Treatment Approaches for Skin Cancer: A Systematic Review. *Current Oncology*, 30(8), 7112–7131. <https://doi.org/10.3390/curroncol30080516>
2. Mitchell, M. J., Billingsley, M. M., Haley, R. M., Wechsler, M. E., Peppas, N. A., & Langer, R. (2021). Engineering precision nanoparticles for drug delivery. *Nature Reviews Drug Discovery*, 20(2), 101–124. <https://doi.org/10.1038/s41573-020-0090-8>
3. Moazzam, M., Zhang, M., Hussain, A., Yu, X., Huang, J., & Huang, Y. (2024). The landscape of nanoparticle-based siRNA delivery and therapeutic development. *Molecular Therapy*, 32(2), 284–312. <https://doi.org/10.1016/j.ymthe.2024.01.005>
4. Joudeh, N., & Linke, D. (2022). Nanoparticle classification, physicochemical properties, characterization, and applications: A comprehensive review for biologists. *Journal of Nanobiotechnology*, 20, 262. <https://doi.org/10.1186/s12951-022-01477-8>
5. Rezić, I. (2022). Nanoparticles for Biomedical Application and Their Synthesis. *Polymers*, 14(22), 4961. <https://doi.org/10.3390/polym14224961>
6. Cheng, X., Xie, Q., & Sun, Y. (2023). Advances in nanomaterial-based targeted drug delivery systems. *Frontiers in Bioengineering and Biotechnology*, 11, 1177151. <https://doi.org/10.3389/fbioe.2023.1177151>
7. Eker, F., Duman, H., Akdaşçi, E., Bolat, E., Sarıtaş, S., Karav, S., & Witkowska, A. M. (2024). A Comprehensive Review of Nanoparticles: From Classification to Application and Toxicity. *Molecules*, 29(15). <https://doi.org/10.3390/molecules29153482>
8. Li, K.-M., Song, P., He, X.-P., & Li, N. (2025). Probing Membrane Structure of Lipid Nanomedicines Using Solution Small-Angle X-Ray Scattering: Applications and Prospects. *Membranes*, 15(12), 382. <https://doi.org/10.3390/membranes15120382>
9. Doye, J. P. K., & Poon, W. C. K. (2006). Protein crystallization *in vivo*. *Current Opinion in Colloid & Interface Science*, 11(1), 40–46. <https://doi.org/10.1016/j.cocis.2005.10.002>
10. Asai, S., Moravcová, J., Žáková, L., Selicharová, I., Hadravová, R., Brzozowski, A. M., Nováček, J., & Jiráček, J. (n.d.). Characterization of insulin crystalline form in isolated β -cell secretory granules. *Open Biology*, 12(12), 220322. <https://doi.org/10.1098/rsob.220322>
11. Yang, J. E., Mitchell, J., Bingman, C., Mosher, D. F., & Wright, E. R. (2024). In-Situ Cryo-EM Structural Studies of Eosinophil Granules. *Microscopy and Microanalysis*, 30(Supplement_1), ozae044.348. <https://doi.org/10.1093/mam/ozae044.348>
12. Höfte, H., & Whiteley, H. R. (1989). Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiological Reviews*, 53(2), 242–255. <https://doi.org/10.1128/mr.53.2.242-255.1989>
13. Sawaya, M. R., Cascio, D., Gingery, M., Rodriguez, J., Goldschmidt, L., Colletier, J.-P., Messerschmidt, M. M., Boutet, S., Koglin, J. E., Williams, G. J., Brewster, A. S., Nass, K., Hattne, J., Botha, S., Doak, R. B., Shoeman, R. L., DePonte, D. P., Park, H.-W., Federici, B. A., ... Eisenberg, D. S. (2014). Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proceedings of the National Academy of Sciences*, 111(35), 12769–12774. <https://doi.org/10.1073/pnas.1413456111>

14. Roos, W. H., Ivanovska, I. L., Evilevitch, A., & Wuite, G. J. L. (2007). Viral capsids: Mechanical characteristics, genome packaging and delivery mechanisms. *Cellular and Molecular Life Sciences: CMLS*, 64(12), 1484. <https://doi.org/10.1007/s00018-007-6451-1>
15. Sevvana, M., Klose, T., & Rossmann, M. G. (2021). Principles of Virus Structure. *Encyclopedia of Virology*, 257–277. <https://doi.org/10.1016/B978-0-12-814515-9.00033-3>
16. Ji, X., Sutton, G., Evans, G., Axford, D., Owen, R., & Stuart, D. I. (2010). How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. *The EMBO Journal*, 29(2), 505–514. <https://doi.org/10.1038/emboj.2009.352>
17. Sampieri, A., Luz-Madrigal, A., Zepeda, J., & Vaca, L. (2015). Identification of fragments from *Autographa Californica* polyhedrin protein essential for self-aggregation and exogenous protein incorporation. *BMC Biochemistry*, 16, 5. <https://doi.org/10.1186/s12858-015-0034-9>
18. Kim, M., Rho, Y., Jin, K. S., Ahn, B., Jung, S., Kim, H., & Ree, M. (2011). pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly. *Biomacromolecules*, 12(5), 1629–1640. <https://doi.org/10.1021/bm200026v>
19. Andrews, S. C. (2010). The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochimica et Biophysica Acta (BBA) - General Subjects, Ferritin: Structures, Properties and Applications*, 1800(8), 691–705. <https://doi.org/10.1016/j.bbagen.2010.05.010>
20. Sato, D., Ohtomo, H., Yamada, Y., Hikima, T., Kurobe, A., Fujiwara, K., & Ikeguchi, M. (2016). Ferritin Assembly Revisited: A Time-Resolved Small-Angle X-ray Scattering Study. *Biochemistry*, 55(2), 287–293. <https://doi.org/10.1021/acs.biochem.5b01152>
21. Chu, S., Wang, A. L., Bhattacharya, A., & Montclare, J. K. (2022). Protein Based Biomaterials for Therapeutic and Diagnostic Applications. *Progress in Biomedical Engineering (Bristol, England)*, 4(1), 012003. <https://doi.org/10.1088/2516-1091/ac2841>
22. Tripathi, N. K., & Shrivastava, A. (2019). Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Frontiers in Bioengineering and Biotechnology*, 7, 420. <https://doi.org/10.3389/fbioe.2019.00420>
23. Olshefsky, A., Richardson, C., Pun, S. H., & King, N. P. (2022). Engineering Self-Assembling Protein Nanoparticles for Therapeutic Delivery. *Bioconjugate Chemistry*, 33(11), 2018–2034. <https://doi.org/10.1021/acs.bioconjchem.2c00030>
24. Gregorio, N. E., Haas, C. M., King, N. P., & DeForest, C. A. (2025). Engineering complexity into protein-based biomaterials for biomedical applications. *Nature Reviews Materials*, 1–15. <https://doi.org/10.1038/s41578-025-00861-8>
25. Haas, C. M., Rankovic, S., Lewis, H. K., Carr, K. D., Weidle, C., Gerdes, S. S., Nuss, L. R., Ruiz, F., Moiz, S., Fiorelli, M., Grey, E., McGowan, J., Kumar, N., Creanga, A., Kang, A., Nguyen, H., Wang, Y., Sankaran, B., Dosey, A., ... King, N. P. (2026). De novo design of protein nanoparticles with integrated functional motifs (p. 2025.12.19.695620). *bioRxiv*. <https://doi.org/10.64898/2025.12.19.695620>
26. Sun, Q., Heater, B. S., Li, T. L., Ye, W., Guo, Z., & Chan, M. K. (2022). Cry3Aa*SpyCatcher Fusion Crystals Produced in Bacteria as Scaffolds for Multienzyme Coimmobilization. *Bioconjugate Chemistry*, 33(2), 386–396. <https://doi.org/10.1021/acs.bioconjchem.2c00003>
27. Biocrystallization: Harnessing natural scaffolds for stable protein crystal formulations.

- (2025). *New Biotechnology*, 90, 257–267. <https://doi.org/10.1016/j.nbt.2025.10.010>
28. Kowalski, A. E., Johnson, L. B., Dierl, H. K., Park, S., Huber, T. R., & Snow, C. D. (2019). Porous protein crystals as scaffolds for enzyme immobilization. *Biomaterials Science*, 7(5), 1898–1904. <https://doi.org/10.1039/C8BM01378K>
 29. Yang, Z., Yang, M., Chow, H.-M., Tsang, S. Y., Lee, M. M., & Chan, M. K. (2021). Cytosolic delivery of CDK4/6 inhibitor p16 protein using engineered protein crystals for cancer therapy. *Acta Biomaterialia*, 135, 582–592. <https://doi.org/10.1016/j.actbio.2021.08.054>
 30. Hartje, L. F., & Snow, C. D. (2019). Protein crystal based materials for nanoscale applications in medicine and biotechnology. *WIREs Nanomedicine and Nanobiotechnology*, 11(4), e1547. <https://doi.org/10.1002/wnan.1547>
 31. Zheng, W., Rädler, J., Sork, H., Niu, Z., Roudi, S., Bost, J. P., Görgens, A., Zhao, Y., Mamand, D. R., Liang, X., Wiklander, O. P. B., Lehto, T., Gupta, D., Nordin, J. Z., & EL Andaloussi, S. (2023). Identification of scaffold proteins for improved endogenous engineering of extracellular vesicles. *Nature Communications*, 14, 4734. <https://doi.org/10.1038/s41467-023-40453-0>
 32. Yang, Z., Sun, J. K.-L., Lee, M. M., & Chan, M. K. (2022). Restoration of p53 activity via intracellular protein delivery sensitizes triple negative breast cancer to anti-PD-1 immunotherapy. *Journal for Immunotherapy of Cancer*, 10(9), e005068. <https://doi.org/10.1136/jitc-2022-005068>
 33. Zhang, W., Yang, Z., Zheng, J., Fu, K., Wong, J. H., Ni, Y., Ng, T. B., Cho, C. H., Chan, M. K., & Lee, M. M. (n.d.). A Bioresponsive Genetically Encoded Antimicrobial Crystal for the Oral Treatment of *Helicobacter Pylori* Infection. <https://doi.org/10.1002/advs.202301724>
 34. Hartje, L. F., Andales, D. A., Gintner, L. P., Johnson, L. B., Li, Y. V., & Snow, C. D. (2023). Textile Functionalization by Porous Protein Crystal Conjugation and Guest Molecule Loading. *Crystals*, 13(2), Article 2. <https://doi.org/10.3390/cryst13020352>
 35. Huber, T. R., McPherson, E. C., Keating, C. E., & Snow, C. D. (2018). Installing Guest Molecules at Specific Sites within Scaffold Protein Crystals. *Bioconjugate Chemistry*, 29(1), 17–22. <https://doi.org/10.1021/acs.bioconjchem.7b00668>
 36. Jones, M. W., Strickland, R. A., Schumacher, F. F., Caddick, S., Baker, James. R., Gibson, M. I., & Haddleton, D. M. (2012). Polymeric Dibromomaleimides As Extremely Efficient Disulfide Bridging Bioconjugation and Pegylation Agents. *Journal of the American Chemical Society*, 134(3), 1847–1852. <https://doi.org/10.1021/ja210335f>
 37. Hermanson, G. T. (2013). Chapter 3—The Reactions of Bioconjugation. In G. T. Hermanson (Ed.), *Bioconjugate Techniques* (Third Edition) (pp. 229–258). Academic Press. <https://doi.org/10.1016/B978-0-12-382239-0.00003-0>
 38. Craenmehr, F. W. B., Gräwe, A., Veenbrink, V. A., Bellan, R., Merckx, M., & Dankers, P. Y. W. (2025). Employing the SpyTag-SpyCatcher Reaction for the Modification of Supramolecular Polymers with Functional Proteins. *Bioconjugate Chemistry*, 36(6), 1197–1207. <https://doi.org/10.1021/acs.bioconjchem.5c00046>
 39. Jones, A. A., & Snow, C. D. (2024). Porous protein crystals: Synthesis and applications. *Chemical Communications*, 60(45), 5790–5803. <https://doi.org/10.1039/D4CC00183D>
 40. Targeted delivery of antimicrobial peptide by Cry protein crystal to treat intramacrophage infection. (2019). *Biomaterials*, 217, 119286. <https://doi.org/10.1016/j.biomaterials.2019.119286>

41. Jutz, G., van Rijn, P., Santos Miranda, B., & Böker, A. (2015). Ferritin: A Versatile Building Block for Bionanotechnology. *Chemical Reviews*, 115(4), 1653–1701. <https://doi.org/10.1021/cr400011b>
42. Reutovich, A. A., Srivastava, A. K., Arosio, P., & Bou-Abdallah, F. (2023). Ferritin nanocages as efficient nanocarriers and promising platforms for COVID-19 and other vaccines development. *Biochimica et Biophysica Acta. General Subjects*, 1867(3), 130288. <https://doi.org/10.1016/j.bbagen.2022.130288>
43. Parwana, K. A. K., Kaur Gill, P., Njanike, R., Yiu, H. H. P., Adams, C. F., Chari, D. M., & Jenkins, S. I. (2024). Investigating Internalization of Reporter-Protein-Functionalized Polyhedrin Particles by Brain Immune Cells. *Materials*, 17(10), 2330. <https://doi.org/10.3390/ma17102330>
44. Wendler, A., James, N., Jones, M. H., & Pernstich, C. (n.d.). Phagocytosed Polyhedrin-Cytokine Cocystal Nanoparticles Provide Sustained Secretion of Bioactive Cytokines from Macrophages. *Biodesign Research*, 2021, 9816485. <https://doi.org/10.34133/2021/9816485>
45. Haas, C. M., Rankovic, S., Lewis, H. K., Carr, K. D., Weidle, C., Gerdes, S. S., Nuss, L. R., Ruiz, F., Moiz, S., Fiorelli, M., Grey, E., McGowan, J., Kumar, N., Creanga, A., Kang, A., Nguyen, H., Wang, Y., Sankaran, B., Dosey, A., ... King, N. P. (2026). De novo design of protein nanoparticles with integrated functional motifs (p. 2025.12.19.695620). *bioRxiv*. <https://doi.org/10.64898/2025.12.19.695620>
46. Rankovic, S., Carr, K. D., Decarreau, J., Skotheim, R., Kibler, R. D., Ols, S., Lee, S., Chun, J.-H., Tooley, M. R., Dauparas, J., Eisenach, H. E., Glögl, M., Weidle, C., Borst, A. J., Baker, D., & King, N. P. (2025). Computational design of bifaceted protein nanomaterials. *Nature Materials*, 24(10), 1635–1643. <https://doi.org/10.1038/s41563-025-02295-7>
47. Li, Z., Wang, S., Nattermann, U., Bera, A. K., Borst, A. J., Yaman, M. Y., Bick, M. J., Yang, E. C., Sheffler, W., Lee, B., Seifert, S., Hura, G. L., Nguyen, H., Kang, A., Dalal, R., Lubner, J. M., Hsia, Y., Haddox, H., Courbet, A., ... Baker, D. (2023). Accurate computational design of three-dimensional protein crystals. *Nature Materials*, 22(12), 1556–1563. <https://doi.org/10.1038/s41563-023-01683-1>
48. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
49. 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., ... Baker, D. (2022). Robust deep learning–based protein sequence design using ProteinMPNN. *Science*, 378(6615), 49–56. <https://doi.org/10.1126/science.add2187>
50. Butcher, J., Krishna, R., Mitra, R., Brent, R. I., Li, Y., Corley, N., Kim, P. T., Funk, J., Mathis, S., Salike, S., Muraishi, A., Eisenach, H., Thompson, T. R., Chen, J., Politanska, Y., Sehgal, E., Coventry, B., Zhang, O., Qiang, B., ... Baker, D. (2025). De novo Design of All-atom Biomolecular Interactions with RFdiffusion3 (p. 2025.09.18.676967). *bioRxiv*. <https://doi.org/10.1101/2025.09.18.676967>
51. Dowling, Q. M., Park, Y.-J., Fries, C. N., Gerstenmaier, N. C., Ols, S., Yang, E. C.,

- Wargacki, A. J., Dosey, A., Hsia, Y., Ravichandran, R., Walkey, C. D., Burrell, A. L., Veessler, D., Baker, D., & King, N. P. (2025). Hierarchical design of pseudosymmetric protein nanocages. *Nature*, 638(8050), 553–561. <https://doi.org/10.1038/s41586-024-08360-6>
52. Hendricks, G. G., Grigoryan, L., Navarro, M. J., Catanzaro, N. J., Hubbard, M. L., Powers, J. M., Mattocks, M., Treichel, C., Walls, A. C., Lee, J., Ellis, D., Wang, J. Y. (John), Cheng, S., Miranda, M. C., Valdez, A., Chao, C. W., Chan, S., Men, C., Johnson, M. R., ... King, N. P. (2025). Computationally designed mRNA-launched protein nanoparticle immunogens elicit protective antibody and T cell responses in mice. *Science Translational Medicine*, 17(820), eadu2085. <https://doi.org/10.1126/scitranslmed.adu2085>
 53. Haas, C. M., Jasti, N., Dosey, A., Allen, J. D., Gillespie, R., McGowan, J., Leaf, E. M., Crispin, M., DeForest, C. A., Kanekiyo, M., & King, N. P. (2025). From sequence to scaffold: Computational design of protein nanoparticle vaccines from AlphaFold2-predicted building blocks. *Proceedings of the National Academy of Sciences*, 122(45), e2409566122. <https://doi.org/10.1073/pnas.2409566122>
 54. Kong, H., Sun, B., Yu, F., Wang, Q., Xia, K., & Jiang, D. (2023). Exploring the Potential of Three-Dimensional DNA Crystals in Nanotechnology: Design, Optimization, and Applications. *Advanced Science*, 10(24), 2302021. <https://doi.org/10.1002/advs.202302021>
 55. Orun, A. R., Slaughter, C. K., Shields, E. T., Vajapayajula, A., Jones, S., Shrestha, R., & Snow, C. D. (2024). Tuning Chemical DNA Ligation within DNA Crystals and Protein–DNA Cocrystals. *ACS Nanoscience Au*, 4(5), 338–348. <https://doi.org/10.1021/acsnanoscienceau.4c00013>
 56. Lin, Q.-Y., Mason, J. A., Li, Z., Zhou, W., O'Brien, M. N., Brown, K. A., Jones, M. R., Butun, S., Lee, B., Dravid, V. P., Aydin, K., & Mirkin, C. A. (2018). Building superlattices from individual nanoparticles via template-confined DNA-mediated assembly. *Science*, 359(6376), 669–672. <https://doi.org/10.1126/science.aag0591>
 57. Zhou, Y., Dong, J., & Wang, Q. (2023). Fabricating higher-order functional DNA origami structures to reveal biological processes at multiple scales. *NPG Asia Materials*, 15(1), 25. <https://doi.org/10.1038/s41427-023-00470-3>
 58. Chen, C., Xia, X., Tian, C., Zhang, Z., Jin, J., Huang, C. Z., Zuo, H., & Mao, C. (2025). Design principles for construction of DNA-based nanostructures. *Advanced Drug Delivery Reviews*, 226, 115679. <https://doi.org/10.1016/j.addr.2025.115679>
 59. Geng, C., & Paukstelis, P. J. (2014). DNA Crystals as Vehicles for Biocatalysis. *Journal of the American Chemical Society*, 136(22), 7817–7820. <https://doi.org/10.1021/ja502356m>
 60. Paukstelis, P. J. (2006). Three-Dimensional DNA Crystals as Molecular Sieves. *Journal of the American Chemical Society*, 128(21), 6794–6795. <https://doi.org/10.1021/ja061322r>
 61. Hao, Y., Kristiansen, M., Sha, R., Birktoft, J. J., Hernandez, C., Mao, C., & Seeman, N. C. (2017). A device that operates within a self-assembled 3D DNA crystal. *Nature Chemistry*, 9(8), 824–827. <https://doi.org/10.1038/nchem.2745>
 62. Zhang, C., Zhao, J., Lu, B., Seeman, N. C., Sha, R., Noinaj, N., & Mao, C. (2023). Engineering DNA Crystals toward Studying DNA–Guest Molecule Interactions. *Journal of the American Chemical Society*, 145(8), 4853–4859.

- <https://doi.org/10.1021/jacs.3c00081>
63. Hu, Q., Li, H., Wang, L., Gu, H., & Fan, C. (2019). DNA Nanotechnology-Enabled Drug Delivery Systems. *Chemical Reviews*, 119(10), 6459–6506.
<https://doi.org/10.1021/acs.chemrev.7b00663>
64. Yang, D., & Mao, C. (2025). Nucleic Acid Nanotechnology: Innovations Shaping the Future of Biomedicine and Diagnostics. *Small Methods*, 9(6), 2500756.
<https://doi.org/10.1002/smtd.202500756>
65. Cai, B., Rong, X., Sun, Y., Liu, L., & Li, Z. (2025). Engineered 3D DNA Crystals: A Molecular Design Perspective. *Small Methods*, 9(6), 2401455.
<https://doi.org/10.1002/smtd.202401455>
66. Oh, T., Park, S. S., & Mirkin, C. A. (2019). Stabilization of Colloidal Crystals Engineered with DNA. *Advanced Materials*, 31(1), 1805480.
<https://doi.org/10.1002/adma.201805480>
67. Pinheiro, A. V., Han, D., Shih, W. M., & Yan, H. (2011). Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology*, 6(12), 763–772.
<https://doi.org/10.1038/nnano.2011.187>
68. Swett, A. A., Kim, Y. H., Park, S., Kesama, M. R., Siraj, N., Mao, C., & Choi, J. H. (2026). DNA Crystals as a Template for Patterned Functional Materials. *Small Structures*, 7(1), e202500693. <https://doi.org/10.1002/ssstr.202500693>
69. Zheng, M., Paluzzi, V. E., Zhang, C., & Mao, C. (2024). Toward Production of Nucleic Acid Nanostructures in Life Cells and Their Biomedical Applications. In *DNA Nanotechnology for Cell Research* (pp. 401–413). John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9783527840816.ch17>
70. Huzar, J., Coreas, R., Landry, M. P., & Tikhomirov, G. (2025). AI-Based Prediction of Protein Corona Composition on DNA Nanostructures. *ACS Nano*, 19(4), 4333–4345.
<https://doi.org/10.1021/acsnano.4c12259>
71. Anusha, Zhang, Z., Li, J., Zuo, H., & Mao, C. (2024). AlphaFold 3 – Aided Design of DNA Motifs To Assemble into Triangles. *Journal of the American Chemical Society*, 146(37), 25422–25425. <https://doi.org/10.1021/jacs.4c08387>
72. Cortes, M., Sun, X., Anusha, Joseph Batchelder-Schwab, E., Li, J., Siraj, N., Jampana, R., Zhang, Y., Bai, Y., & Mao, C. (2025). AlphaFold 3 modeling of DNA nanomotifs: Is it reliable? *Nanoscale Horizons*, 10(7), 1428–1435.
<https://doi.org/10.1039/D5NH00059A>
73. Zhang, Y., Shi, H., Li, Y., Shen, Y., Wang, T., & Liu, J. (2026). Biological Nanoscaffolds from Hierarchical Construction to Applications. *Molecules*, 31(5).
<https://doi.org/10.3390/molecules31050812>
74. Mi, J., Jiang, X., Yin, M., Wang, Y., Tong, Y., & Zhu, Y. (2026). The application of MOF-based drug delivery systems in tumor immunotherapy. *Colloids and Surfaces B: Biointerfaces*, 262, 115506. <https://doi.org/10.1016/j.colsurfb.2026.115506>
75. Du, C., Xu, Y., & Wei, G. (2025). Recent Advances in Biomolecule-Engineered Metal-Organic Frameworks (Bio-MOFs): From Design, Bioengineering, and Structural/functional Regulation to Biocatalytic Applications. *The Chemical Record*, 25(6), e202500001. <https://doi.org/10.1002/tcr.202500001>
76. Liutkus, M., Sasselli, I. R., Rojas, A. L., & Cortajarena, A. L. (2024). Diverse crystalline protein scaffolds through metal-dependent polymorphism. *Protein Science*, 33(5), e4971.
<https://doi.org/10.1002/pro.4971>

77. Xu, X., Wan, J., Sun, J., Wu, L., & Lei, J. (2025). Tailored Metal-Organic Framework-Based Enzyme Hybrids: Immobilization Strategies, Improved Performance, and Biological Applications. *Small Science*, 5(10), 2500222. <https://doi.org/10.1002/smssc.202500222>
78. DeRoo, J. B., Shrestha, R., Jones, A., Rajendran, D., Thai, J. E., Tuttle, R. R., Snow, C. D., & Reynolds, M. M. (2025). Deposition of metal–organic frameworks within a porous protein crystal superstructure. *Journal of Materials Chemistry B*, 13(35), 10864–10870. <https://doi.org/10.1039/D5TB00425J>
79. Wang, S., Chen, Y., Wang, S., Li, P., Mirkin, C. A., & Farha, O. K. (2019). DNA-Functionalized Metal–Organic Framework Nanoparticles for Intracellular Delivery of Proteins. *Journal of the American Chemical Society*, 141(6), 2215–2219. <https://doi.org/10.1021/jacs.8b12705>
80. He, L., Shang, M., Chen, Z., & Yang, Z. (2023). Metal-Organic Frameworks Nanocarriers for Functional Nucleic Acid Delivery in Biomedical Applications. *The Chemical Record*, 23(6), e202300018. <https://doi.org/10.1002/tcr.202300018>
81. Han, Z., Kim, Y. J., & Mirkin, C. A. (2025). Inverse Design of DNA-Programmable Protein Colloidal Crystals. *Journal of the American Chemical Society*, 147(44), 41087–41095. <https://doi.org/10.1021/jacs.5c15424>
82. Stephanopoulos, N. (2020). Hybrid Nanostructures from the Self-Assembly of Proteins and DNA. *Chem*, 6(2), 364–405. <https://doi.org/10.1016/j.chempr.2020.01.012>
83. Hayes, O. G., Partridge, B. E., & Mirkin, C. A. (2021). Encoding hierarchical assembly pathways of proteins with DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 118(40), e2106808118. <https://doi.org/10.1073/pnas.2106808118>
84. Brodin, J. D., Auyeung, E., & Mirkin, C. A. (2015). DNA-mediated engineering of multicomponent enzyme crystals. *Proceedings of the National Academy of Sciences*, 112(15), 4564–4569. <https://doi.org/10.1073/pnas.1503533112>
85. Praetorius, F., & Dietz, H. (2017). Self-assembly of genetically encoded DNA-protein hybrid nanoscale shapes. *Science*, 355(6331), eaam5488. <https://doi.org/10.1126/science.aam5488>
86. McCluskey, J. B., Clark, D. S., & Glover, D. J. (2020). Functional Applications of Nucleic Acid–Protein Hybrid Nanostructures. *Trends in Biotechnology*, 38(9), 976–989. <https://doi.org/10.1016/j.tibtech.2020.02.007>
87. Orun, A. R., Shields, E. T., Dmytriw, S., Vajapayajula, A., Slaughter, C. K., & Snow, C. D. (2023). Modular Protein–DNA Cocystals as Precise, Programmable Assembly Scaffolds. *ACS Nano*, 17(14), 13110–13120. <https://doi.org/10.1021/acsnano.2c07282>
88. Komori, H., Sasai, N., Matsunaga, F., Wada, C., & Miki, K. (1999). Crystallization and Preliminary X-Ray Diffraction Studies of a Replication Initiator Protein (RepE54) of the Mini-F Plasmid Complexed with Iiteron DNA. *Journal of Biochemistry*, 125(1), 24–26. <https://doi.org/10.1093/oxfordjournals.jbchem.a022262>
89. Orun, A. R., Dmytriw, S., Vajapayajula, A., & Snow, C. D. (2022). Stabilizing DNA–Protein Co-Crystals via Intra-Crystal Chemical Ligation of the DNA. *Crystals*, 12(1), Article 1. <https://doi.org/10.3390/cryst12010049>
90. Shields, E. T., Slaughter, C. K., Mekkaoui, F., Magna, E. N., Shepherd, C., Lukeman, P. S., Spratt, D. E., & Snow, C. D. (2026). Modular Scaffold Crystals for Programmable Installation and Structural Observation of DNA-Binding Proteins (p.

2026.03.04.709581). bioRxiv. <https://doi.org/10.64898/2026.03.04.709581>

CRYSTALLINE BIOMATERIALS FOR SITE-SPECIFIC ORGANIZATION OF MODIFIED DNA

Caroline K. Slaughter, Ethan T. Shields, Emma N. Magna, Alexandra P. Platt, Robert S. Paton, Christopher D. Snow.

Conceptualization, writing-drafting, and visualization, C.K.S, C.D.S.; writing-editing, C.K.S., E.T.S., E.N.M., C.D.S.; formal analysis, data curation, acquisition, investigation, and validation; C.K.S., E.T.S., A.P.P., R.S.P., C.D.S.

2.1 Introduction

Protein–DNA co-crystals represent a hybrid class of crystalline nanomaterials that combine the programmability of nucleic acid assemblies with the chemical diversity and other favorable properties of protein crystals^{1–5}. For example, one porous protein crystal with 13 nanometer solvent channels nonetheless provides high-resolution X-ray diffraction (XRD), adequate to see ordered water molecules and the plane of carbonyl peptide bonds (pdb 5W17)^{6,7}. DNA-only crystals with pores large enough for intra-crystal macromolecule diffusion tend to have resolution limits near 5 Å, with higher resolution structures reported with less voluminous pores^{8–12}. Pure-protein crystals have been used for hosting guest molecules for structural determination, but they face difficulty in rapid adaptability for novel targets^{7,13,14}. DNA-based building blocks have also been assembled into two- and three-dimensional crystalline arrays capable of positioning nanoparticles and other guest species at predefined lattice sites^{15–19}. These studies demonstrate the potential of DNA as both a structural component and a chemically addressable scaffold. The chemical programmability of nucleic acids extends beyond canonical Watson–Crick sequence design. Synthetic oligonucleotides can be functionalized through nucleobase modification, alteration of the sugar–phosphate backbone, or site-specific attachment of fluorophores, affinity tags, reactive handles, and other molecular cargos^{20–25}. These chemical

modifications have enabled DNA nanostructures to serve as programmable platforms for biosensing, molecular imaging, targeted delivery, and the spatial organization of functional materials^{18,26–31}. Frameworks that utilize the programmability of DNA and diverse functionality of proteins have been reported, notably as a co-crystallization host-guest system for study of nucleic acid structures from the Georgiadis lab^{32,33}. We have recently reported how protein-DNA assemblies can be crystallized into porous assemblies with modular DNA struts adjacent to solvent channel pores³⁴. This protein-DNA co-crystal platform, referred to here as Co-Crystal 1 (CC1), is composed of repeating protein-DNA building blocks that tile into a three-dimensional crystalline lattice (space group no. 5, I121 or C121). In our recent report, we show that CC1 variants with 10 base pairs added between scaffold protein binding sites (CC1⁺¹⁰) can host DNA-binding protein guests soaked into the lattice post-crystallization. These soaked crystals organized proteins in defined positions, providing diffraction resolution as high as 3 Å, allowing side chain interactions of the guest protein to be modelled (pdb 9YZO)³⁴.

The DNA components of CC1 have programmable modularity that enables diversity in the sequence that abuts the solvent channel pores. In CC1, lattice formation is driven in one direction by protein-protein interactions between monomers of transcription factor RepE54, while DNA helices act as programmable struts that span protein nodes and expand the lattice in a plane lateral to the protein stack. This division of structural roles allows extensive modification of the DNA components while minimizing disruption of the lattice topology. By extending DNA struts in integer numbers of helical turns, CC1 lattices can be transformed from dense assemblies (CC1⁺⁰) into porous frameworks (interpenetrating CC1, CC1⁺¹⁰, CC1⁺²¹) without altering integral protein-protein contacts (Fig. 2.1). Previous work has shown that CC1 tolerates variation in DNA sequence, helix length, terminal phosphate positioning, and sticky-end lengths while

maintaining crystallographic order^{34–37}. These variations, however, focused on static, canonically paired B-form DNA and did not address the extent to that the CC1 family can accommodate chemically modified or structurally dynamic nucleic acid features. Herein we report the successful inclusion of multiple nucleic acid motifs into the CC1 platform, including modifications to single nucleotides, the DNA backbone, and larger changes to the secondary structure (Fig. 2.2). We introduce modifications spanning local base-pair imperfections, altered backbone chemistry, RNA hybridization, pendant fluorophores, and single-stranded DNA elements at defined positions within the DNA struts. By comparing the resulting crystal structures and lattice arrangements, we evaluate how the location and structural magnitude of each perturbation influence CC1 assembly, porosity, and crystallographic order. These studies define practical design constraints for incorporating functional nucleic acid components into protein–DNA crystalline materials and assess the potential of CC1 as a scaffold for the site-specific organization of otherwise flexible molecular features.

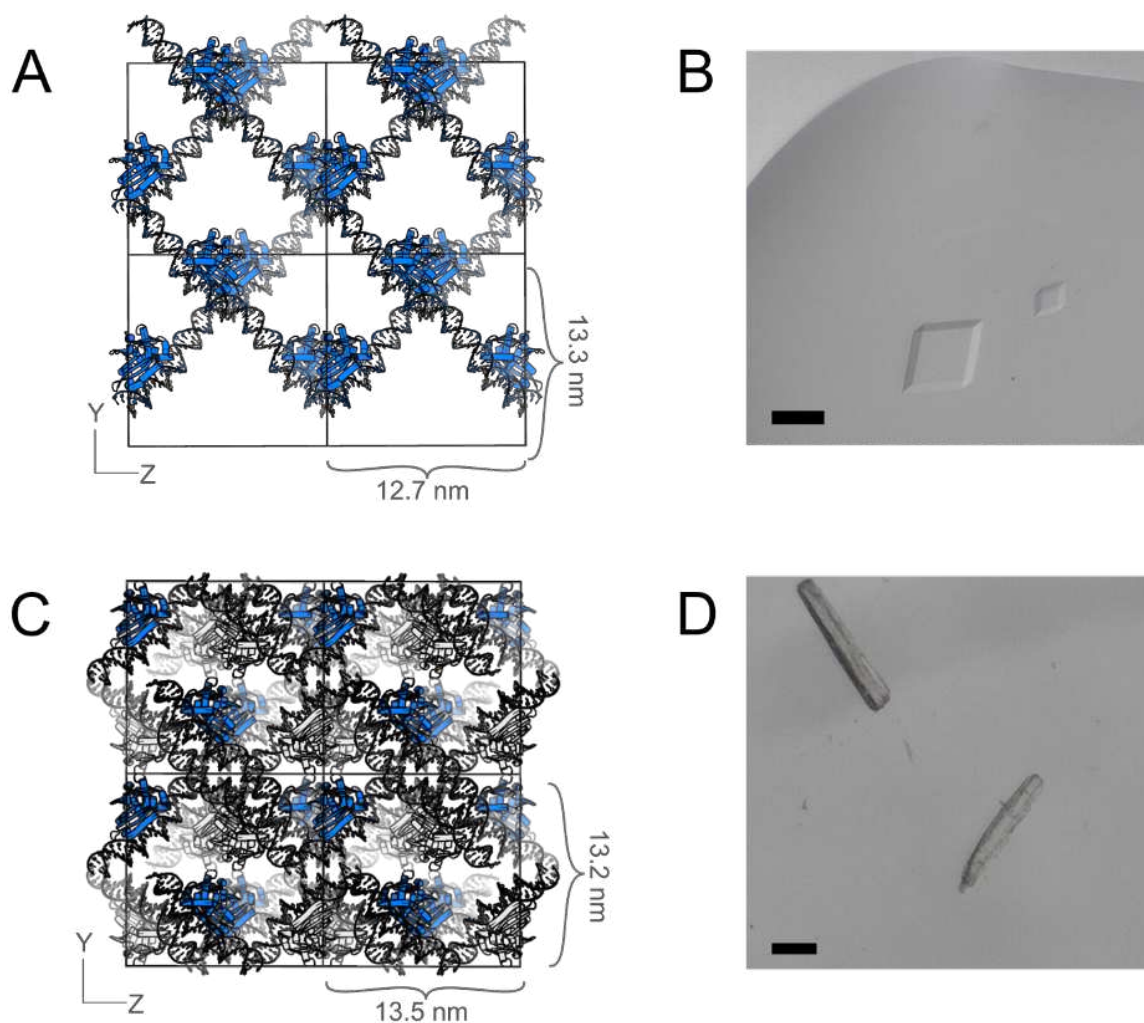


Figure 2.1: Lattice topology of porous CC1^{+10} and interpenetrating CC1^{+10} (ipCC1). Four adjacent unit cells. The solvent channels are sufficient to allow diffusion of macromolecules up to 5.3 nanometer diameter. **B.** CC1^{+10} parallelepiped habit in white light microscopy, scale bar is 100 μm . **C.** Four adjacent unit cells of ipCC1. A second copy of the target lattice (colored blue) that forms porous CC1 is grown in the solvent channel pores (second copy of lattice in grey), reducing porosity. **D.** White light microscopy of ipCC1 containing phosphorothioate modifications, reported in this manuscript.

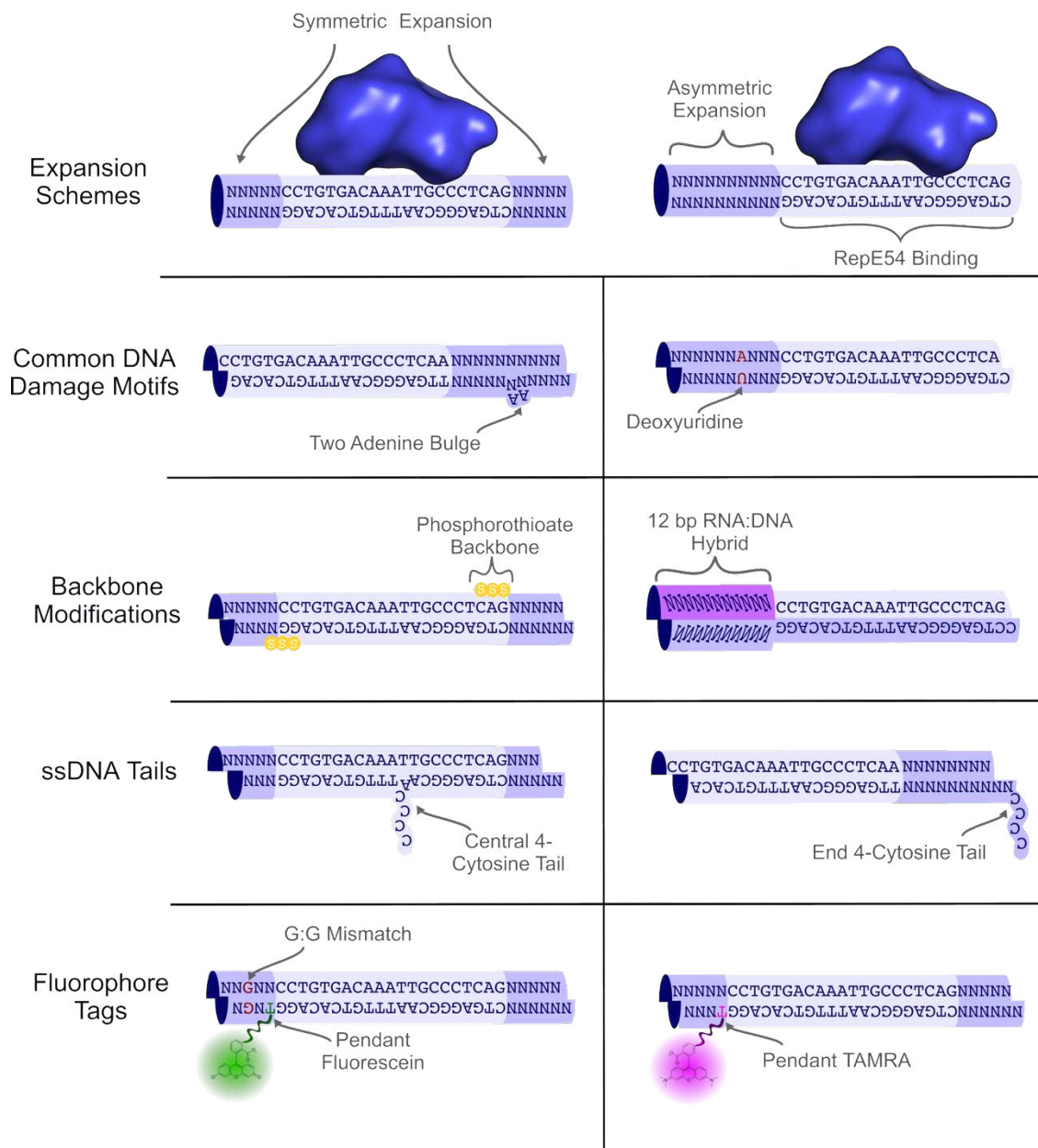


Figure 2.2: Modifications described herein, shown added to symmetric and asymmetric expansions of CC1. In dark blue, RepE54 binds a conserved 19 bp sequence, with 2 flanking nucleotides, shown in light purple. Expansion DNA sections are shown in dark purple. We report structures with simple DNA imperfection motifs, 2-nt bulge, deaminated nt dU, and G:G mismatch (included in the pendant fluorescein design). Backbone modification reported are phosphorothioate bond linkages and an RNA:DNA hybrid helix. We successfully incorporated pendant fluorophores fluorescein and TAMRA attached via linker to dT. Secondary structure motifs reported are ssDNA tails, placed both centrally and at the end of DNA building blocks.

2.2 Results and Discussion

The structures presented herein demonstrate a variety of modifications of the nucleic acid components within CC1⁺¹⁰. We begin with minimal changes to single bases or base-pairs and then proceed to larger changes that influence helical shape and composition. Finally, we show that CC1⁺¹⁰ is capable of site-specific display of nucleic acid secondary structures and detection of guest macromolecules via changes to the electron density maps.

2.2.1 Engineered bases in the expansion

2.2.1.1 Base-level perturbation motifs

To define the sensitivity of the CC1⁺¹⁰ lattice to minimal deviations from canonical B-form DNA, we introduced physiologically relevant imperfections in duplex DNA into the solvent-accessible expansion region. Specifically, we examined incorporation of a single dT to dU exchange, a G:G mismatched base pair, and a two-adenine bulge. Previous structures we reported have shown that sequence changes are tolerated along the “edges” of the RepE binding site (within four bases of the first or last base of the 21 bp binding site). In contrast, base identity changes positioned centrally within the RepE binding site have not yet given successful crystal growth, potentially due to abrogation of RepE binding (Table S2.17). Here we placed our test non-canonical motifs within the expansion region of the CC1+10 DNA struts (Fig. 2.2), a solvent-facing location that is spatially distinct from the RepE54 protein-DNA binding interface. Deoxyuridine introduces a base identity change without altering backbone connectivity, while G:G mismatches and small bulges locally perturb Watson-Crick hydrogen bonding and helical geometry. These DNA damage motifs are found in physiological DNA structures and are recognized with specificity by base excision repair machinery ^{38–44}.

Deoxyuridine and related C5-modified derivatives have been accommodated within several crystallographically characterized DNA duplexes. For example, PDB 2QEF contains 5-

(3-aminopropyl)-2'-deoxyuridine in a B-form duplex, where the modified base remains paired and directs its substituent into the major groove ⁴⁵. G mismatches have also been characterized crystallographically, often in directly stabilized or recognition-bound states. In a ligand-bound DNA structure, G pairs adopted multiple conformations, including syn–syn and anti–syn geometries, accompanied by substantial local backbone rearrangement (pdb 6J0I) ⁴⁶. Protein-bound examples show similarly perturbed conformations. MutS bends mismatch-containing DNA by approximately 60° and stacks a conserved phenylalanine against one mismatched base (PDB 1OH5), whereas DNA polymerase β captures a terminal G mismatch within its active site (PDB 5J2G) ^{43,47}. By comparison, our 3.25 Å structure contains a G mismatch near the center of the expansion region without direct protein or ligand stabilization or obvious lattice disruption, demonstrating that the framework can accommodate a centrally positioned purine–purine mismatch while preserving crystalline order. Previously reported two-nucleotide DNA bulges have likewise been characterized primarily in contexts where the unpaired residues are stabilized by a bound small molecule or protein. In PDB structures 2NEO, 1P96, 1R4E, and 2OEY, the bulged bases form defined pockets around wedge-shaped ligands and adopt distinct intrahelical or extrahelical conformations ^{48–51}. The CC and CT bulges in the RNase T complexes 4KB0 and 4KB1 are instead ordered through direct protein contacts, including insertion of a phenylalanine residue into the bulged region ⁵². In contrast, our 4.1 Å structure contains two consecutive adenines near the center of the expansion region without direct protein or ligand contact. Although the resolution prevents detailed assignment of the individual adenine conformations, the bulged segment is retained within the lattice and remains exposed to the solvent-channel pores. This structure therefore provides an example of a canonical two-base bulge accommodated within a porous protein–DNA crystal without direct host–guest stabilization.

Crystals of all three small change motifs were obtained using standard CC1⁺¹⁰ crystallization conditions. These crystals diffracted to resolutions comparable to canonically paired CC1⁺¹⁰ lattices, 3.21 Å for dU:dT, 3.25 Å for G:G, and a lower resolution 4.1 Å for bulge crystals (Tables S2.1-3). In all cases, crystals retained the same space group and approximate unit cell dimensions ($\pm 10\%$) as the parent porous CC1⁺¹⁰ structure, exhibited no new lattice contacts, and displayed continuous electron density throughout the expansion region. For the two-adenine bulge crystals the bulge was clearly visible in the electron density (Fig. S2.3). These structures indicate that the expansion regions of the lattice can absorb local base-level perturbations without large-scale loss of crystallographic order and without prevent crystal growth. To assess whether these modifications introduced measurable local distortions, we performed comparative helical parameter analysis using x3DNA. Specifically, we tabulated twist and rise, with particular interest paid to junctions between building blocks, comparing to averages from prior CC1⁺¹⁰ structures containing fully canonical base pairing^{34,53}. All three of the structures showed moderate overwinding at the interface between building blocks (Table S18). Interestingly, our prior canonical CC1⁺¹⁰ structures already had a slight overwinding of the helical twist when compared to ideal B-form DNA; the new structures reported here increased helical twist to a greater extent. For base pair steps excluding the junction, the average helical twist was within one standard deviation of canonical CC1⁺¹⁰ structures, implying helical strain is potentially absorbed at the DNA:DNA interfaces.

Our previous work showed the capability of CC1⁺¹⁰ to capture DNA-binders for structural study. Now, with the demonstration of feasibility for deviations from A:T C:G base pairing, this platform can likely be used for future studies of DNA repair proteins by programming their cognate damage target into the CC1⁺¹⁰ expansion.

2.2.1.2 Fluorophore tags

Since CC1⁺¹⁰ could accommodate small imperfections, we decided to determine whether the lattice could tolerate larger modifications. In the same G:G mismatch crystal, a chemically modified thymine displaying a fluorescein molecule via a flexible linker was included (Fig. 2.3). A second design included a TAMRA molecule at the same position (Fig. S2.3). These modifications have more steric bulk, are attached via flexible linker, and contain charged moieties.

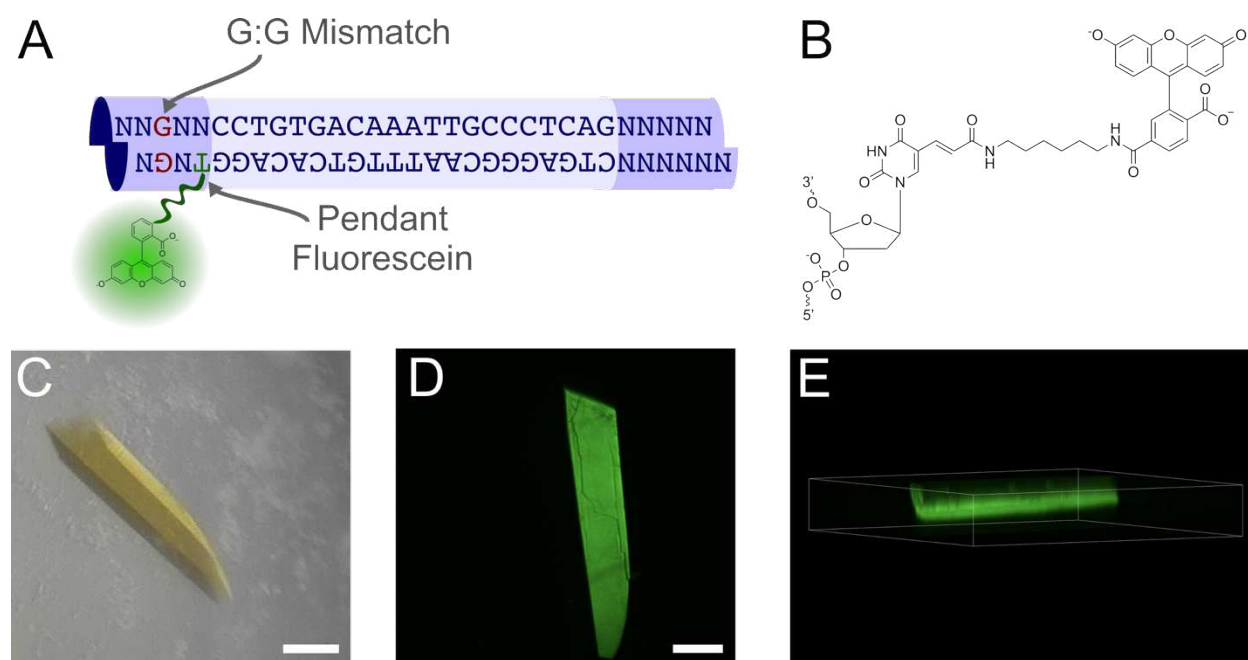


Figure 2.3: CC1⁺¹⁰ with fluorescein appended to each building block. **A.** CC1⁺¹⁰ building block showing symmetric expansion and nucleotide modifications G:G mismatch and pendant fluorescein, both in the same structure. **B.** Internal thymidine modification showing flexible linker and fluorescein moiety. **C.** White light microscopy, scale bar is 100 μm . Confocal microscopy image with 3% illumination by the 488 nanometer laser. **E.** Volume view of emission in all planes of the crystal shown in C and D. Notably, the emission intensity in the bulk of the crystal is non-uniform due to excitation extinction effects as explained in our previous work³⁴. Briefly, our inverted confocal microscope excites fluors from below, and high local fluor concentration leads to extinction as the excitation light advances through the crystal volume. This leads to bright emission along the planes closest to the microscope slide and along overhangs.

TAMRA-tagged building blocks grew into a needle-like habit, a deviation from the typical parallelepiped habit observed from porous CC1⁺¹⁰. These crystals adopted the interpenetrating lattice (ipCC1) formation, with a second copy of the “target” porous lattice growing in the pores (Fig. 2.1). This CC1 variant diffracted to a resolution of 4.2 Å and showed no obvious density corresponding to the flexible linker or TAMRA molecule (Table S2.4). While we hesitate to speculate on the driving force of lattice assembly between ipCC1 and porous CC1, we note that the TAMRA molecule is positioned in the same region of the strut (T27) as interfacial magnesium ions are observed in ipCC1 structures, a phenomenon similarly noted in the phosphorothioate modified structure described below.

CC1⁺¹⁰ crystals containing the fluorescein-modified thymine (T27) readily formed under standard crystallization conditions and diffracted to resolutions comparable to unmodified porous CC1⁺¹⁰ lattices (Table S2.2). The crystals retained the same space group and approximate unit cell parameters as porous CC1, and no new lattice contacts were observed. Continuous electron density was present along the DNA backbone of the expansion region, indicating that incorporation of the pendant modification did not disrupt the helical structure. In addition to well-defined helical DNA density, strong and localized electron density was observed adjacent to the modified thymine, extending beyond the nucleic acid base by approximately 3 Å (Fig. 2.4 A). Similarly, the Fo–Fc difference map contained a strong positive feature at this site that extended beyond the modeled thymine base. Farther from the DNA, near the crystallographic 2-fold axis, a large electron-density feature was present in both the 2Fo–Fc and Fo–Fc maps. We attribute this feature to the proximity of modified thymine 27 to a symmetry mate generated by the crystallographic symmetry operator ($-x + 1, y, -z + 1$). Previous studies have shown π -stacked fluorescein arrangements inferred from powder XRD studies of fluorescein co-crystals, though

single crystal XRD of fluorescein stacks have not been reported^{54–56}. To interrogate the electron density map feature that was unique to the pendant fluorescein crystals, we proceeded to model fluorescein interactions with independent conformational sampling and energy minimization quantum chemical calculations in solution^{57–59}. This conformational modeling suggested that π – π interactions between fluorescein were plausible (Fig. S2.6). Modelling of the fluorescein interacting revealed close apposition of fluorescein moieties, with a symmetry copy across the two-fold axis, with a staggered geometry consistent with π – π interactions between the xanthene cores (Fig. 2.4).

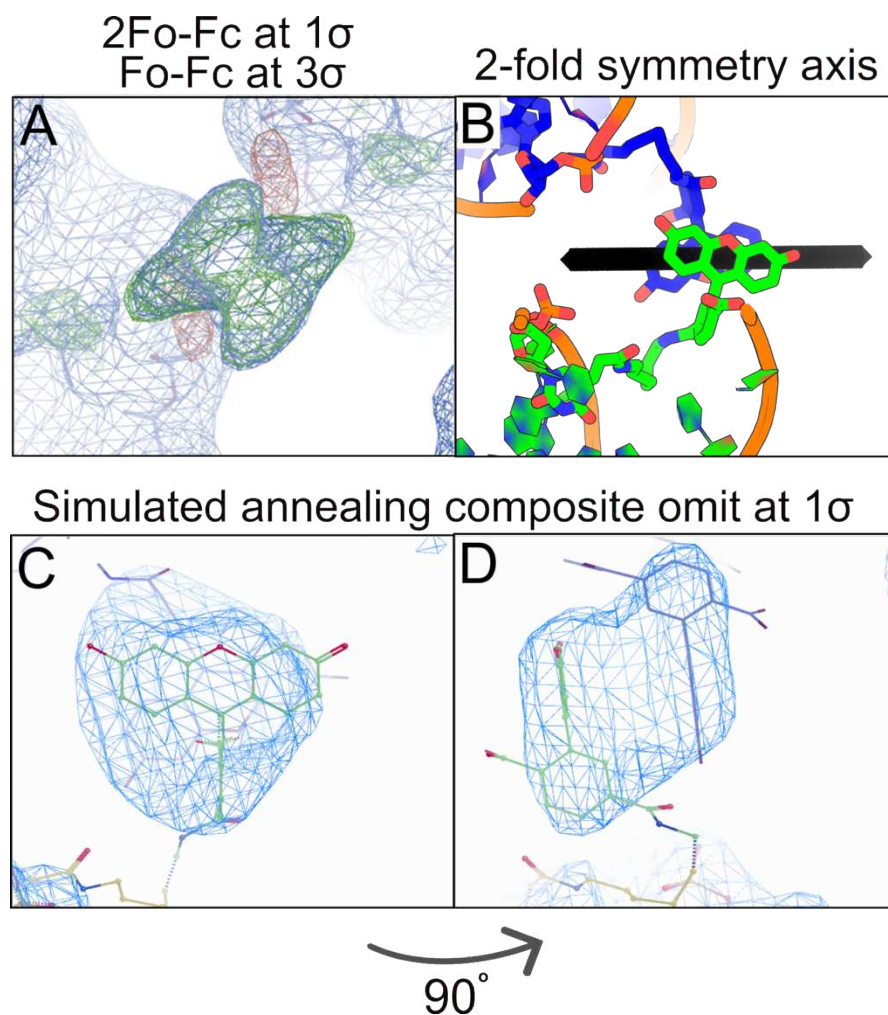


Figure 2.4: Putative fluorescein stacking. **A.** Composite 2Fo–Fc and Fo–Fc map showing observed density adjacent to the modified thymine, contoured at 1 σ and 3 σ , respectively. **B.**

Symmetry axis (black) between two asymmetric units of CC1⁺¹⁰ when fluorescein is modeled into the density shown in panel A. **C.** and **D.** Simulated annealing composite omit 2Fo-Fc map of π -stacked fluorescein contoured at 1 σ .

To our knowledge, this model explains the first electron density feature that can be ascribed to fluorescein-fluorescein aggregation and staggered π - π stacking in single crystal X-ray crystallography. Structural fixation of small molecules *in crystallo* and their subsequent structure determination without requiring *de novo* crystallization of the analyte is a highly desirable capability in chemical structural analysis^{7,60–62}. The fluorescein example here suggests that lattice contacts within CC1⁺¹⁰ may provide a method to observe weak or transient interactions that would be difficult to capture in conventional single-crystal experiments on the small molecule alone. As with the classic crystalline sponge method of Fujita¹⁴, observing such structures requires adventitious interactions. This capability underlies later demonstrations of guest molecule organization within CC1⁺¹⁰ and hints at the possibility of scaffold-assisted structural observation of nucleic acid-associated small molecules and interactions.

2.2.2 Backbone modification and RNA:DNA hybrid

To further probe the chemical tolerance of the CC1⁺¹⁰ platform at the level of the nucleic acid backbone, we next introduced phosphorothioate (PT) linkages into the DNA expansion region. PT modifications replace a non-bridging phosphate oxygen with sulfur, altering duplex stability, backbone electronegativity, torsion angles, and metal coordination behavior while preserving overall duplex architecture^{63–65}. Six PT modifications were incorporated into the expansion region, with three modified linkages on each strand. Because commercially produced PT bonds are a racemic mixture of Rp and Sp stereoisomers, this design ensures a high probability that Rp linkages (the configuration that stalls exonuclease action) are present along

both strands of the duplex ⁶³, although should future work require stereopure DNA to incorporate into crystals, these materials are becoming more accessible ⁶⁶.

Crystals containing phosphorothioate-modified DNA formed reproducibly under standard CC1⁺¹⁰ crystallization conditions (Table S2.14). Successful incorporation of PT bonds opens the door for future applications in which stabilized host crystals might resist exonuclease *in vivo* thereby enabling the tuning of extended-release materials. Notably, these crystals adopted the interpenetrating CC1⁺¹⁰ lattice (ipCC1) in space group I222 rather than the porous I121 lattice. The differences in crystallization conditions leading to the two CC1 crystal packing schemes, leading to the porous CC1 lattice versus the interpenetrating CC1 lattices, was explored extensively in our previous work ³⁴. Typically the ipCC1 architecture was observed when crystallization conditions contain high levels of magnesium (Mg II) ^{34,35}. Despite this change in crystallographic topology due to PT bonds, the crystals diffracted to resolutions comparable to previously reported ipCC1 structures ^{35,36} and exhibited well-ordered DNA electron density throughout the expansion region. Phosphorothioate bonds replace a smaller, less polarizable oxygen with a larger, more polarizable, less negatively-charged sulfur along the DNA backbone, that may alter electrostatic interactions within the lattice. We hypothesize that such changes to the physical chemical character of the backbone could favor tighter packing of neighboring building blocks, biasing assembly toward the denser ipCC1 architecture. Interestingly, the sulfur-containing linkages are located at the same interface where we previously observed magnesium ions to drive stabilization of the interpenetrating lattice. We therefore hypothesize that charge dispersion by either ions or polarizable groups along the backbone may drive ipCC1 formation over porous CC1 (Fig. S2.5). The observed shifts from porous to interpenetrating lattice suggests that backbone chemistry can influence packing preferences during assembly of the CC1⁺¹⁰

system. The ready accessibility of the interpenetrating lattice upon modest changes to the building blocks underscores the continuing need for robust methods to drive formation of the preferred lattice for a given application. Future work may test alternate placements for phosphorothioate modifications, which could restore porous CC1 formation. From a platform perspective, the ability to incorporate phosphorothioate-modified DNA expands the chemical design space of ipCC1 and suggests potential routes for enhancing lattice stability in nuclease-rich environments. Tunable stability could ease adaptation of the CC1 platform to biomedically relevant conditions.

We next examined whether the lattice could accommodate more substantial deviations from canonical B-form DNA helix geometry. B-form DNA completes a helical turn every 10.5 bases (46, 47). Accordingly, CC1 is typically expanded in integer numbers of B-form helical turns ($CC1^{+10}$, $CC1^{+21}$). To push the CC1 variant frontier beyond DNA-only expansions, we introduced an RNA:DNA hybrid helix into the expansion region of the CC1 DNA struts. RNA:DNA hybrids tend to adopt an A-form-like geometry characterized by increased x-displacement, increased helical diameter, and altered backbone torsion angles relative to B-form DNA, representing a significant structural perturbation beyond isolated base modifications^{67–70}. We implemented RNA-DNA hybrid expansion using a three-strand design (Table S2.16) in which the RepE binding site remained a DNA:DNA duplex, preserving native protein-DNA recognition, while one strand was extended beyond the protein interface to form an RNA:DNA hybrid region with 12 nucleotides of RNA (Fig. 2.5). These building blocks had a single sticky base overhang at the edges.

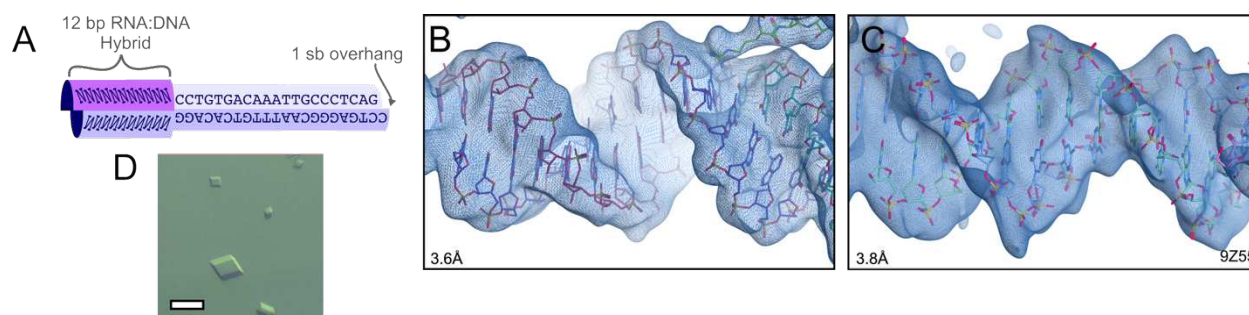


Figure 2.5: RNA:DNA hybrid crystals. CC1⁺¹⁰ building block showing asymmetric expansion in which the bottom strand is contiguous and hybridized to two separate top strands, DNA:DNA for the RepE54 binding footprint and RNA:DNA facing the solvent channel pores. B) RNA:DNA hybrid expansion composite 2Fo–Fc map contoured at 1 σ . C) Comparison DNA:DNA asymmetric expansion composite 2Fo–Fc map contoured at 1 σ .

The RNA:DNA hybrid expansion required adjustment of expansion length to compensate for the tighter helical rise characteristic of A-form-like structures. Specifically, we attempted to crystallize designs containing 9-13 bp of RNA:DNA hybrid expansions, with both one and two sticky base overhangs (Table S2.17). Of these attempts, the lone design to crystallize to date was a 12 bp hybrid expansion with a single sticky base overhang (CC1⁺¹²), an interesting result as literature reports a range of helical pitches/base pairing numbers as complete helical turns of RNA:DNA hybrids^{70–72}. CC1⁺¹² crystals containing the RNA:DNA hybrid expansion formed under standard crystallization conditions (Table S2.14) and diffracted to medium resolution (~ 3.6 Å), comparable to porous CC1⁺¹⁰ lattices containing B-form dsDNA expansions. The crystals retained the expected space group and unit cell parameters, and no new lattice contacts were observed. Electron density within the expansion region clearly resolved the RNA:DNA hybrid duplex (Fig. 2.5). The RNA expansion region adopted a conformation most consistent with the published hybrid structure PDB 479D⁷³. Among five reference RNA structures evaluated (404D, 479D, 1FIX, 9VUN, and 7BPV), 479D provided the best fit to the simulated-annealing omit density for the strut region^{67,73–76}. The initial AlphaFold3 model was therefore morphed toward the 479D geometry in Phenix to generate the final model of the expansion region. Helical geometry examination with x3DNA showed tighter helical twist than a pure-

DNA CC1⁺¹⁰ strut, and negative base slide and x-displacement parameters, more in line with the 479D RNA:DNA hybrid structure (Table S2.19)⁵³. RNA:DNA hybrids are recognized by a distinct class of proteins involved in transcription, replication, and genome maintenance^{77–79}. Therefore, the ability to scaffold RNA-DNA hybrid struts could be useful to capture guest proteins from these families.

2.2.3 ssDNA tail motifs

To explore whether porous CC1⁺¹⁰ lattices could be used to host larger, more diverse nucleic acid motifs, we introduced short ssDNA strands, henceforth referred to as DNA tails, into both the expansion region and opposite the RepE54 binding site (Fig. 2.2). These tails are solvent-exposed, and lack intrinsic structure in the absence of binding partners. The Mao lab and coworkers have shown that ssDNA motifs can be utilized for dynamic processes within tensegrity triangle crystals and provide fixation and release of co-crystallized and post-crystallization DNA motifs^{17,80–82}. Much like these pure-DNA designs, CC1 takes advantage of the porosity of the structure to include functional elements that might otherwise disrupt crystal packing. Of interest to structural biologists, inclusion of ssDNA accessible to the solvent channel pores could provide binding sites for ssDNA-recognizing enzymes and thereby support their structural observation^{83–85}.

The first designs toward this goal included a 6 nt poly-T tail at the end of asymmetric units (“end tailed”) and underneath the RepE54 binding site on the opposite side of the helix from the protein-DNA interface (“centrally tailed”). Crystals formed with both designs, but showed crystal packing that deviated from the expected CC1 contacts (Tables S2.11, S2.12). Where the CC1 platform typically contains two protein:protein interfaces, one protein:DNA interface, and DNA:DNA contacts, we observed preserved DNA:DNA and DNA:protein

interfaces, but only one of the target protein:protein interfaces was retained. The other protein-interface was replaced by a new interface between the ssDNA poly-T tail and RepE54 (Fig. S2.1). Remarkably, despite the crystal packing topology differences for centrally and end tailed crystals ($C222_1$ versus $P4_12_12$), these same interfaces were present in both crystal types (Fig. S2.1). Retroactively, we found an explanation for the new interface in a 2023 study from Wegrzyn *et al.* that showed that RepE54 binds poly-thymine ssDNA specifically and concurrently to dsDNA during its physiological function in bacterial cells⁸⁶. Wegrzyn *et al.* reported that RepE54 binds poly-thymine ssDNA and that this interaction can be disrupted by the substitutions F146E and Y172E. Because the CC1 lattice is assembled through the canonical binding of RepE54 to its cognate double-stranded DNA recognition site, rather than through binding to the appended ssDNA tails, we reasoned that suppressing the poly-T interaction could reduce unintended sequestration or conformational restriction of the tails without eliminating the protein–DNA contacts required for crystallization. We therefore introduced the corresponding Y184E substitution into the CC1 RepE54 construct; the residue number differs because of the N-terminal tags present in the CC1 protein (Fig. S2.2). However, constructs containing either centrally positioned or terminal poly-T tails did not yield crystals with Y184E RepE54. Thus, disruption of the reported poly-T-binding interaction was not sufficient to rescue crystallization of these designs, suggesting that the poly-T tails themselves, or another consequence of their incorporation, remained incompatible with productive lattice formation. Subsequent constructs therefore used ssDNA tails with reduced thymine content, while retaining the Y184E substitution as a precaution against residual nonspecific ssDNA binding by RepE54. Evaluation of the F158E/Y184E double mutant and additional tail compositions remains an area for future work.

Two subsequent ssDNA tail designs were attempted, each consisting of a 4-nt poly-C extension placed centrally (across the helix from RepE54 binding) and at asymmetric unit ends. Crystals containing either ssDNA tail geometry formed readily and diffracted with space group and unit cell parameters consistent with canonical porous CC1⁺¹⁰ lattices (Tables S2.7-S2.10). As expected, in both cases, electron density corresponding to the ssDNA tails was weak or absent, with the first unpaired nucleotide of the end tail and first two unpaired nucleotides of the central tail exhibiting partial density (Fig. 2.7).

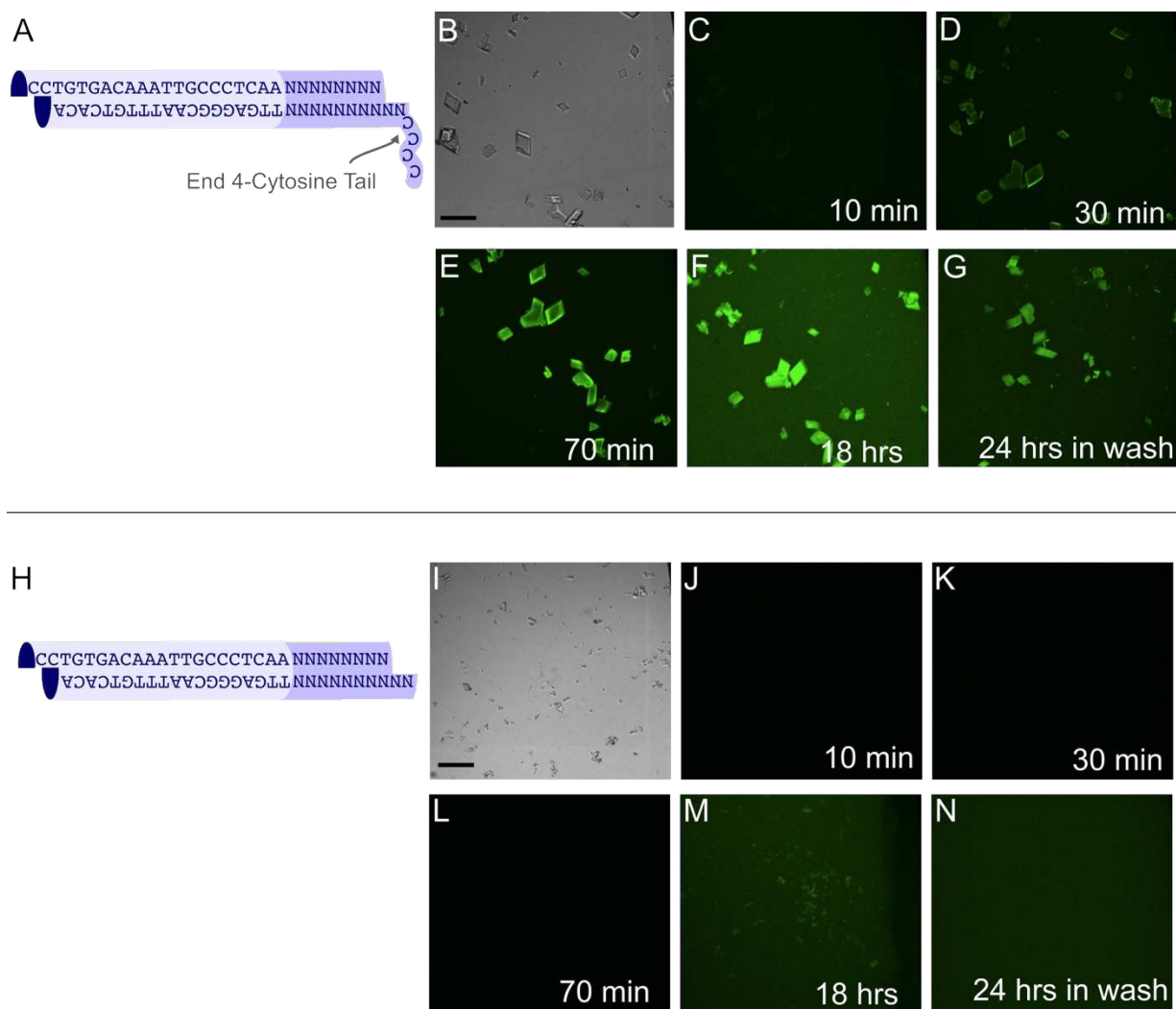


Figure 2.6: Fluorescent oligo loading in end-tailed CC1⁺¹⁰. Schematics in **A.** and **H.** show an end-tailed poly-C CC1⁺¹⁰ and an asymmetric CC1⁺¹⁰ with no tail, respectively. **B.** and **I.** show

white light images of microcrystals of A and H, respectively. Scale bar is 100 μm . Subsequent confocal microscope images are under a 488 nm laser at 3% power in a z-plane through the middle volume of a majority of the microcrystal field. **C-F.** and **J-M.** are time points after addition of a short, fluorescently-tagged oligo guest (FAM-AAGGG) at 10 μM concentration. **G.** and **N.** show the microcrystals 24 hours after the surrounding solution has been swapped for a wash containing no oligo guest.

To assess whether disordered ssDNA tails could capture guest nucleic acids within the CC1^{+10} lattice, crystals containing end C_4 tails were ligated using 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) as described in our previous work^{36,37}. We probed this system with an on-hand fluorescent probe, FAM-AAGGG, that contains three bases complementary to the C_4 tails. Confocal fluorescence microscopy revealed time-dependent accumulation of fluorescence within crystals containing the ssDNA tail, whereas control CC1^{+10} crystals lacking ssDNA tails showed weaker accumulation after 18 hours. After extended incubation (~ 18 hrs), and subsequent transfer to fresh wash solution, ssDNA-containing crystals displayed uniform fluorescence throughout the crystal volume, while control crystals equilibrated (after ~ 24 hr) to the fluorescence intensity of the surrounding solution (Fig. 2.6).

Crystals subjected to guest loading were subsequently analyzed by XRD (Fig. 2.7). Centrally placed ssDNA tails extend into the solvent channel pore with minimal surrounding atoms from symmetry mates. After loading, no additional electron density attributable to the guest oligonucleotide was observed, and the structure was indistinguishable from that of the unloaded control. In contrast, guest loaded crystals containing the end ssDNA tail exhibited strong, localized electron density adjacent to the tail region, of stark contrast to the unloaded control. This density was related by crystallographic symmetry to an equivalent feature on a neighboring lattice repeat. Modeling of two pendant nucleotides yielded a chemically and sterically reasonable fit to the density, consistent with partial ordering of the tail through base

pairing and lattice-imposed symmetry constraints. Unfortunately, proximity to the twofold axis makes it difficult to fully build into the positive difference map features. Obvious electron density features appeared only when a guest nucleic acid was introduced. While we could not directly model the guest in this instance, this clear electron density changes nonetheless hints at the future possibility of scaffold-assisted structure observation of guest molecules. We suspect that fortuitous lattice contacts will only be sufficient to drive structure formation on rare occasions; perhaps dual anchoring of guest nucleic acids will more often be sufficient for routine guest observation. The scaffold ssDNA provides a programmable recognition element that recruits a complementary guest oligonucleotide, while the crystal lattice fixes a subset of otherwise dynamic interactions into a resolvable arrangement. Importantly, structural ordering was observed only for one ssDNA tail geometry, indicating that successful resolution depends on placement within the lattice and local symmetry contacts. This geometry dependence highlights the importance of scaffold design in observing changes to the electron density map and thereby resolving dynamic guests.

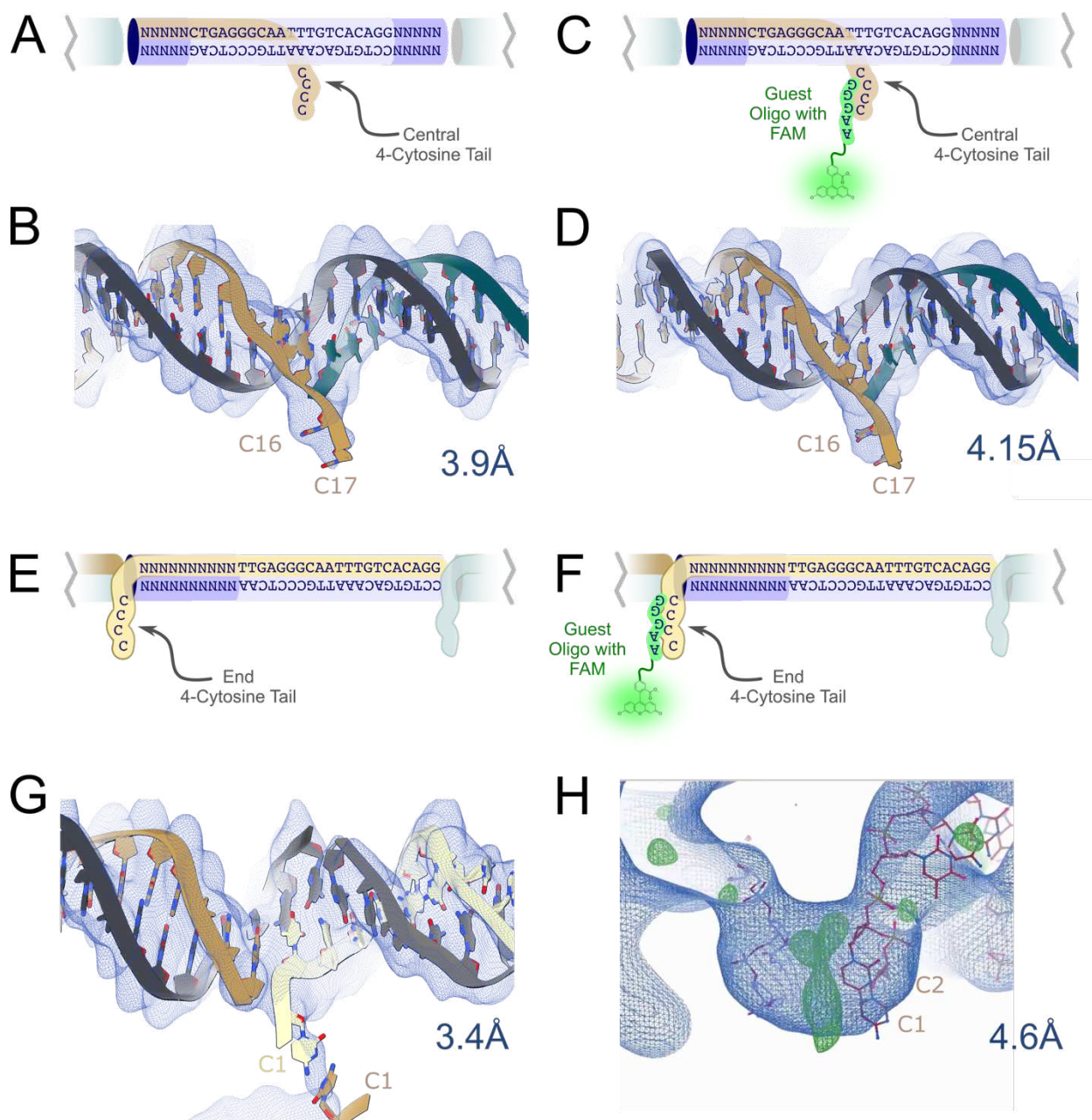


Figure 2.7: Post crystallization loading of oligo guest and subsequent structural observation. **A.** and **B.** show a central C₄ tailed crystal in the absence of oligo guest, the first and second cytosine in the tail could be modelled. **C.** and **D.** show a central C₄ tailed crystal after soaking with 100 μ M short oligo guest, two cytosines can be modelled with no additional observable density for dsDNA from the guest binding. **E.** and **G.** show an end C₄ tailed crystal in the absence of an oligo guest, with one of the tail cytosines modelled. **F.** and **H.** show an end C₄ tailed crystal after soaking with 100 μ M short oligo guest (GGGAA-FAM), with additional density visible corresponding to the entire ssDNA tail, with potential for modelling the dsDNA from the guest binding and/or an interaction across symmetry mates.

This design is reminiscent of the DNA-only crystals developed in the Mao lab that were shown to bind environmental DNA and trigger release⁸⁰, as the dangling strands invite the possibility of the DNA device work. Importantly, DNA origami crystals and colloidal crystals driven by DNA hybridization tend to be more suitable for small angle x-ray scattering due to both limitations in crystal size and structural heterogeneity of the unit cells⁸⁷. These results show that CC1⁺¹⁰ lattices can support post-crystallization loading, retention, and partial structural ordering of guest nucleic acids through rational scaffold design. One route to improve these scaffolds for guest nucleic acid capture may be to extend the length of the DNA tails beyond four nucleotides, allowing for programming of specific capture. Structural observation via XRD provides verification of high occupancy installation of the guest nucleic acid in the crystal pores. This proof of principle positions our platform for downstream biomedical applications (*e.g.* site-specific capture of therapeutic cargo nucleic acids) as well as nanotechnology (*e.g.* monitoring programmed DNA strand pairing with XRD). This capability underscores the versatility of CC1⁺¹⁰ as a modular platform for crystalline organization of functional nucleic acids.

2.3 Conclusions

Here we expand the CC1 platform beyond canonical DNA duplexes to explore tolerance for engineered nucleic acid chemistries, altered helical geometries, and locally distorted motifs that interface with the solvent channels. By introducing these motifs into defined regions of the CC1⁺¹⁰ DNA struts, we probe the limits of lattice robustness and reliable self-assembly under increasing degrees of chemical and structural perturbation. In 6 out of 10 structures presented here, the target porous protein-driven lattice was preserved despite these perturbations. In two of the structures, modifications in the expansion region (phosphorothioate bonds and pendant TAMRA) yielded the interpenetrating version of CC1 (Fig. 2.1 C and D). The two remaining

structures contained a ssDNA poly-T tail that led to crystallographic contact changes, adding a new ssDNA:protein interface (Fig. S2.1). A recent report on the physiological functions of RepE54 identified a propensity for binding T-rich ssDNA as part of the RepE54 function to melt genome elements upstream of the 19 bp dsDNA binding sequence⁸⁶. Importantly, a single amino acid mutation in RepE54 and a new design principle that avoids 4 or more consecutive thymine nucleotides were able to rescue expected porous crystal packing for ssDNA designs. This study also revealed the site-specific display of guest macromolecular features, visible in XRD electron density maps. Beyond the guest DNA-binding proteins we have previously reported, here we show structural observation of ssDNA elements and small molecule interactions³⁴. The combination of solvent accessibility and programmable nucleic acid elements enables organization of flexible or weakly interacting components through symmetry mate contacts imposed by the scaffold, allowing partial observation via XRD. This work expands the library of viable CC1 variants and lays the groundwork for future applications in scaffold-assisted structure determination and rational crystalline material design. The ability of CC1 to incorporate RNA:DNA hybrids highlights an important feature of the platform: tolerance to systematic changes in helical geometry, not merely localized base-level defects. This capability extends the functional scope of CC1 beyond pure-DNA duplexes and positions the platform as a scaffold for studying nucleic acid architectures relevant to RNA biology.

By systematically introducing non-canonical bases, mismatches, bulges, pendant small-molecule modifications, RNA:DNA hybrid helices, phosphorothioate backbones, and single-stranded DNA elements into solvent-exposed regions of the CC1⁺¹⁰ lattice, we demonstrate that these scaffold crystals can accommodate substantial perturbations without loss of lattice integrity and without impeding the delicate process of self-assembly. Additionally, we demonstrate the

capability of CC1⁺¹⁰ to provide visible electron density map changes that emerge when otherwise conformationally heterogeneous elements (*e.g.* pendant fluorophore or pendant ssDNA) interact with symmetry mates and/or guest molecules. In general, dynamic guest domains were not visible if their only constraint was tethering to one site on the scaffold crystal. However, in the presence of at least one additional interaction, density map features emerged. We therefore hypothesize that scaffold-assisted structural observation via CC1 will benefit from multiple rational interactions (*e.g.* dual anchoring). For both structures, interactions with symmetry mates enable structural observation of dynamic motifs that would be difficult to access through direct crystallization of the guest species alone.

Our work adds to the toolbox of programmable crystalline scaffolds, and shows that high porosity, solvent accessibility, and atomistic structural resolution can coexist within a single crystalline material. By expanding the chemical and structural design space of nucleic acids that can be incorporated and structurally resolved within a porous lattice, this work lays a foundation for future efforts in scaffold-assisted structure determination and the rational design of functional protein–DNA crystalline materials.

2.4 Materials and Methods

2.4.1 Protein cloning, expression, purification

Protein sequences located in Table S2.15. Cloning and expression of the protein Replication initiator RepE54 was previously described by Komori *et al.*⁸⁸. Briefly, we cloned RepE54 into pSB3 vector (Addgene plasmid #82027) with an N-terminal His-tag. The CC1 RepE54 was expressed and purified by the Histone Source at Colorado State University as described in our previous ligation study³⁶. In short, the N-terminally hexahistagged protein from PDB code 7RVA was overexpressed in E.coli CodonPlus RIPL competent cells (Agilent #230280). Sonicated cell lysate was purified with Ni Excel Sepharose (Cytiva #17371202) and HiLoad Superdex 200 PG column (Cytiva). The resulting CC1 protein was concentrated to 15 mg/mL in storage buffer (100 mM sodium citrate pH 6.2, 100 mM KCl, 10 mM MgCa₂, and 10% glycerol). Aliquots were flash frozen with liquid nitrogen, and stored at -80°C. This variant of the RepE54 protein was used for all crystal structures other than Y184E (that used the point mutant described immediately below).

In hopes of disrupting any favorable intermolecular contacts at the ssDNA binding cleft identified in pdb 8AAN we added a negatively charged side chain⁸⁶. Variant RepE54 (Y184E) was cloned into the pETDuet plasmid (Novagen, Sigma-Aldrich # 71146-3) using Gibson cloning (Hi-Fi assembly New England Biolabs #E2621L). The protein was overexpressed with a T7 promoter in E. coli BL21(DE3) competent cells (New England Biolabs #C2527H). Upon addition of 0.5 mM IPTG, the cells were incubated at 25 °C for 20 h. The cell pellets were sonicated in lysis buffer (1X PBS, 300 mM NaCl, 25 mM imidazole, pH 7.4) and applied to HisTrap (HisPur™ Ni-NTA resin, ThermoFisher #88222) equilibrated with HisTrap buffer (1X PBS, 300 mM NaCl, 25 mM imidazole, pH 7.4). The protein was eluted with 100 mM imidazole

in HisTrap buffer. Fractions containing the desired protein were pooled and further purified at the CSU Histone Source facility with HiLoad Superdex 200 PG column (Cytiva #28989335). RepE54 variant (Y184E) was concentrated to 15 mg/mL in CC1 storage buffer (100 mM sodium citrate pH 6.2, 100 mM KCl, 10 mM MgCa₂, and 10% glycerol), flash frozen in liquid nitrogen, and stored at −80°C.

2.4.2 Duplex annealing and co-crystal complexing

Individual oligomers were synthesized and HPLC purified by Integrated DNA Technologies. The DNA oligomer sequences used for co-crystallization are listed in Tables S2.1 and S2.4. With limited exceptions, each oligomer contained the 19-bp iteron sequence TGTGACAAATTGCCCTCAG, but typically also contained the core 21-bp DNA block used in PDB:1REP from Komori *et al.*⁸⁸: CCTGTGACAAATTGCCCTCAG (omitting a dangling 3' T). In Table S16, the standard 21-bp footprint is shaded in gray, and (purposeful) deviations are shaded in green. Flanking variable DNA sequences depended on the expansion scheme and desired sequence. Notable modifications are marked with underlines and italics. The oligomers were resuspended in a standardized buffer (henceforth CC1 oligo buffer: 50 mM Tris HCl, 100 mM KCl, pH 6.7) and combined in equal molar ratio (1:1) with the complementary strand. The strands were annealed by heating to 94 °C for 2 min and slowly cooling to room temperature over approximately 60 minutes. The final concentration of all CC1 dsDNA blocks following duplexing was 4 mM. The scaffold protein and DNA were mixed in a 1.2:1 ratio (440:364 μM, 55 μL total) and were incubated on ice for 30 minutes prior to crystallization via sitting drop vapor diffusion.

2.4.3 Crystallization

Crystallization trials used CrysChem sitting-drop plates (Hampton Research HR3-159), with 400 μ L reservoir solution per well and 4 μ L crystallization drops prepared by mixing 2 μ L each of the protein–DNA complex and reservoir solution directly on the platform; drops were mixed by gentle pipetting. Plates were sealed immediately after setup and incubated at 25 °C under static conditions. Crystallization conditions were 10–80 mM Magnesium Acetate, 0.3–1.8 M Lithium Sulfate, and 50 mM MES pH 6.5. A detailed list of individual crystals, PDB codes, and corresponding growth conditions are reported in Table S2.14. Crystals grew to full size in a range of 24 hours to 30 days (most often in 1–2 days). Crystal images were taken with a Moticam X5 Plus camera attached to a Motic SMZ-168 stereozoom microscope.

2.4.4 Crosslinking

Co-crystals were washed in conditions similar to their growth conditions (0.3–0 M Lithium sulfate, 20–50 mM Magnesium acetate, and 50 mM MES, pH 6.5). The co-crystals were washed in 200 μ L on 9-well glass plates (Hampton) to remove additional protein and DNA monomers. Note: For EDC ligation it is important to avoid buffers with 1°/2° amines or carboxylic acids. 1-Ethyl-3(3-dimethylaminopropyl)carbodiimide (EDC) (Advanced Chemtech CAS#:25952-53-8) was resuspended in the crosslinking solution (1.0 M Lithium sulfate, 50 mM Magnesium acetate, and 50 mM MES, pH 6.0) to final concentration values of 30–60 mg/mL and used immediately. The co-crystals were crosslinked in a 200 μ L EDC solution volume overnight (16–24 hrs). The crosslinking reaction was ended by looping the co-crystals back into the wash solution.

2.4.5 Confocal Microscopy

Crystals loaded with fluorescently labeled oligos were imaged using a Nikon Eclipse Ti Yokagawa CSU-X1 spinning-disk confocal microscope equipped with an Andor iXon Ultra 879U

EMCCD camera. Fluorescein signal was collected using 488 nm laser excitation and a 515 nm emission filter. TAMRA signal was collected using 561 nm laser and a 575 nm emission filter. Image acquisition and processing were performed using Nikon NIS-Elements software.

2.4.6 X-ray Diffraction collection and refinement

Crystals were flash frozen in liquid nitrogen after briefly transferring the crystal to a cryo-protectant solution (a 1:1 mixture of the reservoir solution and 50% glycerol stock (1:1 100% glycerol and diH₂O) for a final concentration of 25% glycerol). Single crystal X-Ray diffraction data was collected at either ALS beamline 4.2 (CMOS detector), ALS beamline 8.2 (Pilatus3 detector), ALS beamline 8.1 (Dectris EIGER2 detector), or APS beamline 24-IDE (Dectris EIGER2 detector) from 0 to 180° with an omega delta of 0.2° and an exposure of 0.5 s or 0.25 s. Data was processed with XDS ⁸⁹ and refined using PHENIX and COOT ^{90,91}. The first porous CC1+10 dataset (PDB 9Z08) was refined using newly generated RFree flags ³⁴. All subsequent isomorphous structures were refined using the same RFree flags to preserve a fixed cross-validation benchmark across datasets. Decisions about truncating the structure factor data were made on a case-by-case basis, jointly considering the following thresholds $CC1/2 > 0.3$, $\text{mean } \langle I/\sigma(I) \rangle > 1.0$, and completeness $> 85\%$ as target guidelines for the highest-resolution shell. Most datasets satisfy all three thresholds in the highest-resolution shell. TLS restraints were autogenerated in Phenix and included in REMARK comments in pdb entries. RNA:DNA hybrid structure was morphed towards 479D reference structure using phenix.morph and end C4 loaded with fluorescent guest structure was modelled computationally in house python scripts.

2.4.7 Quantum Chemistry and Posing Comparison

Conformational sampling for the di-anionic form of the fluorescein dimer was done using the global optimizer algorithm (GOAT) within ORCA version 6.0.0 using xTB version 6.7.0 at

the GFN2-xTB semi-empirical level of theory with implicit solvation in water using the analytical linearized Poisson-Boltzmann (ALPB) solvation model^{92–102}. This initial conformer search gave an ensemble of 49 conformers. The conformational ensemble output from the GOAT calculation was then filtered for duplicate structures using the CREGEN module of CREST version 3.0^{103–105}. Filtering was done with an energy window of 25 kcal/mol, an RMSD threshold of 0.8 Å, an energy threshold between conformer pairs of 0.8 kcal/mol, and a lower bound for the rotational constant of 0.08 (8%). After filtering, the conformational ensemble for the di-anion fluorescein dimer was reduced to 46 conformers. Starting with this filtered conformational ensemble, AIMNET2 was used through Auto3D version 3.1 to perform a single point energy calculation for each conformer to get an updated energy for each conformer and reorder the ensemble accordingly^{57,58}. Using the AIMNET2 predicted single point energies, the 10 lowest energy conformers were selected for optimization with density functional theory (DFT). DFT optimizations for these structures were done at the M06-2X(D3)/6-31+G(d) level of theory with implicit solvation in water using the SMD solvation model^{106–116}. Next, Goodvibes software version 3.0.1 was used to find the thermochemical free energies for each conformer and provide relative energies of the conformational ensemble⁵⁹.

To assess whether the relative orientation of the two fluorophores in the crystal structure is consistent with the DFT-optimized conformational ensemble, the crystallographic dimer was compared against each of the 10 lowest-energy conformers following DFT optimization at the M06-2X(D3)/6-31+G(d) level of theory with implicit solvation (SMD, water). RMSD calculations were restricted to the 50 heavy atoms shared between both fluorescein cores of the crystal and computed structures, excluding the crystallographic amine-linker atoms absent from the computational model. For each DFT-optimized structure, atom correspondence to the crystal

structure was established by molecular graph matching, accounting for all symmetry-equivalent mappings arising from the pseudo-C₂ symmetry of the xanthene ring system, and RMSD was computed using a single rigid-body superposition (Kabsch algorithm) applied to both fluorophores simultaneously.

BIBLIOGRAPHY

1. Arthur Jones, A. & D. Snow, C. Porous protein crystals: synthesis and applications. *Chem. Commun.* **60**, 5790–5803 (2024).
2. Cumberworth, A. & Reinhardt, A. Models and simulations of structural DNA nanotechnology reveal fundamental principles of self-assembly. *Chem. Soc. Rev.* **54**, 2344–2368 (2025).
3. Kowalski, A. E. *et al.* Porous protein crystals as scaffolds for enzyme immobilization. *Biomater. Sci.* **7**, 1898–1904 (2019).
4. Stephanopoulos, N. Hybrid Nanostructures from the Self-Assembly of Proteins and DNA. *Chem* **6**, 364–405 (2020).
5. Zheng, J. *et al.* From Molecular to Macroscopic via the Rational Design of a Self-Assembled 3D DNA Crystal. *Nature* **461**, 74–77 (2009).
6. Blow, D. *Outline of Crystallography for Biologists*. (OUP Oxford, 2002).
7. Huber, T. R., McPherson, E. C., Keating, C. E. & Snow, C. D. Installing Guest Molecules at Specific Sites within Scaffold Protein Crystals. *Bioconjug. Chem.* **29**, 17–22 (2018).
8. Lu, B. *et al.* 3D Hexagonal Arrangement of DNA Tensegrity Triangles. *ACS Nano* **15**, 16788–16793 (2021).
9. Ohayon, Y. P. *et al.* Designing Higher Resolution Self-Assembled 3D DNA Crystals via Strand Terminus Modifications. *ACS Nano* **13**, 7957–7965 (2019).
10. Vecchioni, S. *et al.* The Rule of Thirds: Controlling Junction Chirality and Polarity in 3D DNA Tiles. *Small* **19**, 2206511 (2023).
11. Woloszyn, K. *et al.* Augmented DNA Nanoarchitectures: A Structural Library of 3D Self-Assembling Tensegrity Triangle Variants. *Adv. Mater.* **34**, 2206876 (2022).
12. Zhang, L., Chu, M., Ji, C., Tan, J. & Yuan, Q. Preparation, applications, and challenges of functional DNA nanomaterials. *Nano Res.* **16**, 3895–3912 (2023).
13. Ernst, P., Plückthun, A. & Mittl, P. R. E. Structural analysis of biological targets by host:guest crystal lattice engineering. *Sci. Rep.* **9**, 15199 (2019).
14. Fujita, D. *et al.* Protein encapsulation within synthetic molecular hosts. *Nat. Commun.* **3**, 1093 (2012).
15. Kong, H. *et al.* Exploring the Potential of Three-Dimensional DNA Crystals in Nanotechnology: Design, Optimization, and Applications. *Adv. Sci.* **10**, 2302021 (2023).
16. Laramy, C. R., O’Brien, M. N. & Mirkin, C. A. Crystal engineering with DNA. *Nat. Rev. Mater.* **4**, 201–224 (2019).
17. Li, X., Siraj, N., Sha, R. & Mao, C. Crystallographic Study of DNA T-Junction via Crystal Engineering. *Angew. Chem. Int. Ed.* **64**, e18174 (2025).
18. Liang, H. *et al.* Functional DNA-Containing Nanomaterials: Cellular Applications in Biosensing, Imaging, and Targeted Therapy. *Acc. Chem. Res.* **47**, 1891–1901 (2014).
19. Paukstelis, P. J. & Seeman, N. C. 3D DNA Crystals and Nanotechnology. *Crystals* **6**, 97 (2016).

20. Lee, Y.-H., Yu, E. & Park, C.-M. Programmable site-selective labeling of oligonucleotides based on carbene catalysis. *Nat. Commun.* **12**, 1681 (2021).
21. López-Tena, M., Chen, S.-K. & Winssinger, N. Supernatural: Artificial Nucleobases and Backbones to Program Hybridization-Based Assemblies and Circuits. *Bioconjug. Chem.* **34**, 111–123 (2023).
22. Rothenbühler, S., Iacovache, I., Langenegger, S. M., Zuber, B. & Häner, R. Complex DNA Architectonics—Self-Assembly of Amphiphilic Oligonucleotides into Ribbons, Vesicles, and Asterosomes. *Bioconjug. Chem.* **34**, 70–77 (2023).
23. Talha-Dean, T. *et al.* Nanoscale Control of DNA-Linked MoS₂-Quantum Dot Heterostructures. *Bioconjug. Chem.* **34**, 78–84 (2023).
24. Taylor, S. K., Kostic, N. & Stojanovic, M. N. Oligonucleotide-Blocked Streptavidin for Biotinylation Analysis. *Bioconjug. Chem.* **34**, 92–96 (2023).
25. Vecchioni, S. *et al.* Six-Letter DNA Nanotechnology: Incorporation of Z-P Base Pairs into Self-Assembling 3D Crystals. *Nano Lett.* **24**, 14302–14306 (2024).
26. Dunn, K. E. & Elfick, A. Harnessing DNA Nanotechnology and Chemistry for Applications in Photonics and Electronics. *Bioconjug. Chem.* **34**, 97–104 (2023).
27. Geng, C. & Paukstelis, P. J. DNA Crystals as Vehicles for Biocatalysis. *J. Am. Chem. Soc.* **136**, 7817–7820 (2014).
28. Hu, Q., Li, H., Wang, L., Gu, H. & Fan, C. DNA Nanotechnology-Enabled Drug Delivery Systems. *Chem. Rev.* **119**, 6459–6506 (2019).
29. Ouyang, Y., Zhang, P. & Willner, I. Dynamic Catalysis Guided by Nucleic Acid Networks and DNA Nanostructures. *Bioconjug. Chem.* **34**, 51–69 (2023).
30. Wu, Y. *et al.* Using Spectroscopy to Guide the Adaptation of Aptamers into Electrochemical Aptamer-Based Sensors. *Bioconjug. Chem.* **34**, 124–132 (2023).
31. Yang, D. & Mao, C. Nucleic Acid Nanotechnology: Innovations Shaping the Future of Biomedicine and Diagnostics. *Small Methods* **9**, 2500756 (2025).
32. Hoshika, S., Shukla, M. S., Benner, S. A. & Georgiadis, M. M. Visualizing “Alternative Isoinformational Engineered” DNA in A- and B-Forms at High Resolution. *J. Am. Chem. Soc.* **144**, 15603–15611 (2022).
33. Zhang, Z. *et al.* Assembly of Protein-DNA Framework Nanostructures: Structurally Defining Protein-DNA Interfaces With Aptamer. *Angew. Chem. Int. Ed.* **n/a**, e9624043.
34. Shields, E. T. *et al.* Modular Scaffold Crystals for Programmable Installation and Structural Observation of DNA-Binding Proteins. 2026.03.04.709581 Preprint at <https://doi.org/10.64898/2026.03.04.709581> (2026).
35. Orun, A. R. *et al.* Modular Protein–DNA Cocrystals as Precise, Programmable Assembly Scaffolds. *ACS Nano* **17**, 13110–13120 (2023).
36. Orun, A. R. *et al.* Tuning Chemical DNA Ligation within DNA Crystals and Protein–DNA Cocrystals. *ACS Nanosci. Au* <https://doi.org/10.1021/acsnanoscienceau.4c00013> (2024) doi:10.1021/acsnanoscienceau.4c00013.

37. Ward, A. R., Dmytriw, S., Vajapayajula, A. & Snow, C. D. Stabilizing DNA–Protein Co-Crystals via Intra-Crystal Chemical Ligation of the DNA. *Crystals* **12**, 49 (2022).
38. Eroshenko, D. A., Diatlova, E. A., Golyshev, V. M., Endutkin, A. V. & Zharkov, D. O. Aberrant Repair of 8-Oxoguanine in Short DNA Bulges. *Dokl. Biochem. Biophys.* **513**, S82–S86 (2023).
39. Freier, S. M. & Altmann, K.-H. The ups and downs of nucleic acid duplex stability: Structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.* **25**, 4429–4443 (1997).
40. Gohil, D., Sarker, A. H. & Roy, R. Base Excision Repair: Mechanisms and Impact in Biology, Disease, and Medicine. *Int. J. Mol. Sci.* **24**, 14186 (2023).
41. Krokan, H. E. & Bjørås, M. Base Excision Repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012583 (2013).
42. Moon, J. *et al.* DNA Damage and Its Role in Cancer Therapeutics. *Int. J. Mol. Sci.* **24**, 4741 (2023).
43. Natrajan, G. *et al.* Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.* **31**, 4814–4821 (2003).
44. Yang, W. Structure and mechanism for DNA lesion recognition. *Cell Res.* **18**, 184–197 (2008).
45. Wang, F. *et al.* A Bridging Water Anchors the Tethered 5-(3-Aminopropyl)-2'-deoxyuridine Amine in the DNA Major Groove Proximate to the N+2 C·G Base Pair: Implications for Formation of Interstrand 5'-GNC-3' Cross-Links by Nitrogen Mustards,. *Biochemistry* **47**, 7147–7157 (2008).
46. Satange, R., Chuang, C.-Y., Neidle, S. & Hou, M.-H. Polymorphic G:G mismatches act as hotspots for inducing right-handed Z DNA by DNA intercalation. *Nucleic Acids Res.* **47**, 8899–8912 (2019).
47. Batra, V. K., Beard, W. A., Pedersen, L. C. & Wilson, S. H. Structures of DNA Polymerase Mispaiored DNA Termini Transitioning to Pre-catalytic Complexes Support an Induced-Fit Fidelity Mechanism. *Structure* **24**, 1863–1875 (2016).
48. Hwang, G.-S., Jones, G. B. & Goldberg, I. H. Solution Structure of a Wedge-Shaped Synthetic Molecule at a Two-Base Bulge Site in DNA. *Biochemistry* **42**, 8472–8483 (2003).
49. Hwang, G.-S., Jones, G. B. & Goldberg, I. H. Stereochemical Control of Small Molecule Binding to Bulged DNA: Comparison of Structures of Spirocyclic Enantiomer-Bulged DNA Complexes,. *Biochemistry* **43**, 641–650 (2004).
50. Stassinopoulos, A., Ji, J., Gao, X. & Goldberg, I. H. Solution Structure of a Two-Base DNA Bulge Complexed with an Enediyne Cleaving Analog. *Science* **272**, 1943–1946 (1996).
51. Zhang, N., Lin, Y., Xiao, Z., Jones, G. B. & Goldberg, I. H. Solution Structure of a Designed Spirocyclic Helical Ligand Binding at a Two-Base Bulge Site in DNA. *Biochemistry* **46**, 4793–4803 (2007).

52. Hsiao, Y.-Y., Fang, W.-H., Lee, C.-C., Chen, Y.-P. & Yuan, H. S. Structural Insights Into DNA Repair by RNase T—An Exonuclease Processing 3' End of Structured DNA in Repair Pathways. *PLOS Biol.* **12**, e1001803 (2014).
53. Li, S., Olson, W. K. & Lu, X.-J. Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res.* **47**, W26–W34 (2019).
54. Arhangelskis, M. *et al.* Rationalization of the Color Properties of Fluorescein in the Solid State: A Combined Computational and Experimental Study. *Chem. – Eur. J.* **22**, 10065–10073 (2016).
55. Osborn, R. S. & Rogers, D. The crystal and molecular structure of the 1:1 complex of acetone with the lactoid form of fluorescein. *Acta Crystallogr. B* **31**, 359–364 (1975).
56. Răducă, M. & Mădălan, A. M. Cocrystals versus Salts of Fluorescein. *Crystals* **11**, (2021).
57. Anstine, D., Zubatyuk, R. & Isayev, O. AIMNet2: A Neural Network Potential to Meet your Neutral, Charged, Organic, and Elemental-Organic Needs. *ChemRxiv* **2024**,.
58. Liu, Z., Zubatyuk, T., Roitberg, A. & Isayev, O. Auto3D: Automatic Generation of the Low-Energy 3D Structures with ANI Neural Network Potentials. *J. Chem. Inf. Model.* **62**, 5373–5382 (2022).
59. Luchini, G., Alegre-Requena, J. V., Funes-Ardoiz, I. & Paton, R. S. GoodVibes: automated thermochemistry for heterogeneous computational chemistry data. Preprint at <https://doi.org/10.12688/f1000research.22758.1> (2020).
60. Cardenal, A. D. & Ramadhar, T. R. Application of Crystalline Matrices for the Structural Determination of Organic Molecules. *ACS Cent. Sci.* **7**, 406–414 (2021).
61. Habib, F., Tocher, D. A. & Carmalt, C. J. Applications of the crystalline sponge method and developments of alternative crystalline sponges. *Mater. Today Proc.* **56**, 3766–3773 (2022).
62. Inokuma, Y., Yoshioka, S., Ariyoshi, J., Arai, T. & Fujita, M. Preparation and guest-uptake protocol for a porous complex useful for ‘crystal-free’ crystallography. *Nat. Protoc.* **9**, 246–252 (2014).
63. Crooke, S. T., Vickers, T. A. & Liang, X. Phosphorothioate modified oligonucleotide–protein interactions. *Nucleic Acids Res.* **48**, 5235–5253 (2020).
64. Egli, M. & Manoharan, M. Chemistry, structure and function of approved oligonucleotide therapeutics. *Nucleic Acids Res.* **51**, 2529–2573 (2023).
65. Jahns, H. *et al.* Chirality matters: stereo-defined phosphorothioate linkages at the termini of small interfering RNAs improve pharmacology *in vivo*. *Nucleic Acids Res.* **50**, 1221–1240 (2022).
66. Nassir, M. *et al.* An Improved P(V) Thio-Oligonucleotide Synthesis Platform. *Org. Lett.* **27**, 97–102 (2025).
67. Conn, G. L., Brown, T. & Leonard, G. A. The crystal structure of the RNA/DNA hybrid r(GAAGAGAAGC)·d(GCTTCTCTTC) shows significant differences to that found in solution. *Nucleic Acids Res.* **27**, 555–561 (1999).

68. Ho, P. S. & Carter, M. DNA Structure: Alphabet Soup for the Cellular Soul. in *DNA Replication - Current Advances* (IntechOpen, 2011). doi:10.5772/18536.
69. Liu, J.-H. *et al.* Structural Flexibility of DNA-RNA Hybrid Duplex: Stretching and Twist-Stretch Coupling. *Biophys. J.* **117**, 74–86 (2019).
70. McDonald, A. R., Denning, E. J. & MacKerell, A. D. Jr. Impact of Geometry Optimization on Base–Base Stacking Interaction Energies in the Canonical A- and B-Forms of DNA. *J. Phys. Chem. A* **117**, 1560–1568 (2013).
71. Endo, M. *et al.* RNA-templated DNA origami structures. *Chem. Commun.* **49**, 2879–2881 (2013).
72. Ko, S. H. *et al.* Synergistic self-assembly of RNA and DNA molecules. *Nat. Chem.* **2**, 1050–1055 (2010).
73. Xiong, Y. & Sundaralingam, M. Crystal structure of a DNA·RNA hybrid duplex with a polypurine RNA r(gaagaagag) and a complementary polypyrimidine DNA d(CTCTTCTTC). *Nucleic Acids Res.* **28**, 2171–2176 (2000).
74. Horton, N. C. & Finzel, B. C. The Structure of an RNA/DNA Hybrid: A Substrate of the Ribonuclease Activity of HIV-1 Reverse Transcriptase. *J. Mol. Biol.* **264**, 521–533 (1996).
75. Jhan, C.-R. *et al.* Targeting the ALS/FTD-associated A-DNA kink with anthracene-based metal complex causes DNA backbone straightening and groove contraction. *Nucleic Acids Res.* **49**, 9526–9538 (2021).
76. Wang, J. *et al.* Unlocking the potential of diarylmethane dyes for living systems. *Chem* **12**, (2026).
77. Aguilera, A. & García-Muse, T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol. Cell* **46**, 115–124 (2012).
78. Nowotny, M. *et al.* Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. *Mol. Cell* **28**, 264–276 (2007).
79. Sarafianos, S. G. *et al.* Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**, 1449–1461 (2001).
80. Hao, Y. *et al.* A device that operates within a self-assembled 3D DNA crystal. *Nat. Chem.* **9**, 824–827 (2017).
81. Vecchioni, S. *et al.* Metal-Mediated DNA Nanotechnology in 3D: Structural Library by Templated Diffraction. *Adv. Mater.* **35**, 2210938 (2023).
82. Zhang, C. *et al.* Engineering DNA Crystals toward Studying DNA–Guest Molecule Interactions. *J. Am. Chem. Soc.* **145**, 4853–4859 (2023).
83. Changela, A., DiGate, R. J. & Mondragón, A. Structural Studies of E. coli Topoisomerase III-DNA Complexes Reveal A Novel Type IA Topoisomerase-DNA Conformational Intermediate. *J. Mol. Biol.* **368**, 105–118 (2007).
84. Kouno, T. *et al.* Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nat. Commun.* **8**, 15024 (2017).

85. Tompkins, K. J. *et al.* Molecular underpinnings of ssDNA specificity by Rep HUH-endonucleases and implications for HUH-tag multiplexing and engineering. *Nucleic Acids Res.* **49**, 1046–1064 (2021).
86. Wegrzyn, K. *et al.* Rep protein accommodates together dsDNA and ssDNA which enables a loop-back mechanism to plasmid DNA replication initiation. *Nucleic Acids Res.* **51**, 10551–10567 (2023).
87. Li, X. *et al.* Crystalline Assemblies of DNA Nanostructures and Their Functional Properties. *Angew. Chem. Int. Ed.* **64**, e202416948 (2025).
88. Komori, H., Sasai, N., Matsunaga, F., Wada, C. & Miki, K. Crystallization and Preliminary X-Ray Diffraction Studies of a Replication Initiator Protein (RepE54) of the Mini-F Plasmid Complexed with Iteron DNA. *J. Biochem. (Tokyo)* **125**, 24–26 (1999).
89. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
90. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
91. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).
92. Bannwarth, C. *et al.* Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **11**, e1493 (2021).
93. de Souza, B. GOAT: A Global Optimization Algorithm for Molecules and Atomic Clusters. *Angew. Chem. Int. Ed.* **64**, e202500393 (2025).
94. Ehlert, S., Stahn, M., Spicher, S. & Grimme, S. Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods. *J. Chem. Theory Comput.* **17**, 4250–4261 (2021).
95. Lehtola, S., Steigemann, C., Oliveira, M. J. T. & Marques, M. A. L. Recent developments in libxc — A comprehensive library of functionals for density functional theory. *SoftwareX* **7**, 1–5 (2018).
96. Lehtola, S. & Marques, M. A. L. Reproducibility of density functional approximations: How new functionals should be reported. *J. Chem. Phys.* **159**, 114116 (2023).
97. Marques, M. A. L., Oliveira, M. J. T. & Burnus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* **183**, 2272–2281 (2012).
98. Neese, F. Approximate second-order SCF convergence for spin unrestricted wavefunctions. *Chem. Phys. Lett.* **325**, 93–98 (2000).
99. Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2**, 73–78 (2012).
100. Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Comput. Mol. Sci.* **12**, e1606 (2022).
101. Neese, F., Wennmohs, F., Becker, U. & Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **152**, 224108 (2020).
102. Valeev, E. evaleev/libint. (2026).

103. Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *J. Chem. Theory Comput.* **15**, 2847–2862 (2019).
104. Pracht, P., Bohle, F. & Grimme, S. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.* **22**, 7169–7192 (2020).
105. Pracht, P. *et al.* CREST-A program for the exploration of low-energy molecular chemical space. *J. Chem. Phys.* **160**, 114110 (2024).
106. Binning Jr., R. C. & Curtiss, L. A. Compact contracted basis sets for third-row atoms: Ga–Kr. *J. Comput. Chem.* **11**, 1206–1216 (1990).
107. Blaudeau, J.-P., McGrath, M. P., Curtiss, L. A. & Radom, L. Extension of Gaussian-2 (G2) theory to molecules containing third-row atoms K and Ca. *J. Chem. Phys.* **107**, 5016–5021 (1997).
108. Ditchfield, R., Hehre, W. J. & Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **54**, 724–728 (1971).
109. Francl, M. M. *et al.* Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **77**, 3654–3665 (1982).
110. Gordon, M. S. The isomers of silacyclopropane. *Chem. Phys. Lett.* **76**, 163–168 (1980).
111. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104 (2010).
112. Hariharan, P. C. & Pople, J. A. Accuracy of AH *n* equilibrium geometries by single determinant molecular orbital theory. *Mol. Phys.* **27**, 209–214 (1974).
113. Hehre, W. J., Ditchfield, R. & Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **56**, 2257–2261 (1972).
114. Rassolov, V. A., Pople, J. A., Ratner, M. A. & Windus, T. L. 6-31G* basis set for atoms K through Zn. *J. Chem. Phys.* **109**, 1223–1229 (1998).
115. Rassolov, V. A., Ratner, M. A., Pople, J. A., Redfern, P. C. & Curtiss, L. A. 6-31G* basis set for third-row atoms. *J. Comput. Chem.* **22**, 976–984 (2001).
116. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241 (2008).

CHAPTER 3: ENZYMATIC AND CHEMICAL LIGATION FOR STABILIZATION OF THE CC1 PLATFORM

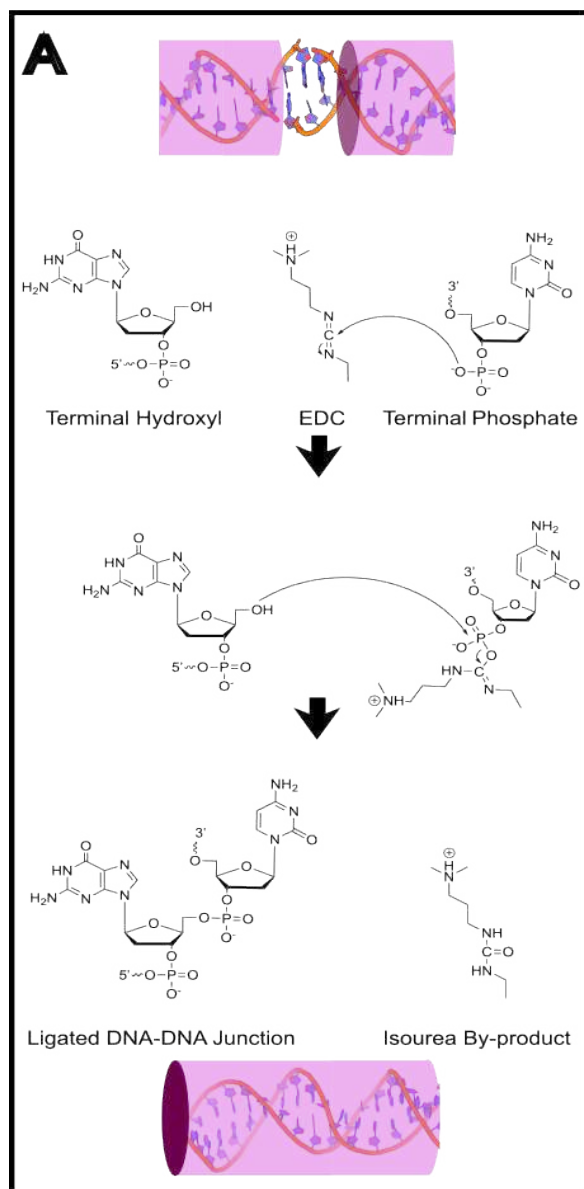
3.1 Introduction

3.1.1 Motivation for post-assembly ligation in protein–DNA co-crystals

Crystals composed of protein and DNA are versatile biomaterials that enable precise molecular ordering through programmed self-assembly and tunable growth properties. In CC1, protein–DNA self-assembly produces porous crystalline lattices with solvent-accessible DNA components, making these materials attractive scaffolds for host-guest structural biology and, more broadly, for biomolecular capture, storage, and remodeling (1-3). However, the same noncovalent interactions that enable ordered self-assembly also impose a central limitation, as-grown biomolecular crystals are often fragile outside of their native crystallization conditions, or mother liquor (4, 5). This limitation is especially pronounced for DNA-containing crystals, where lattice stability depends strongly on ionic strength, divalent cation concentration, and the electrostatic environment surrounding the nucleic acid scaffold (6). Consequently, DNA-based crystals often cannot tolerate the solution exchanges required for downstream applications, including guest soaking, enzymatic treatment, or biologically relevant buffer challenges (7).

Our previous work established that CC1+10 crystals can function as scaffolds for structure determination of DNA-binding guest molecules (3). A key enabling feature of that work was the ability to transfer crystals from mother liquor into alternative buffer conditions that both promoted guest binding after soaking in guest solution and preserved lattice order for high-resolution XRD. This stability was achieved through post-crystallization treatment with EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (Fig. 3.1 A). EDC can promote two chemically distinct forms of covalent stabilization 1. Phosphate-hydroxyl coupling across DNA nicks and 2. Carboxyl-amine crosslinking between nearby protein functional groups (8-11). Together, these

reactions provide a route to reinforce the lattice after crystal growth. In prior CC1+10 guest-soaking experiments, EDC ligation provided stability sufficient to identify buffer conditions compatible with the binding of multiple classes of DNA-binding proteins while retaining diffraction-quality order (3). However, extending CC1 to a broader library of guest proteins via different binding solutions will require crystals that tolerate a wider range of solution conditions. Improving post-assembly ligation is therefore a critical step toward expanding CC1 as a robust scaffold for structural biology and biomolecular materials applications.



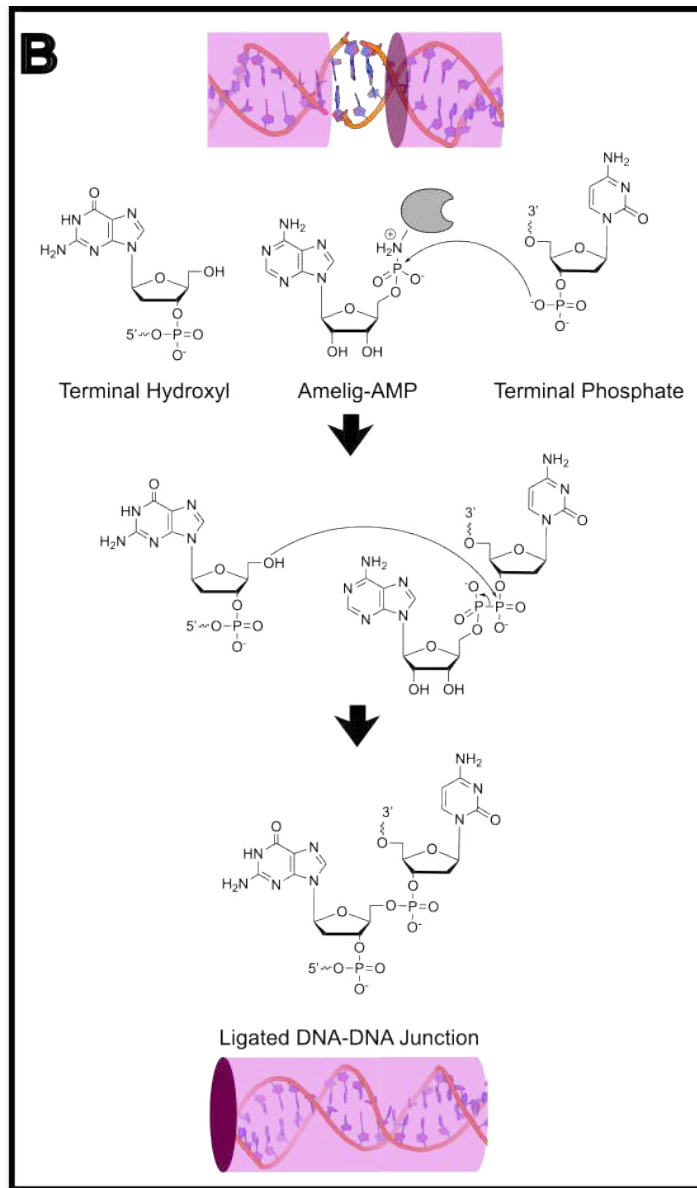


Figure 3.1: Ligation Chemistries. **A.** EDC ligation across a terminal hydroxyl and terminal phosphate. This results in a new phosphodiester bond between nucleotides and an isourea by-product. **B.** AmeLig ligation using ATP, forming an enzyme-AMP intermediate, resulting in a new phosphodiester bond between nucleotides.

To date, our most well-characterized stabilization strategy has been post-assembly ligation across DNA-DNA junctions within the crystal lattice. Covalent linkage at these junctions effectively “weaves” adjacent DNA building blocks together, converting a lattice held primarily by reversible interactions into one reinforced by covalent bonds. Chemical ligation

with EDC is particularly useful because it can stabilize crystals through both intended DNA-DNA junction ligation and additional protein crosslinking. In previous studies of EDC-treated interpenetrating CC1 (where a second copy of the lattice grows in the solvent channel pores, or ipCC1), optimized conditions produced DNA-DNA ligation in 58% of building blocks (12). Importantly, EDC ligation requires a terminal phosphate (13). However, ipCC1 variants lacking terminal phosphates were still stabilized by EDC, relative to untreated controls, suggesting that EDC contributes additional stabilization beyond the designed DNA-DNA junctions. Although structural evidence of discrete protein-protein crosslinks was not resolved crystallographically, the stabilization of phosphate-deficient crystals indicates that carboxyl-amine crosslinking likely contributes to the mechanical reinforcement of the lattice (12).

In pursuit of precise DNA-DNA ligation, and to potentially increase ligation yields, here we investigate the use of a ligase enzyme. Enzymatic ligation offers a complementary strategy because DNA ligases selectively form phosphodiester bonds at defined 5'-phosphate/3'-hydroxyl nicks (Fig. 3.1 B) (14, 15). However, enzymatic ligation inside CC1 requires that the ligase diffuse through the solvent channels, bind the DNA nick, and complete catalysis; each of these steps can be stymied by steric hinderance of the surrounding lattice. These constraints are strongly dependent on the lattice architecture. CC1+10 contains solvent channels predicted to accommodate a biomolecule with a maximum diameter of ~5.3 nm, as calculated by MAP_CHANNELS, that likely excludes or sterically restricts larger ligases such as T4 DNA ligase, that is known to contact an 18-nucleotide footprint (3, 16, 17). Although T4 ligase has been used previously to ligate DNA-only tensegrity triangle crystals, its size and binding geometry are poorly matched to the narrower pores of CC1+10 (Fig S3.1) (18). For this reason, we investigated AmeLig, a smaller DNA ligase from *Alteromonas mediterranea*, as a candidate

enzyme for in-crystal ligation (19). Structural modeling suggested that AmeLig could access multiple nick sites in CC1+10 and over more, would be readily accommodated in the larger-pore CC1+21 architecture (Fig S3.1).

Finally, we extend the concept of enzyme accessibility beyond ligation by demonstrating restriction enzyme activity in CC1+21 crystals using endonuclease HhaI (20). This experiment addresses a long-standing goal of dynamic lattice remodeling, in which DNA building blocks could be selectively cleaved and then removed/replaced through mass-action exchange. In this context, ligating the crystal and then enzymatically cutting the DNA scaffold represents a form of programmed “controlled demolition.” HhaI treatment of AmeLig-ligated CC1+21 crystals provides evidence that enzyme-mediated dissolution is possible. By contrast, EDC-ligated crystals remain macroscopically intact after HhaI treatment, likely because nonspecific protein crosslinks preserve the crystal structure even if DNA cleavage occurs. Together, these results show that chemical and enzymatic ligation serve distinct but complementary roles. EDC provides robust stabilization of CC1+10 for solution exchange and diffraction, whereas CC1+21 is capable of undergoing more precise enzymatic ligation and cleavage.

3.1.2 Chapter objectives

This chapter evaluates chemical and enzymatic ligation strategies in CC1+10 and CC1+21 to improve covalent stabilization of the lattice. In CC1+10, we screen EDC ligation yields across junction architectures - varying sticky base overhang length and terminal phosphate position - then testing whether removing magnesium from the crosslink buffer affects yield. Free-solution splint ligation controls directly compare EDC and AmeLig, revealing that AmeLig is far more efficient, establishing that EDC's in-crystal yields are bolstered by the ligation being templated by the crystal lattice. Despite this advantage, AmeLig fails to ligate CC1+10 building

blocks *in crystallo*, implicating steric exclusion from junction sites within the pores. This motivates moving to CC1+21, where expanded pores permit AmeLig to diffuse more freely. We demonstrate successful enzymatic ligation in CC1+21 and use the restriction enzyme HhaI to then selectively reverse that ligation. AmeLig-ligated crystals dissolve upon HhaI treatment, establishing a proof of concept for controlled crystal disassembly.

3.2 Ligation in CC1+10 Expanded Crystals

3.2.1 Experimental Design for EDC Ligation Assays

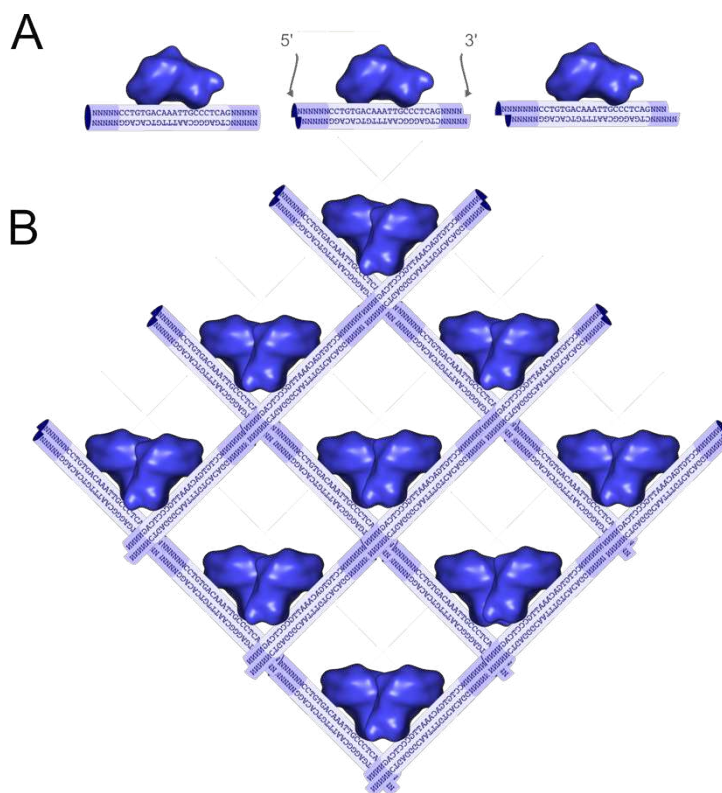


Figure 3.2: Constructs of CC1 with varied junction architecture. **A.** Building blocks with 0, 1, and 2 sticky base overhangs and indications of the 5' and 3' ends of the DNA where a phosphate can be added. **B.** CC1 lattice showing one of the protein-protein interfaces and the DNA:DNA junctions.

Crystals were assembled from a CC1+10 building block in which RepE54 binds a two-turn DNA recognition sequence that is extended by 10 bp to generate solvent channel pores, resulting in linear DNA stacks aligned end-to-end across crystal contacts. Junctions were

engineered with defined structural features, including phosphates on either the 5' or 3' termini and 0-, 1-, or 2-base single-stranded overhangs, yielding predictable ligation products upon successful ligation (*e.g.*, dimers, trimers, and higher-order multimers) (Fig. 3.2). Intact crystals were soaked in solutions containing EDC. Importantly, here we balance high concentrations and multiple doses of EDC with conditions that are less harsh on crystals. Multiple doses of EDC requires changing solution multiple times, that may result in higher ligation yields, but may lower XRD resolution through crystal cracking or building block diffusion out of the crystal. Following EDC incubation, crystals were dissolved and reaction products were resolved by gel electrophoresis. Ligation yields were quantified by gel densitometry, measuring the molar fractions of 1-mer, 2-mer, 3-mer, and higher-order species. Unligated 1-mer building blocks were used as internal controls to distinguish true ligation, while DNA gel ladders enabled accurate sizing and quantification of higher-order ligation products, allowing direct comparison of chemical ligation outcomes across junction architectures. Band intensities were quantified by densitometry after local background subtraction. Product yields were calculated as the integrated intensity of ligated products divided by the total lane intensity.

For accuracy of reporting, we refer to these ligation calculations as relative ligation yields. Gel densitometry is a semi-quantitative estimate of DNA ligation yield rather than an absolute measurement of product concentration. Ligation efficiency was estimated from the fraction of total lane intensity assigned to ligated product, with substrate and relevant intermediate bands included in the same lane. This within-lane normalization is useful for comparing reaction conditions because it reduces the effect of modest loading differences and directly reports the relative conversion of starting material into higher-molecular-weight product.

External DNA standards can improve absolute quantification, especially for confirming that band intensities fall within the linear detection range of the staining and imaging system. However, standards are not always necessary when the goal is relative comparison among ligation conditions, because substrate, product, and intermediate species are measured within the same lane. In this context, a standard can help validate detection range but does not necessarily improve the internal comparison between bands in a given reaction.

A key limitation is that DNA band intensity reflects both DNA amount and fragment length, since longer DNA species contain more nucleotides and can bind more fluorescent dye per molecule. Therefore, raw band intensity does not directly correspond to molar yield when ligation generates products of different lengths, and gel-based quantification is most reliable when standards are similar in size to the sample DNA. For this reason, our densitometric ligation yield is interpreted as an apparent or relative yield.

In the present analysis, the goal was to compare ligation outcomes across reaction conditions rather than determine absolute product concentration. We therefore used internal lane normalization and, where appropriate, length-weighted intensity calculations to estimate the relative fraction of DNA converted into ligated product. This approach supports comparison between conditions while avoiding overinterpretation of gel densitometry as a fully quantitative measurement.

3.2.2 Effect of Junction Architecture on EDC Ligation Yield

Our laboratory has previously demonstrated that DNA ligation efficiency is strongly influenced by phosphate placement (3' versus 5'), sticky base length (0, 1, or 2 bases), as well as EDC concentration and dosing frequency (9, 12, 13). However, these studies were conducted using the ipCC1 platform. In this chapter, we evaluate how porous CC1 junction changes alter

ligation yield using the most effective EDC dosing strategy identified for ipCC1, two doses of 30 mg/mL EDC. Interestingly, ligated porous CC1 crystals exhibit reduced macrostructural stability in deionized water as compared to ligated ipCC1, suggesting that lattice topology differences alter the stabilizing effects of chemical crosslinking. One possible contributor to this behavior is off-target carboxyl-amine crosslinking by EDC, that may provide additional mechanical reinforcement in ipCC1 but is less effective in the porous lattice, perhaps due to less protein-protein contacts per unit cell.

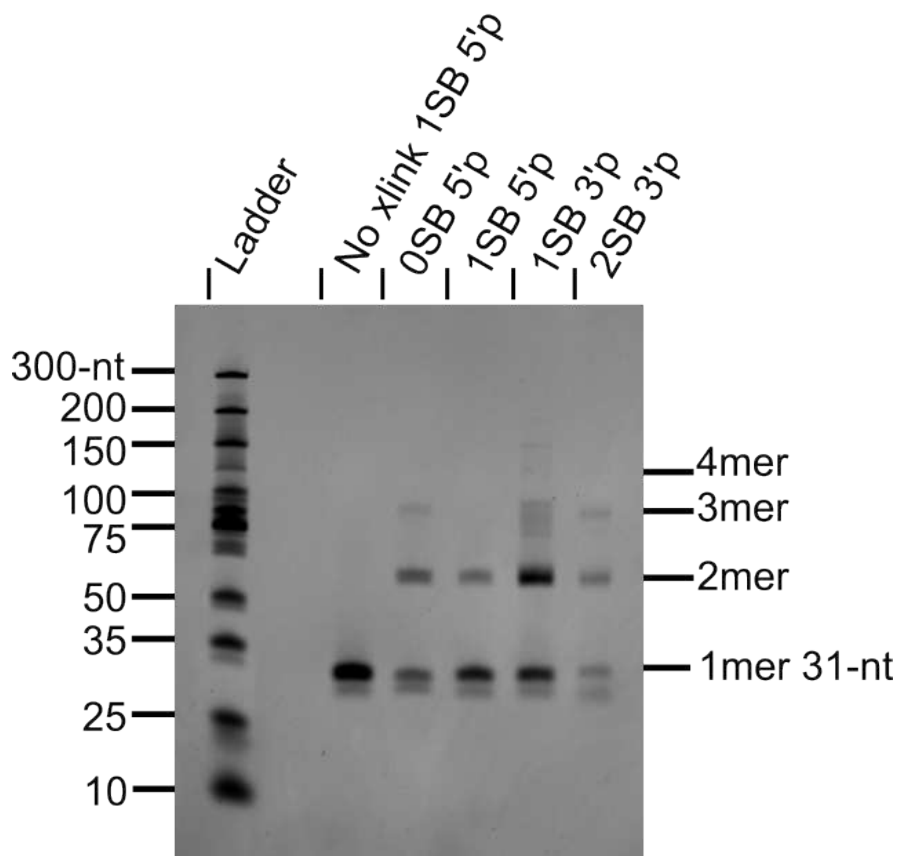


Figure 3.3: TBE-Urea denaturing gel electrophoresis of chemically ligated CC1+10. Lane 1 shows DNA ladder some bands identified by length in nucleotides. Lane 3 shows a dissolved CC1+10 that was not subjected to EDC treatment. Lanes 4-7 show CC1+10 bp grown with different junction architectures (0, 1, 2 sb and 3' or 5' phosphate) that have been treated with 2 doses of 30 mg/mL EDC then dissolved via proteinase and 3.5 M Urea in 1X TBE. 1-mer building blocks are 31 nt, 2-mer is 62 nt, etc. Sequences provided in Table S3.1.

Informed by the trends established in the ipCC1 system, we initially anticipated that junctions containing a 3'-phosphate with a 2-base sticky overhang would exhibit the highest relative ligation yield (11, 12). Instead, gel-based analysis reveals that 2-base sticky junctions 3' phosphate (2 sb 3'p) and 0 sb 5'p junctions display comparable levels of apparent ligation. 0 sb 5'p crystals exhibited 27% of building blocks as 2-mer or higher species and 2 sb 3'p crystals showed 29% of building blocks as 2-mer or higher species. The most efficiently ligated crystal type was the 1 sb 3'p, producing the greatest abundance of higher-order ligation products at 37% (Table 3.1). In contrast, 1 sb 5'p junctions exhibited the lowest relative ligation yield, with only 15% 2-mer species and no discernable signal for higher-order products.

Table 3.1: Distribution of DNA block sizes as a function of DNA-DNA junction architecture. The data shown correspond with the gel lanes in Fig. 3.3. The crosslinking protocols were 2 doses of 30 mg/mL EDC for 12 h each, as compared to dissolved crystals that were not subjected to EDC doses. The values in this table are weighted so that the final value was a molar fraction. Densitometry fits are in Fig. S3.2-S3.5.

CC1+10 Junction Comparison					
DNA Block Size	No Crosslink [%]	0 SB 5'P [%]	1 SB 5'P [%]	1 SB 3'P [%]	2 SB 3'P [%]
1	100	73	85	64	71
2	0	24	15	28	23
3	0	3	0	8	6

Relevant to our goals of high throughput CC1 crystal growth for use as a broadly applicable scaffold for structure determination, crystals containing 5' P tend to form crystals quickly. To increase ligation, and potentially lattice stability, future work may focus on optimization of more rapid growth of 3' P CC1 crystals. Although higher EDC concentrations or additional dosing rounds might further increase apparent DNA-DNA junction ligation efficiency, we have avoided this strategy. The goal of ligation is to enable crystal survival in varied solution conditions and high-resolution XRD data collection. More than one EDC treatment risks

potential loss of building blocks from solution changes and building block equilibration, which would reduce diffraction resolution. The two doses of 30 mg/mL shows how “optimal” ligation conditions from the ipCC1 studies translates to porous CC1, but for practical ligation prior to guest soaking and XRD we hope to use a single dose of 45 mg/mL EDC. Doses higher than 45 mg/mL caused visible degradation of the crystal, while lower single doses did not impart sufficient crystal structural stability in downstream guest soaking.

3.2.3 Effect of Magnesium on EDC Ligation in CC1+10

Initial investigation of EDC ligation in CC1+0 (the original structure pdb code 7RVA) found that MgCl_2 reduced ligation yields (12). An exact mechanism of Mg(II) interference in the reaction was not determined, but it was hypothesized that perhaps high magnesium concentrations might reduce the availability of nucleophilic phosphate groups during the reaction. As discussed above, EDC reacts with carboxylate groups. These two concepts make the Mg(OAc)_2 present in the porous CC1 growth buffer doubly problematic. To assess whether we could improve relative ligation yields with removal of Mg(OAc)_2 , two replacement counterions were tested: sodium chloride (motivated by previous studies in ipCC1) and manganese chloride (motivated by previous studies of EDC DNA ligation in solution) (9, 12, 13). To adapt solution-optimized EDC ligation conditions to the CC1+10 lattice, we introduced 100 mM MES pH 6.0, a buffer condition previously reported for efficient EDC DNA ligation in solution, while retaining Li_2SO_4 and Mg(OAc)_2 (9). All buffers contained 1 M Li_2SO_4 , as crystals dissolved rapidly in lower solutions, and higher salt concentrations may interfere in the reaction. Crystals containing 5' P and 0 sticky base overhangs were subjected to one dose of 45 mg/mL EDC in the different buffers, then dissolved and analyzed using denaturing gel electrophoresis (Fig. 3.4).

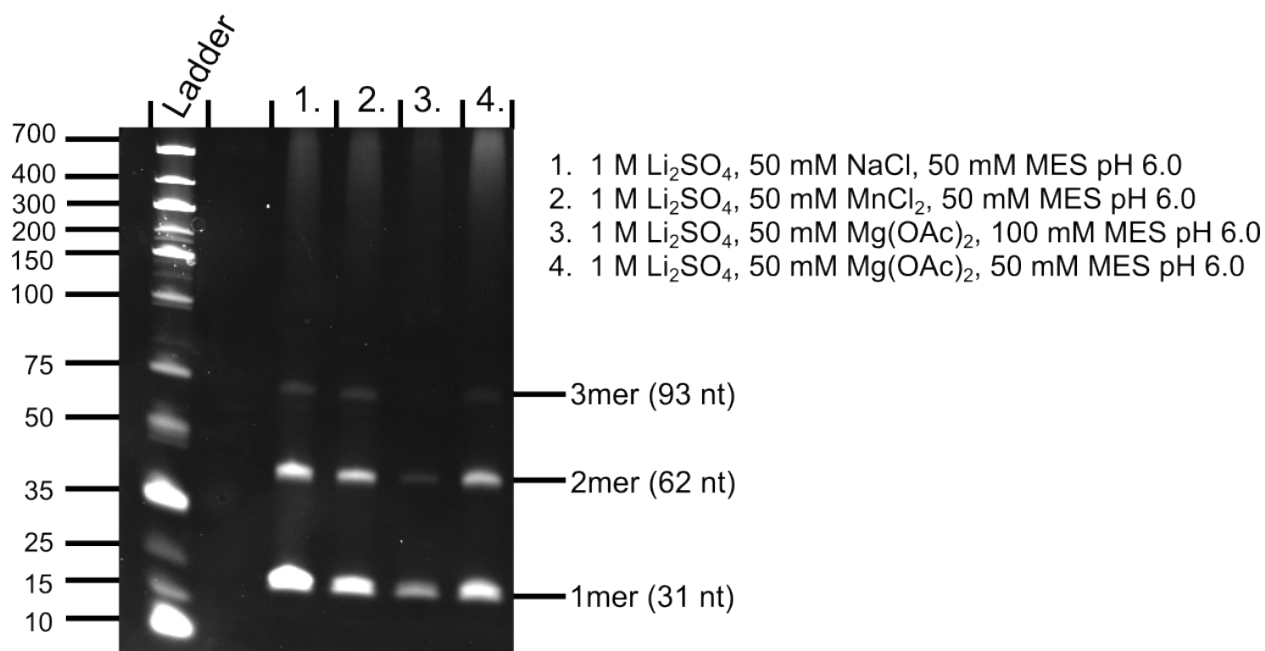


Figure 3.4: Buffer condition comparison. Buffer 1 substitutes NaCl for Mg(OAc)₂. Buffer 2 substitutes MnCl₂ for Mg(OAc)₂. Buffer 3 substitutes 100 mM MES pH 6.0 for 50 mM MES pH 6.0. Buffer 4 is the crosslink buffer used in our previous structural observation paper (3).

Changing the salt or divalent counterion composition did not produce an apparent improvement in EDC-mediated ligation efficiency in CC1+10 crystals (Table S3.2). In the NaCl substitution buffer, ligation yield was reduced relative to the standard porous CC1 crosslinking buffer, as indicated by a higher fraction of monomeric DNA building block remaining after treatment: 88% monomer in NaCl compared with 80% monomer in the standard Mg(OAc)₂ condition. A more pronounced reduction in ligation was observed when the original Li₂SO₄ and Mg(OAc)₂ salt identities were retained but the MES concentration was increased to 100 mM, with 94% of DNA remaining as monomer and only 6% detected as dimer. This result suggests that simply adopting the higher MES concentration used in prior solution-based EDC ligation conditions does not improve ligation in the CC1+10 crystal environment. Instead, increasing buffer concentration may reduce ligation efficiency, potentially by altering the local protonation environment, ionic activity, or EDC lifetime within the crystal-containing solution.

Table 3.2: Distribution of DNA block sizes in differing buffer conditions. Crystals were treated with 45 mg/mL EDC dissolved in the buffer then dissolved and analyzed via denaturing gel electrophoresis. Buffer 1 is 1 M Li₂SO₄, 50 mM NaCl, 50 mM MES pH 6.0. Buffer 2 is 1 M Li₂SO₄, 50 mM MnCl₂, 50 mM MES pH 6.0. Buffer 3 is 1 M Li₂SO₄, 50 mM Mg(OAc)₂, 100 mM MES pH 6.0. Buffer 4 is 1 M Li₂SO₄, 50 mM Mg(OAc)₂, 50 mM MES pH 6.0. Densitometry fits can be found in Fig. S3.6-S3.9

CC1+10 EDC Buffer Comparison				
DNA Block Size	Buffer 1 (NaCl) [%]	Buffer 2 (MnCl ₂) [%]	Buffer 3 (100 mM MES) [%]	Buffer 4 (Mg(OAc) ₂) [%]
1	88	80	94	80
2	11	18	6	19
3	1	2	0	1

Interestingly, substitution with MnCl₂ produced a ligation profile nearly identical to that of the standard Mg(OAc)₂ condition. The MnCl₂ and Mg(OAc)₂ buffers yielded identical monomer fractions, 80% in both cases, and similar levels of dimer and trimer formation. This indicates that replacing Mg(OAc)₂ with MnCl₂ did not measurably enhance EDC-mediated DNA-DNA coupling under these conditions. Because acetate is a carboxylate-containing counterion and could, in principle, compete with phosphate activation by EDC, we initially expected that removing acetate might improve ligation. However, the similarity between the MnCl₂ and Mg(OAc)₂ conditions suggests that acetate is not a dominant inhibitor in the current CC1+10 crosslinking buffer, likely because the Mg(OAc)₂ concentration is relatively low compared with DNA tile crystallization systems that use hundreds of millimolar Mg(OAc)₂. Together, these buffer comparisons indicate that the standard porous CC1 crosslinking buffer does not contain an obvious, easily substitutable component that strongly limits DNA-DNA ligation yield. This further motivates exploring enzymatic ligation to increase ligation yields.

3.2.4 Splint Ligation Controls: AmeLig vs EDC Yields

We expressed the small ligase AmeLig using the plasmid from Williamson *et al.*, 2018 (Fig. S3.10) (19). To assess the feasibility for using AmeLig in the same crosslinking buffer in which we do EDC crosslinking, we used a splint ligation assay made of three oligos (Fig. 3.5). The oligos have differing lengths (A is 26-nt, B is 20-nt, C is 24-nt), and when complexed a 5' P on oligo B is adjacent to the 3' hydroxyl of A, held together by C as a splint. If ligation is successful, the resulting species is A+B at 46-nt, leaving C at 24-nt. Separate oligo and all combinations of oligos are shown via denaturing gel electrophoresis (Fig. S3.11). As seen in Fig. 3.5, the original 3 strand stock has contaminating bands at approximately the double length of the A, B, and C oligos, a common impurity in solid-phase phosphoramidite DNA synthesis (20). This may contribute to the extra higher molecular weight bands present in subsequent reaction lanes. As such, ligation efficiency can be assessed by the presence or absence of separate A and B oligos, that become the higher weight 46-nt species when ligated.

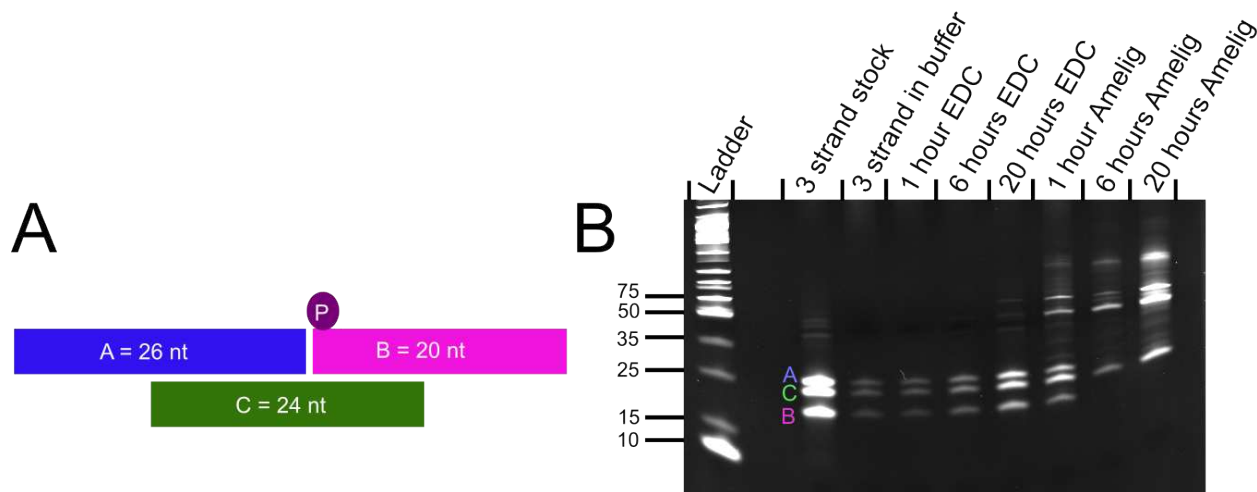


Figure 3.5: Splint ligation assay. Each reaction had 20 μM of the splint ligation duplex. EDC ligations had 45 mg/mL EDC dissolved in the crosslinking buffer. Enzymatic ligation reactions had 7.6 μM AmeLig and 1 mM ATP added to the crosslinking buffer. **A.** Splint ligation oligos and schematic. **B.** Denaturing gel electrophoresis shows the 3 strand stock with A, B, and C marked. Next is the three strand stock in the crosslinking buffer (1 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, 50 mM MES pH 6.0). The next three lanes are three separate EDC ligation reactions stopped at three time points: 1 hour, 6 hours, and 20 hours. The next three lanes are three separate AmeLig + ATP ligation reactions stopped at three time points: 1 hour, 6 hours, and 20 hours. The final

two lanes (6 hours and 20 hours elapsed in the AmeLig reaction) show complete depletion of the A and B species.

This splint ligation led to two essential conclusions. Firstly, AmeLig showed substantive activity in the CC1 crosslinking buffer. Secondly, AmeLig is far more efficient at ligating DNA free in solution than EDC. After 20 hours, the EDC reaction shows strong band intensity corresponding to the separate A and B oligos. In contrast, the AmeLig reaction shows no visible bands corresponding to the separate A and B oligos after just 6 hours. This effective ligation using AmeLig in the CC1 crosslinking buffer facilitates the transition from free solution splint ligation to testing *in crystallo* ligation.

3.2.5 Attempted Enzymatic Ligation in CC1+10

As mentioned above, steric calculations using our computational tool GuestScan model builder using a crystal structure of AmeLig bound to dsDNA (pdb code 6GDR) predicted that AmeLig would have sufficient room to bind within the CC1+10 lattice (Fig. S3.12) (3, 19). Crystals with a CG rich 10 base pair expansion and a 1 sticky base overhang were grown to potentially serve a dual purpose as restriction enzyme HhaI substrates (HhaI recognizes and cleaves GCG/C) (Table S3.1) (20). The structure of these crystals was determined via XRD at a resolution of 3.3 Å (Table S3.2). We tested AmeLig's ability to ligate dsDNA building blocks free in solution and observed remarkable ligation (Fig. 3.6 A). Thus, to assess *in crystallo* ligation we implemented stringent wash steps to differentiate ligation products that were free in solution from crystal templated ligation. After crystals were looped into the crosslinking buffer supplemented with 7.6 µM AmeLig and 1 mM ATP and left to react overnight they were looped into a wash solution (with the same salts and buffers as the crosslinking solution). After allowing the crystals to equilibrate for 10 minutes they were looped into a second wash solution and

allowed to equilibrate again. After the double wash, crystals were dissolved and analyzed via denaturing gel electrophoresis (Fig 3.6 B).

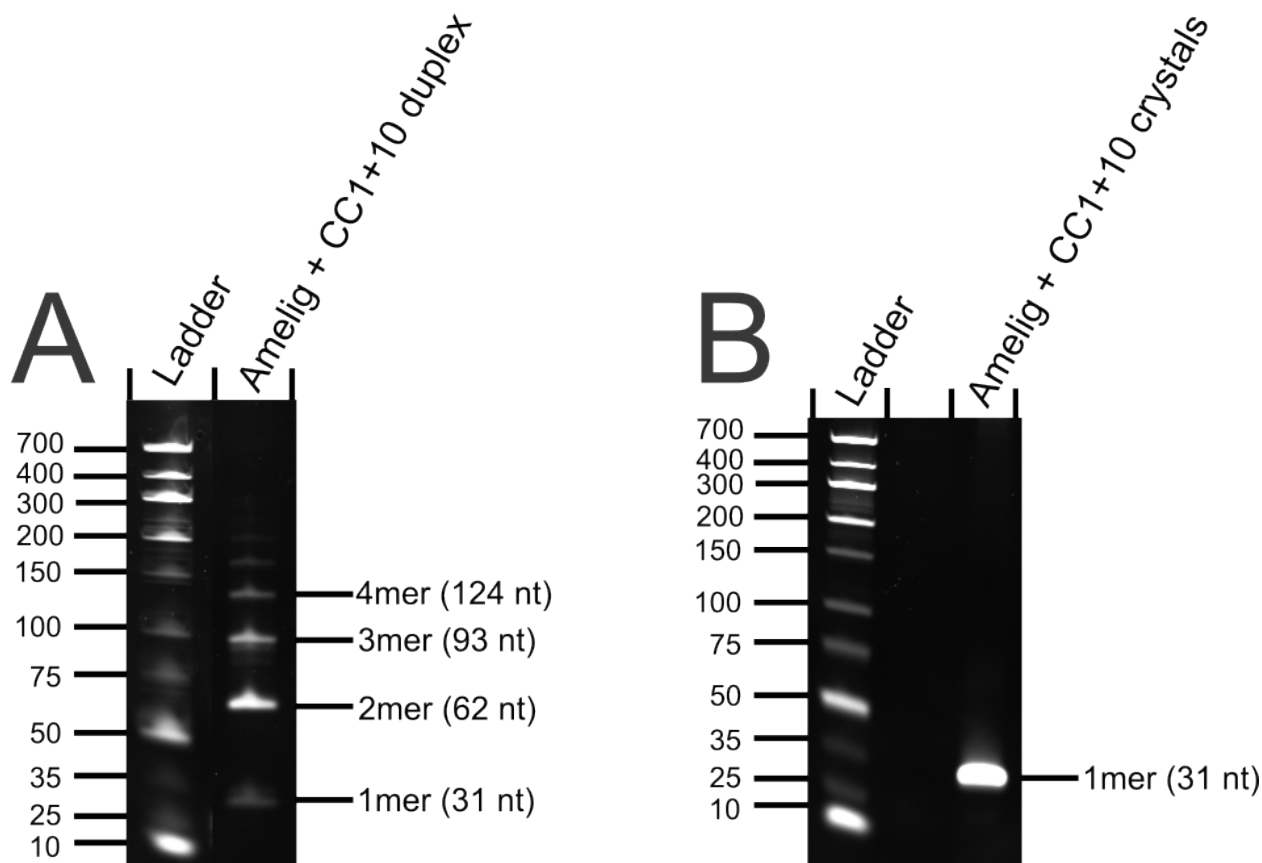


Figure 3.6: Enzymatic ligation on CC1+10 duplexes in free solution and *in crystallo*. **A.** Free solution CC1+10 building blocks with 1 sticky base overhang were reacted overnight with AmeLig. **B.** CC1+10 crystals of the same building blocks from solution described in A. were reacted overnight with AmeLig, washed twice, then dissolved.

From this comparison we can see that AmeLig works on the crystal building block substrate in solution but not in the crystal. Despite the predictive modelling showing several possible binding sites for AmeLig, this lack of *in crystallo* ligation yield is likely due to steric occlusion within the lattice. The 5.3 nm diameter pore size is an average/estimate based on MAP_CHANNELS calculations and local geometry at junctions may be more constrained (16). To determine if the pore size contributes to the lack of ligation within the crystal structure, we

next investigated enzymatic ligation with CC1+21. CC1+21 contains 21 basepairs of expansion DNA, resulting in ~8.5 nm diameter solvent channel pores (pdb 9YZK) (Fig. S3.1).

3.3 Ligation in CC1+21 Crystals

3.3.1 Design and Growth of CC1+21 Crystals

To design CC1+21 for enzymatic ligation and potential HhaI cleavage, we built upon the successful CC1+21 sequence from pdb 9YZK. Oligos had 2 sticky base overhangs, 5' P, and CG content at the edges of the symmetric expansion (Table S3.1). Symmetrically expanded CC1 building blocks distribute the expansion basepairs on either side of the centrally positioned RepE54 bindings site. In comparison, asymmetrically expanded CC1 building blocks have all the expansion content on one side of the RepE54 binding site, positioning the nick closely to the RepE54 binding site on the adjacent asymmetric unit.

Repeating successful growth conditions for 9YZK with the exact sequence from 9YZK yielded crystals with a typical parallelepiped habit, typically diagnostic of porous CC1 (Fig. S3.20). When the newly designed oligos containing the HhaI binding sites were used in the same crystallization screen, we observed triangular crystals, an unexpected crystal habit (Fig. 3.7). We have not yet validated the structure of the triangular crystals via XRD.



Figure 3.7: CC1+21 with 2 sticky base overhangs and high GC content at the junctions. Crystals formed a triangular habit.

3.3.2 EDC Chemical Ligation in CC1+21

Previous work in our lab has found that EDC does not ligate crystal building blocks (that have short sticky base overhangs) free in solution, and the splint ligation results presented here supports the need for crystal templating for *in crystallo* EDC ligation. To assess whether DNA-DNA junctions in the triangular CC1+21 were positioned close enough for ligation we treated crystals with one dose of 45 mg/mL EDC, washed, then dissolved crystals for analysis on denaturing gel electrophoresis. Despite the crystal habit deviating from the typical porous CC1, EDC was able to produce higher order ligation species, suggesting that the crystal packing contains DNA-DNA junctions (Fig. 3.8, densitometry fit in Fig. S3.16). Though 92% of building blocks remained monomers, we do observe 2-mer and 3-mer species (8% and 1%, respectively) (Table 3.3). We repeated this process with 9YZK crystals, that are confirmed to form the porous CC1+21, and observed similar ligation yields (Fig. S3.20-23). This suggests that despite the differing crystal habit, the triangular CC1+21 may be forming the target porous lattice.

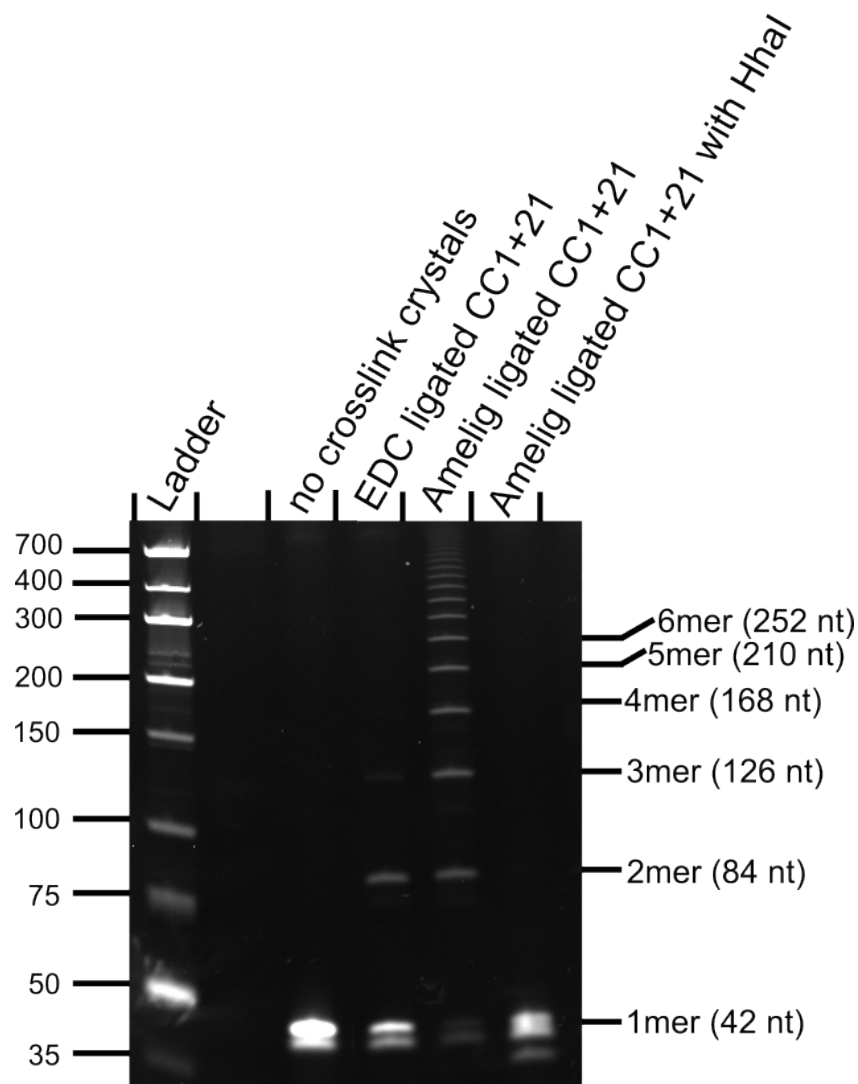


Figure 3.8: CC1+21 ligation and cleavage. Denaturing gel electrophoresis of triangular 21 bp expanded CC1. The first experimental lane contains dissolved crystals that weren't crosslinked. The next lane is EDC ligated dissolved crystals, followed by AmeLig ligated crystals that were washed twice before being dissolved. The last lane contains crystals that were ligated with AmeLig and then treated with HhaI endonuclease. Densitometry fits can be found in Fig. S3.16 and S3.17.

Table 3.3: Distribution of DNA block sizes in CC1+21 crystals as grown, after EDC ligation, after AmeLig ligation, and after AmeLig ligation followed by HhaI digestion.

CC1+21 Ligation and Cleavage				
DNA Block Size	No Crosslink [%]	EDC [%]	AmeLig [%]	HhaI [%]
1	100	92	42	100
2	0	8	22	0
3	0	1	14	0
4	0	0	8	0
5	0	0	5	0
6	0	0	3	0
7	0	0	2	0
8	0	0	2	0
9	0	0	1	0
10	0	0	1	0
> 10	0	0	2	0

3.3.3 Enzymatic Ligation

As seen in Fig. 3.8 and Table 3.3, AmeLig can ligate within the CC1+21 crystal. After treatment with 7.6 μ M AmeLig and 1 mM ATP in the crosslinking buffer the crystals were double washed before being dissolved. Denaturing gel electrophoresis shows that only 42% of building blocks remain monomers, while there are observable bands up to 10-mers and larger (densitometry fit in Fig. S3.17). This drastic difference from the attempts to ligate within CC1+10 indicates that indeed the smaller pore size seems to be a limiting factor for AmeLig to function correctly. Interestingly, when we trace labelled AmeLig with NHS fluorescein and added it to solution with CC1+10 and CC1+21 we observed fluorescence accumulate within both crystals (Fig. S3.18). This is in accordance with predictive modelling using in house tool GuestScan that showed AmeLig could diffuse through the solvent channel pores of CC1+10 based on size alone, and points to a steric issue in the CC1+10 design. Perhaps shifting the nick site of the CC1+10 design could rescue AmeLig function in the smaller pore size.

CC1+21 building blocks are 42 bp long, with a 21 bp RepE binding site positioned between flanking 10 and 11 bp expansion regions. To test whether this larger building block, that presumably would tumble slower when free solution, would still be readily ligated we added 7.6 μ M AmeLig and 1 mM ATP to solution with duplexed CC1+21 building blocks. As described above, these oligos have 2 sticky base overhangs, 5' terminal phosphates, and high CG content at the junctions. After observing substantial ligation over night at room temperature, we then repeated the experiment at 40 °C (Fig. S3.19). Despite the building blocks having only 2-nt overhangs to bind each other, at 40 °C we still observed extensive ligation, with almost all the 1-mer building blocks depleted. At 53 °C we observed almost no ligation, with a weak band at the 2-mer location. When CC1+21 crystals were incubated with AmeLig and ATP at 53 °C, the crystals dissolved and significant protein precipitation was observed.

3.3.4 Restriction Enzyme Digestion

Predictive modelling using GuestScan and a crystal structure of HhaI bound to dsDNA (pdb code 6UKF) did not return any registers within the CC1+10 system, and several in CC1+21 (Fig. S3.14 and S3.15) (21). Of note, GuestScan is built to assess the steric fit for scaffold assisted structural determination, meaning it considers steric clashes between symmetry mates of the guest protein. This means successful binding registers would sterically allow the guest to bind with 100% occupancy. For enzymatic function, 100% concurrent occupancy is not required. Thus, GuestScan can be useful to find binding registers for functional enzymes but may not provide a binary yes/no on steric fit when the assay for this fit is a measure of enzymatic function.

HhaI is a commercially available Type II endonuclease that is known to bind and cleave the sequence 5' GCG/C 3' on both strands, creating a double stranded break with two sticky base overhangs (21). Type II endonucleases use a divalent metal counterion as a cofactor, typically Mg(II) (22). We tested the function of HhaI on ligated crystal building blocks free in solution. These building blocks were designed with the HhaI binding sequence at the nick site, so that when ligated by AmeLig and then cleaved by HhaI we obtain back a 1-mer building block. Due to the proprietary nature of the optimization the commercially obtained HhaI has undergone, we included the suggested 1X NEBuffer in the mother liquor solution. When we added ligated building blocks (crystal monomers with dsDNA alone, no RepE54) and HhaI to this mother liquor with cutting buffer solution, we did not observe a change in building block distribution after it was left to react overnight, indicating it was not actively cleaving building blocks. In contrast, when HhaI and ligated building blocks were added to deionized water supplemented with 1X NEBuffer, all higher order species were cleaved and a prominent band at the 1-mer height was observed.

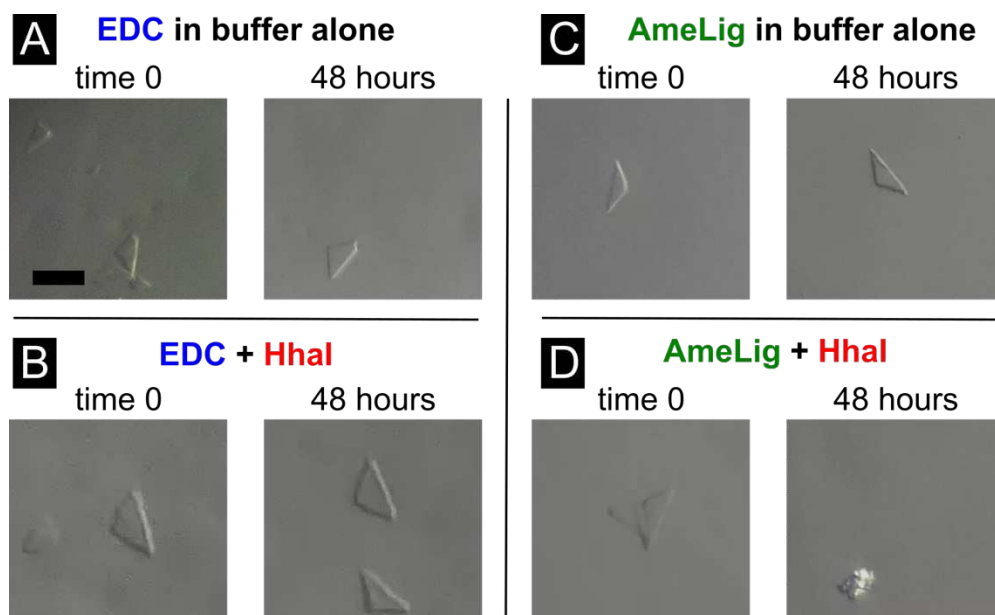


Figure 3.9: Crystal stability in water based cutting buffer. Scale bar is 100 μm . Crystal stability in water based cutting buffer. **A.** CC1+21 ligated with EDC in the absence of HhaI before and after 48 hours in solution. **B.** CC1+21 ligated with EDC in the presence of HhaI before and after 48 hours in solution. **C.** CC1+21 ligated with AmeLig in the absence of HhaI before and after 48 hours in solution. **D.** CC1+21 ligated with AmeLig in the presence of HhaI before and after 48 hours in solution.

As mentioned previously, crosslinked crystals stability outside of mother liquor conditions must be considered when moving crystals to different solutions. Fortunately, once CC1+21 crystals had been ligated (either by EDC or by AmeLig) they did not dissolve in the deionized water based cutting buffer (Fig 3.9). In comparison, when sister crystals (from the same batch) were added to the cutting buffer with HhaI and allowed to incubate for 48 hours, the AmeLig ligated-HhaI digested crystals dissolved (Fig. 3.9). The EDC ligated-HhaI digested crystals did not dissolve. The lack of dissolution after EDC ligation bolsters the idea that EDC is providing reinforcement beyond the DNA-DNA junction ligation. Fig 3.7 shows that once AmeLig ligated crystals were incubated with HhaI, no higher order species remain. It is important to note, all the crystals in these digestion trials were washed thoroughly after undergoing ligation before being added to the digestion reaction - this ensures that any ligation products that were free in solution were not still present when crystals were digested. Since these

crystals did not dissolve in the water-based cutting buffer in the absence of HhaI, it can be deduced that HhaI is acting directly on crystal building blocks within the crystal, not simply on building blocks free in solution from dissolved crystals. Additionally, we confirmed that HhaI was cleaving GCG/C specifically. 9YZK crystals, that do not contain GCG/C in the expansion sequence, were grown and ligated with AmeLig. After incubation with HhaI in the water-based cutting no cleavage of higher order species was observed (Fig. S3.20).

3.4 Discussion and Conclusions

3.4.1 Chemical Ligation in CC1+10

In accordance with previous studies from our lab, we see that junction architecture has great effect on relative EDC ligation yields. While we saw in the ipCC1 system 2 sticky base overhangs with a 3' terminal phosphate yielding the highest percentage of ligated building blocks, in porous CC1 we saw 1 sticky base with a 3' terminal phosphate with the highest percentage of ligation. The increased apparent ligation yields when the phosphate is located on the 3' termini is a previously known phenomenon in EDC DNA ligation (9, 12, 13). The change from 2 sticky base to 1 sticky base showing the highest yield in our porous system is likely due to local phosphate position differences between ipCC1 and porous CC1. While the 1 sticky base 3'P variant showed the greatest yield of ligation, there is a tradeoff. 5'P variants of CC1+10 have been observed to grow more rapidly than 3'P variants. When considering the CC1 platform for scalability as a high throughput crystallization scaffold, ease of crystal growth becomes a significant consideration. Another consideration for employing EDC ligation in crystal stabilization is balancing the concentration of EDC doses with the resulting quality of the crystal structure. Too much EDC results in visible cracking of the crystal, while ligation yields may be higher. Therefore, downstream optimization should be focused on screens for expediting 3'P

variant growth, as well as buffer conditions that improve ligation chemistry without being overly harsh on the crystal.

3.4.2 Enzymatic Ligation in CC1+10 and CC1+21

This chapter demonstrates the first successful enzymatic ligation in the CC1 system. We also see here that pore size estimates and predictive modelling can be suggestive of steric fit within the crystal pores, but that functionality of enzymes may require more space. Despite modelling that shows AmeLig has ample space within CC1+10 - particularly when considering that one enzyme may function on many building blocks (and does not need to occupy all building blocks simultaneously) - we saw failure of the enzyme. When all other variables are held constant (buffer components, enzyme concentration, crystal growth conditions) but the pore size is increased, we see rescue of AmeLig function. This strongly implies that steric fit at the DNA nick site was the hindering factor in CC1+10. We have seen that varying sticky base overhangs and nick site location is quite feasible in the CC1+10 system. Future designs that move the nick site within the expansion region may allow AmeLig to function properly within the smaller pore size CC1. If we consider CC1 from a nanotechnology perspective, where enzymes could be installed for downstream functions such as biosensing, a deeper understanding of the steric constraints in crystal pores allows more precise crystal engineering.

3.4.3 Restriction Enzyme Digestion of CC1+21

We sought to demonstrate that two enzymes can function within CC1+21, a ligase to fortify the crystal structure and an endonuclease to reverse that fortification in a controlled manner. This cleavage that is specific to the DNA components of CC1 has already given rise to an important result: when crystals are ligated via DNA-DNA junctions alone, once digested by

HhaI they are vulnerable to dissolution. Yet when crystals are ligated via EDC, they are not similarly vulnerable after HhaI digestion. This evidences the hypothesis that EDC may be acting on the protein:protein interfaces of CC1, despite a lack of structural observation of any specific crosslink in crystallographic models. Though this off-target crosslinking may be non-specific, its function to rigidify the CC1 lattice may be integral to retaining high resolution XRD data as we further develop the scaffold assisted structural determination capabilities of the platform.

3.4.4 Future Directions

There are two branches of CC1 applications that clearly benefit from the work presented in this chapter: 1. Scaffolded structural biology and 2. Protein-DNA nanotechnology. Firstly, in the structural biology field, if further development of CC1 is to enable further innovations by allowing determination/observation of guest proteins without the need to directly crystallize that guest, we must fortify the scaffold/parent lattice. Our previous manuscript in this area makes use of a buffer that contains KCl, CaCl₂, glycerol, and Tris HCl in a narrow range of concentrations. This buffer allowed EDC ligated CC1 crystals to retain high resolution XRD data (~ 3 Å), while facilitating guest protein binding at high enough occupancy to observe their structures in the resulting electron density maps. To apply scaffold crystallography to proteins who function in disparate solution conditions, the structure of CC1 needs reinforcement. Here we see how gains in DNA-DNA ligation can be achieved through variance in the sticky base overhangs and phosphate location, but future work to optimize EDC ligation chemistry across the protein:protein interfaces may have a more substantial effect on crystal stability and survival in varied guest-loading solutions.

Secondly, the verification of ligase and nuclease activity within CC1 points toward a broader opportunity to engineer dynamic crystalline biomaterials. Future work could explore

whether ligation and restriction digestion can be coupled within pre-formed co-crystals to enable selective removal and replacement of DNA building blocks. In this model, nuclease cleavage would generate noncovalently retained or partially-detached lattice components, while ligase activity would stabilize newly incorporated DNA segments. Establishing this type of enzyme-mediated exchange would demonstrate that CC1 is not merely a static crystalline scaffold, but a dynamic biomaterial capable of framework renewal, repair, or functional remodeling after crystallization.

3.5 Materials and Methods

3.5.1 Protein cloning, expression, purification

Cloning and expression of the protein Replication initiator RepE54 was previously described by Komori *et al.* (23). Cloning and expression of the protein AmeLig was previously described by Williamson *et al.* (19). The protein was overexpressed with a T7 promoter in *E. coli* BL21(DE3) competent cells (New England Biolabs #C2527H). Upon addition of 0.5 mM IPTG, the cells were incubated at 25 °C for 20 h. The cell pellets were sonicated in lysis buffer (1X PBS, 300 mM NaCl, 25 mM imidazole, pH 7.4) and applied to HisTrap (HisPur™ Ni-NTA resin, ThermoFisher #88222) equilibrated with HisTrap buffer (1X PBS, 300 mM NaCl, 25 mM imidazole, pH 7.4). The protein was eluted with 100 mM imidazole in HisTrap buffer. Fractions containing the desired protein were pooled and further purified at the CSU Histone Source facility with HiLoad Superdex 200 PG column (Cytiva #28989335). RepE54 was concentrated to 12 mg/mL in CC1 storage buffer (100 mM sodium citrate pH 6.2, 100 mM KCl, 10 mM MgCa₂, and 10% glycerol), flash frozen in liquid nitrogen, and stored at –80°C. AmeLig was concentrated to 10 mg/mL in CC1 storage buffer (100 mM sodium citrate pH 6.2, 100 mM KCl, 10 mM MgCa₂, and 10% glycerol), flash frozen in liquid nitrogen, and stored at –80°C.

3.5.2 Duplex annealing and co-crystal complexing

Individual oligomers were synthesized and HPLC purified by Integrated DNA Technologies. The DNA oligomer sequences used for co-crystallization are listed in Tables S3.1. Each oligomer contained the 19-bp iteron sequence TGTGACAAATTGCCCTCAG, but typically also contained the core 21-bp DNA block used in PDB:1REP from Komori *et al.* (23): CCTGTGACAAATTGCCCTCAG (omitting a dangling 3' T). In Table S3.1, the standard 21-bp

footprint is shaded in gray, and (purposeful) deviations are shaded in cyan. Flanking variable DNA sequences depended on the expansion scheme and desired sequence. The oligomers were resuspended in a standardized buffer (henceforth CC1 oligo buffer: 50 mM Tris HCl, 100 mM KCl, pH 6.7) and combined in equal molar ratio (1:1) with the complementary strand. The strands were annealed by heating to 94 °C for 2 min and slowly cooling to room temperature over approximately 60 minutes. The final concentration of all CC1 dsDNA blocks following duplexing was 4 mM. The scaffold protein and DNA were mixed in a 1.2:1 ratio (440:364 µM, 40 µL total) and were incubated on ice for 30 minutes prior to crystallization via sitting drop vapor diffusion.

3.5.3 Crystallization

Crystallization trials used CrysChem sitting-drop plates (Hampton Research HR3-159), with 400 µL reservoir solution per well and 4 µL crystallization drops prepared by mixing 2 µL each of the protein–DNA complex and reservoir solution directly on the platform; drops were mixed by gentle pipetting. Plates were sealed immediately after setup and incubated at 25 °C under static conditions. Crystallization conditions for the porous lattices were 10–80 mM Magnesium Acetate, 1–1.8 M Lithium Sulfate, and 50 mM MES pH 6.5 (or occasionally 50 mM sodium cacodylate pH 6.5). A detailed list of individual crystals, PDB codes, and corresponding growth conditions are reported in Table S3.3. Crystals grew to full size in a range of 24 hours to 30 days (most often in 1–2 days). Crystal images were taken with a Moticam X5 Plus camera attached to a Motic SMZ-168 stereozoom microscope.

3.5.4 EDC Crosslinking

Co-crystals were washed in conditions like their growth conditions (1-2.0 M Lithium sulfate, 20-50 mM Magnesium acetate, and 50 mM MES, pH 6.5). The co-crystals were washed in 200 μ L on 9-well glass plates (Hampton) to remove additional protein and DNA monomers. Note: For EDC ligation it is important to avoid buffers with 1°/2° amines or carboxylic acids. 1-Ethyl-3(3-dimethylaminopropyl)carbodiimide (EDC) (Advanced Chemtech CAS#:25952-53-8) was resuspended in the crosslinking solution (1.0 M Lithium sulfate, 50 mM Magnesium acetate, and 50 mM MES, pH 6.0 unless otherwise noted) to final concentration values of 30-60 mg/mL and used immediately. The co-crystals were crosslinked in a 200 μ L EDC solution volume overnight (16-24 hrs). If two doses of EDC were used, the first solution was pipetted off and replaced with a freshly prepared EDC solution. The crosslinking reaction was ended by looping the co-crystals back into the wash solution. If crystals were dissolved for gel electrophoresis then they were looped into 40 μ L dissolution buffer, 3.5 M Urea in 1X TBE supplemented with 20 μ g of proteinase K.

3.5.5 AmeLig Crosslinking

Co-crystals were looped into a 40 μ L drop of crosslinking solution (1.0 M Lithium sulfate, 50 mM Magnesium acetate, and 50 mM MES, pH 6.0 unless otherwise noted) supplemented with AmeLig at a final concentration of 0.25 mg/mL and ATP at a final concentration of 1 mM on 9-well glass plates (Hampton). Crystals were left to ligate overnight. The reaction was stopped by washing crystals. Washing consisted of looping into a fresh 40 μ L of crosslinking buffer (with no ligase or ATP), allowing crystals to equilibrate for 1 hour. Crystals were then looped into a second 40 μ L of fresh crosslinking buffer.

3.5.6 HhaI Digestion

Ligated co-crystals (either via EDC or AmeLig) were looped into 40 μ L drop of cutting buffer (1X NEBuffer in DI water) supplemented with 1 μ L HhaI (NEB R0139S) at concentration 20,000 units/mL on 9-well glass plates (Hampton). If crystals were not dissolved after digestion the reaction was stopped by looping into 40 μ L dissolution buffer, 3.5 M Urea in 1X TBE supplemented with 20 μ g of proteinase K. Crystals already dissolved after HhaI digestion had 5X DNA binding buffer (Zymo D4029) added to solution in preparation for oligo cleaning before gel electrophoresis.

3.5.7 DNA Gel Electrophoresis and Densitometry

Cross-linked co-crystals were dissolved in 1 $\times$ TBE supplemented with 3.5 M urea and 20 μ g of proteinase K. Dissolved oligos were cleaned using Zymo DNA Clean and Concentrate (Zymo D4029). Prior to loading oligos were normalized to each other using Nanodrop One DNA concentration readings. The crystals were analyzed with denaturing gels (6% Novex TBE-Urea Polyacrylamide Gel Electrophoresis) formulated with 7 M urea and run with 1 $\times$ TBE running buffer. Splint ligation oligos were analyzed with 10% Novex TBE-Urea Polyacrylamide Gel Electrophoresis. The co-crystal DNA ladder was the TriDye Ultra Low Range DNA Ladder (NEB). Gels were incubated with 1 $\times$ SYBR Gold Nucleic Acid Gel Stain and imaged with a FluorChem E System by ProteinSimple on the ethidium bromide setting. For gel densitometry, we used the technique established in our previous ligation study (Python scripts for fitting densitometry peaks with Gaussians) (7, 8).

3.5.8 NHS-fluorescein Conjugation

AmeLig or HhaI was dialyzed into labeling buffer (50 mM Na₂PO₄, 300 mM NaCl, pH 8.1) using Slide-A-Lyzer™ dialysis pucks (1 kDa MWCO; ThermoFisher) to remove free amines and

equilibrate the sample for N-hydroxysuccinimide (NHS) conjugation. The protein concentration was 114 μ M (AmeLig) and 180 μ M (HhaI) prior to labeling. NHS-fluorescein (Thermo Scientific; 5 mg/mL, ~10 mM) was added at a 1:5 dye:protein molar ratio, and the reaction was incubated overnight at 4 °C with gentle agitation. To minimize photobleaching, dye-handling steps were performed in opaque 1.5 mL microcentrifuge tubes and under low-light conditions. Unreacted dye was removed by dialysis against fresh labeling buffer (50 mM Na_2PO_4 , 300 mM NaCl, pH 8.1, 1 kDa MWCO) with multiple buffer exchanges until no residual dye was visible in the dialysate. The labeled protein was concentrated to the working concentration using a 3 kDa MWCO Amicon centrifugal filter unit (Millipore) and stored on ice for immediate use, or aliquoted and flash-frozen for long-term storage at -80 °C.

3.5.9 Confocal Microscopy

Crystals loaded with NHS trace labelled AmeLig or HhaI were imaged using a Nikon Eclipse Ti Yokagawa CSU-X1 spinning-disk confocal microscope equipped with an AndoriXon Ultra 879U EMCCD camera. Fluorescein signal was collected using 488 nm laser excitation and a 515 nm emission filter. Image acquisition and processing were performed using Nikon NIS-Elements software.

3.5.10 Guest Loading

Crosslinked co-crystals were washed at room temperature to remove components that could interfere with protein binding and to equilibrate the crystals to the guest loading soak solution (1.2 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, 50 mM MES pH 7.1). Equilibrated crystals were moved to 9 μ L fresh buffer solution in a micro-Insert 4 Well in μ -Dish (Ibidi Cat.No:80406). Proteins (either AmeLig or HhaI, combined 20:1 unlabeled to trace labeled) of 20 μ M were added on top of the

well. The co-crystals sitting in guest loading solutions were sealed for 24-72 hours to allow guest species to diffuse and reach binding equilibrium within the crystal lattice.

3.5.11 X-ray Diffraction collection and refinement

Crystals were flash frozen in liquid nitrogen after briefly transferring the crystal to a cryo-protectant solution (a 1:1 mixture of the reservoir solution and 50% glycerol stock for a final concentration of 25% glycerol). Single crystal X-Ray diffraction data was collected at APS beamline 24-IDE (Dectris EIGER2 detector) from 0 to 180° with an omega delta of 0.2° and an exposure of 0.5 s or 0.25 s. Data was processed with XDS (24) and refined using PHENIX and COOT (25, 26). The first porous CC1+10 dataset (PDB 9Z08) was refined using newly generated RFree flags. All subsequent isomorphous structures were refined using the same RFree flags to preserve a fixed cross-validation benchmark across datasets. Decisions about truncating the structure factor data were made on a case-by-case basis, jointly considering the following thresholds $CC1/2 > 0.3$, mean $\langle I/\sigma(I) \rangle > 1.0$, and completeness $> 85\%$ as target guidelines for the highest-resolution shell. Most datasets satisfy all three thresholds in the highest-resolution shell.

CHAPTER THREE BIBLIOGRAPHY

1. Ward, A. R., & Snow, C. D. (2020). Porous crystals as scaffolds for structural biology. *Current Opinion in Structural Biology, Folding and Binding - Proteins*, 60, 85–92. <https://doi.org/10.1016/j.sbi.2019.12.008>
2. Orun, A. R., Shields, E. T., Dmytriw, S., Vajapayajula, A., Slaughter, C. K., & Snow, C. D. (2023). Modular Protein–DNA Cocrystals as Precise, Programmable Assembly Scaffolds. *ACS Nano*, 17(14), 13110–13120. <https://doi.org/10.1021/acsnano.2c07282>
3. Shields, E. T., Slaughter, C. K., Mekkaoui, F., Magna, E. N., Shepherd, C., Lukeman, P. S., Spratt, D. E., & Snow, C. D. (2026). Modular Scaffold Crystals for Programmable Installation and Structural Observation of DNA-Binding Proteins (p. 2026.03.04.709581). *bioRxiv*. <https://doi.org/10.64898/2026.03.04.709581>
4. Hiromoto, T., Ikura, T., Honjo, E., Blaber, M., Kuroki, R., & Tamada, T. (2022). Creation of Cross-Linked Crystals With Intermolecular Disulfide Bonds Connecting Symmetry-Related Molecules Allows Retention of Tertiary Structure in Different Solvent Conditions. *Frontiers in Molecular Biosciences*, 9. <https://doi.org/10.3389/fmolb.2022.908394>
5. Kuzmanic, A., & Zagrovic, B. (2014). Dependence of Protein Crystal Stability on Residue Charge States and Ion Content of Crystal Solvent. *Biophysical Journal*, 106(3), 677–686. <https://doi.org/10.1016/j.bpj.2013.12.022>
6. Nasiri, M., Bahadorani, M., Dellinger, K., Aravamudhan, S., Vivero-Escoto, J. L., & Zadehan, R. (2024). Improving DNA Nanostructure Stability: A Review of the Biomedical Applications and Approaches. *International Journal of Biological Macromolecules*, 260(Pt 1), 129495. <https://doi.org/10.1016/j.ijbiomac.2024.129495>
7. Kong, H., Sun, B., Yu, F., Wang, Q., Xia, K., & Jiang, D. (2023). Exploring the Potential of Three-Dimensional DNA Crystals in Nanotechnology: Design, Optimization, and Applications. *Advanced Science*, 10(24), 2302021. <https://doi.org/10.1002/advs.202302021>
8. Edeleva, E., Salditt, A., Stamp, J., Schwintek, P., Boekhoven, J., & Braun, D. (2019). Continuous nonenzymatic cross-replication of DNA strands with in situ activated DNA oligonucleotides. *Chemical Science*, 10(22), 5807–5814. <https://doi.org/10.1039/C9SC00770A>
9. Obianyor, C., Newnam, G., Clifton, B. E., Grover, M. A., & Hud, N. V. (2020). Towards Efficient Nonenzymatic DNA Ligation: Comparing Key Parameters for Maximizing Ligation Rates and Yields with Carbodiimide Activation. *ChemBioChem*, 21(23), 3359–3370. <https://doi.org/10.1002/cbic.202000335>
10. Nakajima, N., & Ikada, Y. (1995). Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. *Bioconjugate Chemistry*, 6(1), 123–130. <https://doi.org/10.1021/bc00031a015>
11. Lepvrier, E., Doigneaux, C., Moullintraffort, L., Nazabal, A., & Garnier, C. (2014). Optimized protocol for protein macrocomplexes stabilization using the EDC, 1-ethyl-3-

- (3-(dimethylamino)propyl)carbodiimide, zero-length cross-linker. *Analytical Chemistry*, 86(21), 10524–10530. <https://doi.org/10.1021/ac502561e>
12. Ward, A. R., Dmytriw, S., Vajapayajula, A., & Snow, C. D. (2022). Stabilizing DNA–Protein Co-Crystals via Intra-Crystal Chemical Ligation of the DNA. *Crystals*, 12(1), Article 1. <https://doi.org/10.3390/cryst12010049>
 13. Orun, A. R., Slaughter, C. K., Shields, E. T., Vajapayajula, A., Jones, S., Shrestha, R., & Snow, C. D. (2024). Tuning Chemical DNA Ligation within DNA Crystals and Protein–DNA Cocrystals. *ACS Nanoscience Au*, 4(5), 338–348. <https://doi.org/10.1021/acsnanoscienceau.4c00013>
 14. Lehman, I. R. (1974). DNA ligase: Structure, mechanism, and function. *Science*, 186(4166), 790–797. <https://doi.org/10.1126/science.186.4166.790>
 15. Williamson, A., & Leiros, H.-K. S. (2019). Structural intermediates of a DNA–ligase complex illuminate the role of the catalytic metal ion and mechanism of phosphodiester bond formation. *Nucleic Acids Research*, 47(14), 7147–7162. <https://doi.org/10.1093/nar/gkz596>
 16. Juers, D. H., & Ruffin, J. (2014). MAP_CHANNELS: A computation tool to aid in the visualization and characterization of solvent channels in macromolecular crystals. *Journal of Applied Crystallography*, 47(Pt 6), 2105–2108. <https://doi.org/10.1107/S160057671402281X>
 17. Shi, K., Bohl, T. E., Park, J., Zasada, A., Malik, S., Banerjee, S., Tran, V., Li, N., Yin, Z., Kurniawan, F., Orellana, K., & Aihara, H. (2018). T4 DNA ligase structure reveals a prototypical ATP-dependent ligase with a unique mode of sliding clamp interaction. *Nucleic Acids Research*, 46(19), 10474–10488. <https://doi.org/10.1093/nar/gky776>
 18. Li, Z., Liu, L., Zheng, M., Zhao, J., Seeman, N. C., & Mao, C. (2019). Making Engineered 3D DNA Crystals Robust. *Journal of the American Chemical Society*, 141(40), 15850–15855. <https://doi.org/10.1021/jacs.9b06613>
 19. Williamson, A., Grgic, M., & Leiros, H.-K. S. (2018). DNA binding with a minimal scaffold: Structure–function analysis of Lig E DNA ligases. *Nucleic Acids Research*, 46(16), 8616–8629. <https://doi.org/10.1093/nar/gky622>
 20. Applied Biosystems. (2002). Evaluating and isolating synthetic oligonucleotides (p. 20). Applied Biosystems.
 21. Horton, J. R., Yang, J., Zhang, X., Petronzio, T., Fomenkov, A., Wilson, G. G., Roberts, R. J., & Cheng, X. (2020). Structure of HhaI endonuclease with cognate DNA at an atomic resolution of 1.0 Å. *Nucleic Acids Research*, 48(3), 1466–1478. <https://doi.org/10.1093/nar/gkz1195>
 22. Pingoud, A., & Jeltsch, A. (2001). Structure and function of type II restriction endonucleases. *Nucleic Acids Research*, 29(18), 3705–3727. <https://doi.org/10.1093/nar/29.18.3705>
 23. Komori, H., Sasai, N., Matsunaga, F., Wada, C., & Miki, K. (1999). Crystallization and Preliminary X-Ray Diffraction Studies of a Replication Initiator Protein (RepE54) of the

- Mini-F Plasmid Complexed with Iteron DNA1. *The Journal of Biochemistry*, 125(1), 24–26. <https://doi.org/10.1093/oxfordjournals.jbchem.a022262>
24. Kabsch, W. (2010). XDS. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 125–132. <https://doi.org/10.1107/S0907444909047337>
25. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., ... Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*, 75(10), 861–877. <https://doi.org/10.1107/S2059798319011471>
26. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, 66(4), 486–501. <https://doi.org/10.1107/S0907444910007493>

CHAPTER 4: PROGRAMMABLE CAPTURE AND TOEHOLD MEDIATED RELEASE OF NUCLEIC ACID CARGO IN POROUS PROTEIN-DNA CO-CRYSTALS

4.1 Introduction

4.1.1 Porous protein-DNA co-crystals as programmable host materials

CC1 possesses large, continuous solvent channels that permit the diffusion of macromolecular species throughout the crystal lattice, making it a promising platform for crystalline storage of nucleic acid cargo. As discussed in previous chapters, CC1 can be made rigid and robust to the surrounding solution via chemical crosslinking and enzymatic ligation (1,2). Because the DNA binding sites are periodically repeated within the rigid protein-DNA scaffold, guest binding can occur without degrading the resultant lattice homogeneity (3,4). We have previously utilized diffusion through the lattice pores to introduce DNA-binding proteins as guest molecules for structural studies, where their high, lattice-wide occupancy generated sufficient signal for structural resolution. This behavior indicates that biomolecules with a defined installation site can be spatially organized with precision inside CC1. In this chapter, we extend this concept to nucleic acid sequestration by leveraging the sequence programmability of the DNA components of the protein-DNA building block. Specifically, we engineer solvent-exposed ssDNA tails that project into the crystal channels and function as modular capture elements for complementary ssDNA or RNA guests. Hybridization to these anchors affixes the cargo to the lattice, while the porous solvent channel architecture allows both loading and removal to occur by diffusion. Using this framework, we first demonstrate a reversible DNA capture-release cycle inspired by prior crystalline DNA devices, and then show controlled capture, storage, and toehold-mediated release of a biologically relevant dsRNA species,

miR-34a. Together, these results establish CC1 as a porous, sequence-modular crystalline reservoir capable of nucleic acid capture, storage, and controlled release, with clear relevance to therapeutic delivery applications.

4.1.2 From DNA crystal devices to functional guest handling in protein-DNA co-crystals

Prior work from the Mao and Seeman laboratories established that nucleic-acid-mediated state changes can be programmed to occur within a three-dimensional crystal lattice. In their 2017 paper, “A device that operates within a self-assembled 3D DNA crystal,” the authors showed that ssDNA tails designed to extend into solvent channels enable strand-displacement–based switching through the diffusion of guest DNA molecules (5). In this system, the kinetics of state change are dominated by diffusion through the crystal rather than by strand length, and XRD combined with kinetic modeling of macrolevel color change were used to monitor these transitions. However, these pioneering devices were built from pure-DNA tensegrity triangles and operated exclusively in their mother liquor, rendering them vulnerable to dissolution outside of growth conditions (6,7). The absence of wash steps to remove free building blocks necessitated large excesses of guest strands, and sequence specificity of release was not explicitly tested using scrambled controls.

In this chapter, we adapt this crystalline DNA cycling concept to porous protein–DNA co-crystals (Fig. 4.1). We build on the DNA device in two important aspects: 1. Chemical ligation converts the co-crystal from a dynamic assembly held together by non-covalent interactions into a biomaterial able to be moved between varied conditions, including wash solutions and guest loading soaks. 2. By studying scrambled versions of release strands, we demonstrate that cargo displacement and release are sequence specific (Fig. 4.3).

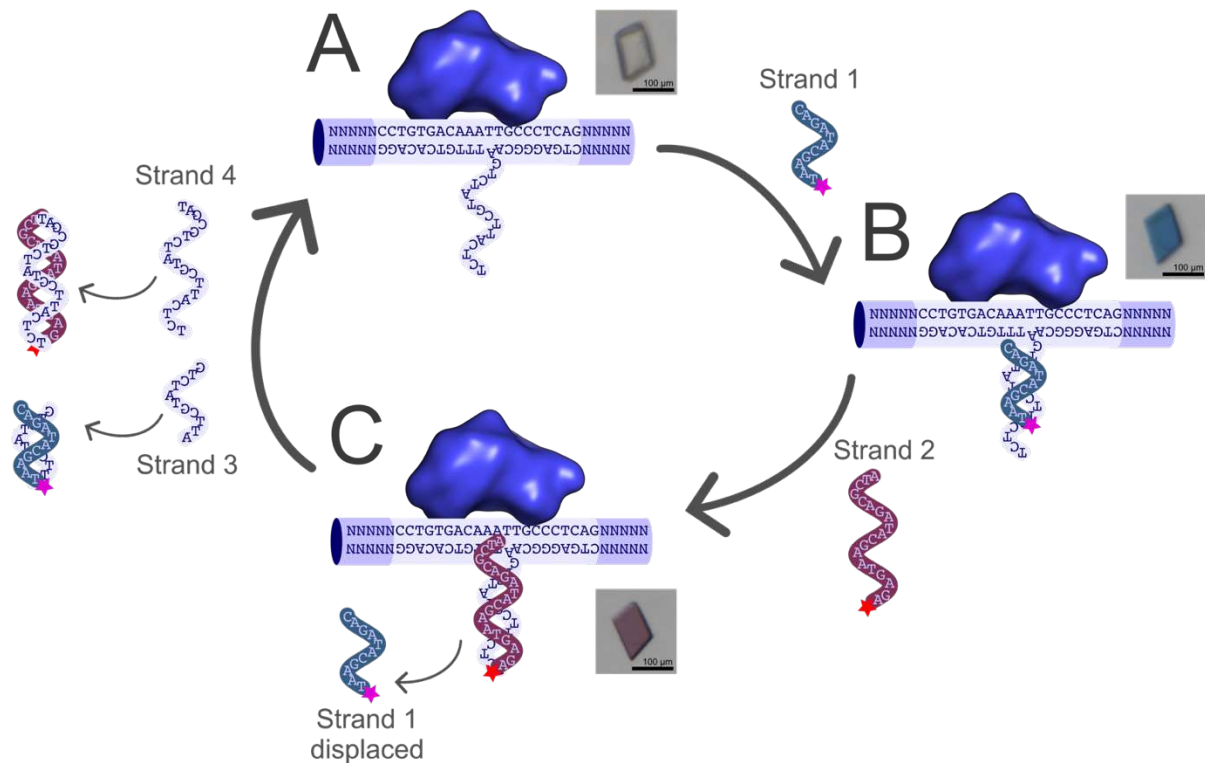


Figure 4.1: DNA device cycle in protein-DNA co-crystals. **A.** An unpaired 15-nt tail is incorporated under the RepE54 binding footprint. Strand 1 with a Cy5 fluor binds to the unpaired tail, leaving a 4-nt tail unpaired and coloring the crystal blue to the eye as seen in **B.** When Strand 2 with a Cy3 fluor is added it hybridizes with the full 15-nt tail and extending a 4-nt unpaired segment past the tail. This displaces Strand 1 and turns the crystal pink as seen in **C.** Strand 3 hybridizes with left over Strand 1, sequestering it. Strand 4 hybridizes with the unpaired 4-nt tail of Strand 2 and peels it off the 15-nt tail, cycling the crystal pack to optically clear as seen in **A.**

We demonstrate two types of nucleic acid release using toehold mediated strand displacement (TMSD) (Fig. 4.2). The three oligos participating in TMSD are the incumbent, that is part of the original duplex, the substrate strand that is also part of the original duplex, but contains an unpaired toehold region, and the invading strand that binds the substrate strand in the toehold region, promoting branch migration and subsequent dissociation of the incumbent strand (8, 9). In the DNA capture device, we demonstrate release of both ssDNA and dsDNA from the crystal by alternating the identity of the substrate strand (anchor strand vs guest strand). After establishing that CC1 is amenable to functioning as a device for capture, storage, and triggered

release of DNA species we look to installing and delivering a dsRNA functional nucleic acid guest, the broadly anti-oncogenic mir34a (Fig. SF 4.X) (10).

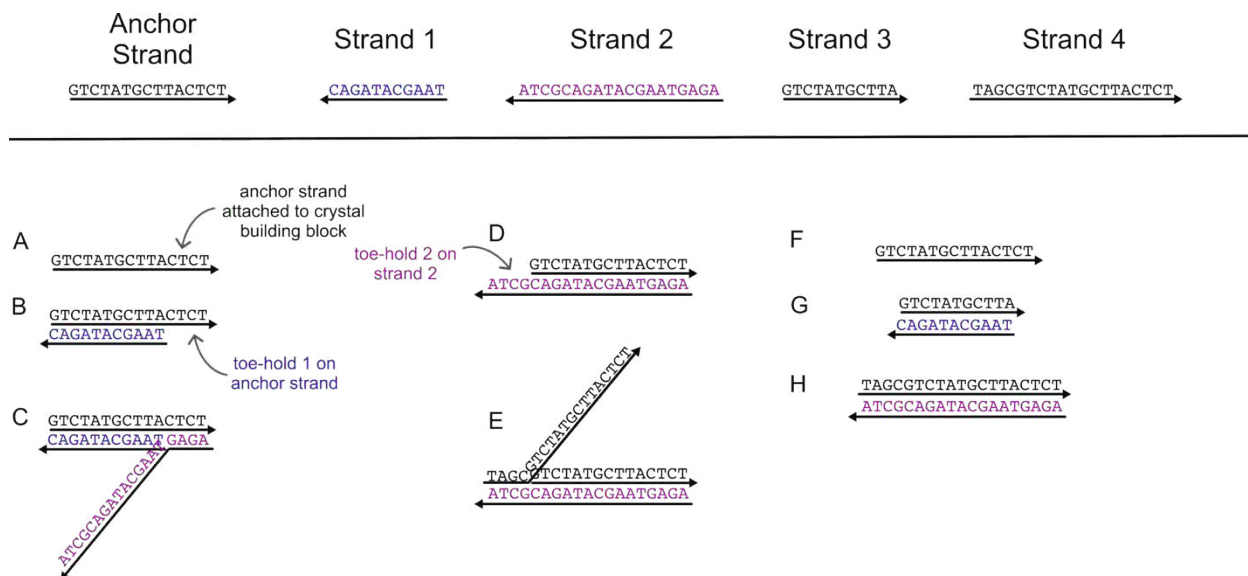


Figure 4.2: Strand displacement mechanisms in the DNA capture device. Arrows indicate 5' to 3' direction. **A.** The ssDNA 15-nt anchor strand extending from the DNA component of a CC1 crystal. **B.** Strand 1 (containing a Cy5 fluor) hybridizes with the anchor strand, leaving 4-nt of the anchor unbound. **C.** Strand 2 (containing a Cy3 fluor) invades and displaces Strand 1 using the 4-nt toehold, releasing ssDNA Strand 1 into solution. **D.** Strand 2 extends 4-nt beyond the anchor strand, leaving a second toehold. Strand 3 (ssDNA complement to Strand 1) is added to hybridize to and sequester any lingering Strand 1 within the crystal, resulting in the dsDNA in G. **E.** Strand 4 invades and binds the unpaired 4-nt on Strand 2, releasing dsDNA into solution, resulting in **F.** ssDNA anchor strand, **G.** Hybridized Strand 1 and Strand 3, and **H.** Hybridized Strand 2 and Strand 4.

4.1.3 Motivation for therapeutic RNA capture and release

miR-34a was selected as a benchmark RNA cargo for evaluating programmed capture and release in CC1 because it is both biologically well characterized and therapeutically relevant. miR-34a is a tumor-suppressive microRNA that was identified as a direct p53-responsive miRNA and has been implicated in p53-associated programs including cell-cycle arrest, apoptosis, senescence, and suppression of tumor-promoting pathways across multiple cancer contexts (10-18). Its translational relevance is underscored by MRX34, a liposomal miR-34a mimic that became the first miRNA mimic evaluated in human clinical trials, although that trial

was ultimately terminated early because of serious immune-mediated adverse events (19-21). These clinical results highlight both the promise of miRNA replacement therapy and the persistent delivery challenges associated with RNA therapeutics, including nuclease susceptibility, immune activation, off-target effects, and the need for carriers that preserve RNA integrity while enabling controlled release (22-26). Both relevant to therapeutic efforts, miRNA replacement and miRNA inhibition use different molecular formats. Synthetic miRNA mimics are typically double-stranded RNAs designed to resemble endogenous post-Dicer duplex intermediates, whereas anti-miR therapeutics are generally single-stranded antisense oligonucleotides that bind mature miRNAs and block their activity (27-32). Endogenous miRNAs are generated from hairpin precursors that are exported from the nucleus, processed by Dicer into duplex intermediates, and loaded into Argonaute-containing RISC, where one strand is retained as the guide strand and the passenger strand is discarded or degraded (33). Therefore, in this chapter, miR-34a is used to test whether the co-crystal platform can capture, retain, and release a fragile RNA cargo in a duplex-forming format relevant to miRNA mimic delivery.

4.2 Design of nucleic acid capture devices in porous protein-DNA co-crystals

4.2.1 Architecture of the capture-enabled co-crystal

The kinetics of TMSD have been extensively characterized (8, 9, 34-42). In pure-DNA reactions (where incumbent, substrate, and invading strands are all ssDNA oligos) and hybrid reactions (where one or more of the incumbent, substrate, and invading strands are ssRNA oligos) it's been reported that 5' toeholds (where the substrate contains a toehold at the 5' of the oligo) are faster than 3' toeholds when sequences are identical, but that changes in sequence identity drive speed of reaction more than location of toehold (34-37). The DNA capture device

presented here utilizes both 5' and 3' TMSD. The RNA capture device utilizes a 5' toehold where the substrate is RNA, the invading strand is RNA, and the incumbent is DNA (Fig. S4.1). Of interest, hybrid reaction rates are highly dependent on which strand contains RNA, as RNA:DNA and RNA:RNA complexes have differing free energy of hybridization (36, 43). The RNA device presented here has incumbent:substrate binding between RNA:DNA species, with an invading RNA strand. RNA:RNA complexes are more stable than RNA:DNA, meaning as branch migration occurs there is a downhill free energy landscape, increasing the rate of reaction (40). In contrast, branch migration in a pure-DNA TMSD is a 1:1 change in base pairing energy as branch migration proceeds. A commonality between pure-DNA and hybrid TMSD is the increase in rate of reaction as toehold length increases from 1 to 6-nt, with a plateau in rate change after 6-nt (36, 41). The DNA device presented here utilizes 4-nt toeholds as a proof-of-concept recapitulation of the 2017 Mao DNA device, while the RNA device utilizes a 7-nt toehold (5). This 7-nt toehold design builds on the successfully crystallized anchor length of the DNA device (15-nt) to bind 22-nt ssRNA miR34.

The successful growth of CC1 variants bearing 4-nt ssDNA tails is described in Chapter 2 of this thesis, where we demonstrated that short single-stranded extensions can be incorporated modularly without disrupting the expected crystal packing arrangement. A critical prerequisite for this compatibility was the Y184E mutation in RepE54, that abrogates its propensity to bind ssDNA concurrently with dsDNA and thereby prevents unintended protein-tail interactions (44). Building on these results, we sought to extend the length of the ssDNA tails incorporated into CC1 to increase the sequence complexity and specificity available for nucleic acid capture. In designing this capture element, careful consideration was given to the placement of the ssDNA sequence within the DNA building block. Based on the structures determined of previously

characterized 4-nt tail variants, we selected a register located in the middle of the 31 bp building block, directly beneath the RepE54 binding footprint. This placement is predicted to position the ssDNA tail along the crystallographic solvent channel axis rather than toward lattice contacts, minimizing steric interference with symmetry-related neighbors while maximizing solvent accessibility, thusly maximizing accessibility for hybridization with incoming nucleic acid guests (Fig 4.3). By situating the capture handle within the solvent channel while preserving the native packing interactions of the lattice, this design was anticipated to remain compatible with robust crystal growth and diffraction, while enabling the installation of modular, sequence-programmable tails within the porous scaffold.

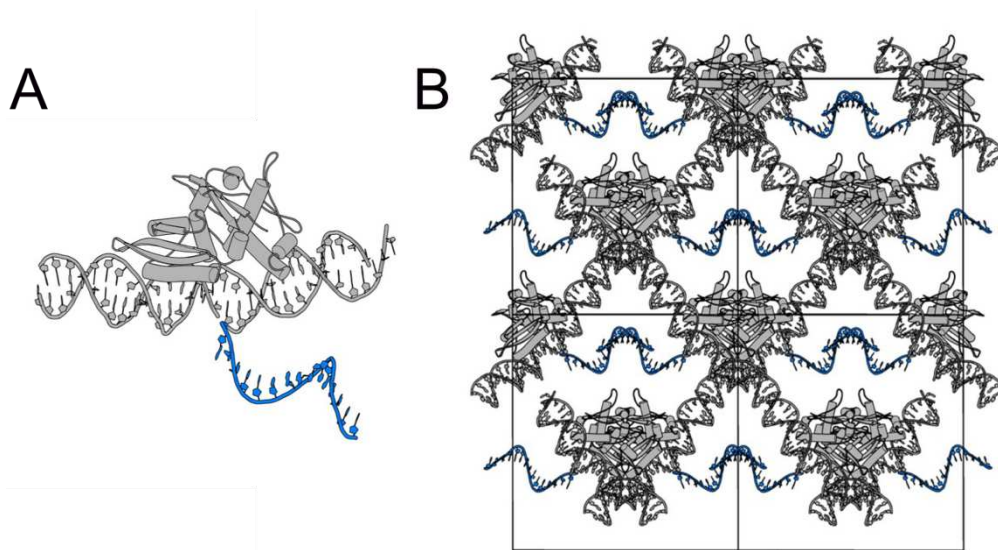


Figure 4.3: Hypothetical model of centrally placed 15-nt tail. **A.** AlphaFold3 prediction of RepE54 and the three strands of DNA composing the RNA capture device. **B.** Actual experimentally determined unit cell dimensions of RNA capture crystal with the AlphaFold3 model superimposed, showing how the 15-nt tail extends into the solvent channel pore when placed underneath the binding footprint of RepE54. Crystal models of tailed crystals did not have sufficient electron density signal to model the full 15-nt tail, indicating that the tail is labile in the solvent channel pore (Fig. S4.8)

4.2.2 DNA capture construct

As an initial test of functional equivalence between pure-DNA tensegrity triangles and the CC1 platform, we designed a CC1 variant incorporating the anchor sequence reported in the 2017 Seeman and Mao study (5). Robust crystal growth was achieved, and X-ray diffraction confirmed that the crystals adopted the expected packing arrangement, space group number 5 (Table S4.1). As shown in Fig. 4.1, crystals grown with the anchor strand unpaired are optically clear. Because the guest strands bind at lattice-wide, regularly spaced positions, fluorophore exchange on the anchor strands produces an optical color change visible to the eye. Upon addition of a Cy5-tagged complementary strand that leaves four nucleotides of the anchor unpaired, the guest strand hybridizes within the crystal lattice, imbuing a visible blue color in the crystal (Fig. 4.1, B), consistent with other direct visualizations of this dye in microscopy (45).

Confocal microscopy, performed using a 640 nm excitation laser and detecting in the 660 nm emission spectra, shows accumulation of the Strand 1 oligo after 18 hours of incubation in a 100 μ m bath of Strand 1 (Fig. 4.2). After transfer to a wash buffer (1.2 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, 50 mM MES pH 6.5) and equilibration for 18 hours, the fluorescent signal remains localized within the crystal pores, indicating stable retention of the bound guest strand. Introduction of a Cy3-tagged perfect complement to the 15-nt anchor strand, extending four nucleotides beyond the anchor for a total length of 19-nt, results in displacement of the Cy5-labeled guest (Fig. 4.2) and a corresponding shift in crystal color from blue to pink (Fig. 4.1, C) (45).

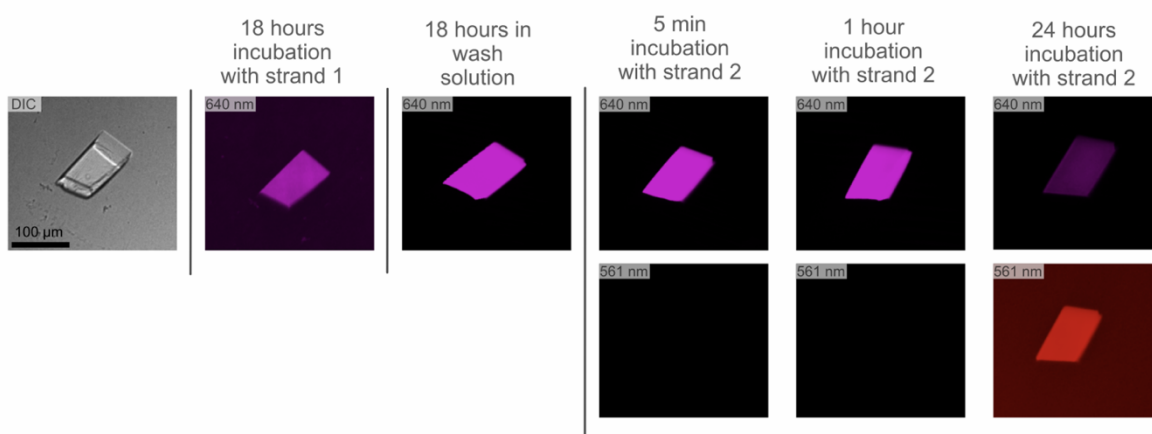


Figure 4.4: TMSD from Strand 1 to Strand 2. A crystal is incubated with 100 μ M Strand 1 and after 18 hours shows accumulation. The loaded crystal is transferred to a new drop and left to equilibrate for 18 hours. When 100 μ M Strand 2 is added to solution, Strand 1 is displaced and diffuses out of the crystal, equilibrating with the surrounding solution. Strand 2 is now bound within the crystal, showing emission when excited with a 561 nm wavelength laser.

To assess the sequence specificity of this displacement, a control experiment was performed in which a scrambled version of the Cy3-tagged strand was added to a crystal preloaded with strand 1. Under these conditions, no change in fluorescence or macro-scale crystal color was observed (Fig. S4.2). The lack of fluorescence change observed under scrambled-strand conditions provides evidence against nonspecific strand exchange or concentration-driven competitive displacement and indicates that guest release is mediated by sequence-specific hybridization.

To complete a full capture–release cycle, a perfect complement to strand 1 was first added to sequester any residual guest strand 1 within the pores. Subsequent addition of the full 19-nt perfect complement to strand 2 displaced the Cy3-tagged species from the anchor strand (Fig. 4.3). Once released, the guest strands diffuse out of the solvent channels and equilibrate with the surrounding solution, restoring the crystal to its optically clear state (Fig. 4.1, A).

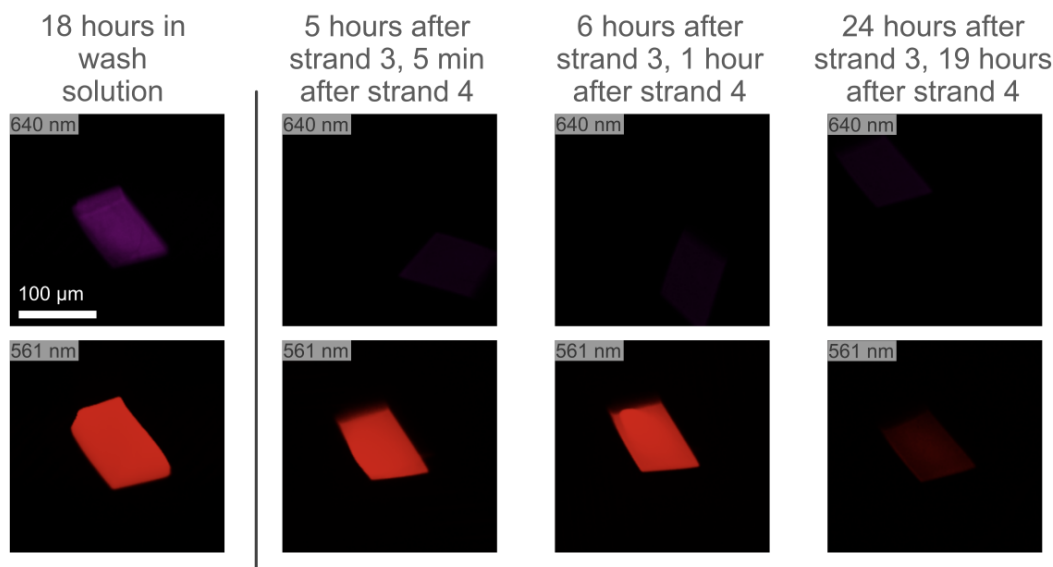


Figure 4.5: TMSD to complete cycle. After Strand 2 is loaded the crystal is moved to a new wash buffer and allowed to equilibrate for 18 hours. Strand 3 (perfect complement to Strand 1) is added at 100 μM and after 5 hours all residual Strand 1 is cleared from the crystal, while the fluorescent signal from Strand 2 is unaffected. Strand 4 (perfect complement to Strand 2) peels off Strand 2, that then diffuses out of the crystal and dissipates the fluorescent signal. After 24 hours the fluorescence emission after excitation with the 561 nm laser is nearly at background levels. We present multiple cycles with the same crystal in Fig. S4.3.

4.2.3 RNA/miR-34a capture construct

After establishing programmable strand displacement, toehold switching, and environmental capture and release of DNA cargo, we next sought to adapt CC1 into a reservoir for a functional RNA species. As discussed above, miR-34a is a well-characterized microRNA that functions as a tumor suppressor by regulating gene expression associated with cell-cycle control and apoptosis and is frequently downregulated across multiple cancer cell lines (10-15). To enable sequence-specific RNA capture, we designed a CC1 variant bearing a 15-nt ssDNA

tail complementary to one strand of miR-34a, with the goal of using toehold mediated release of double stranded RNA (Fig. 4.4).

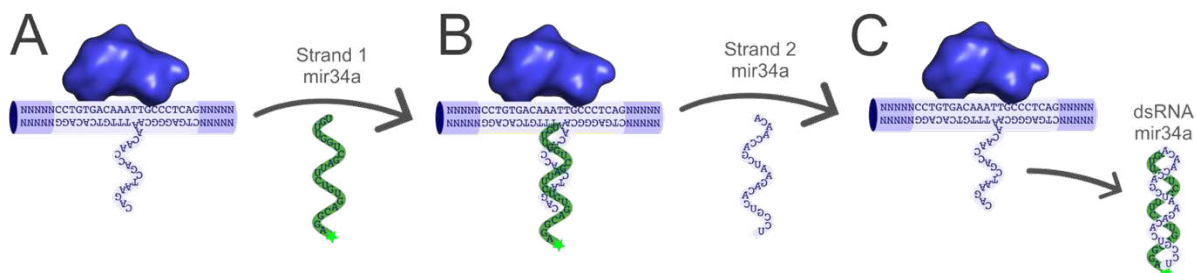


Figure 4.6: Loading and triggered release of dsRNA miR-34a. **A.** CC1 is grown with an unpaired 15-nt tail that has sequence compatibility with strand 1 of mir-34a. Strand 1 diffuses through the crystal and binds as shown in **B.** Strand 1 is 22-nt, leaving 7-nt of ssRNA. Strand 2 of mir-34a is added to solution, and as it diffuses through the crystal it binds to Strand 1. Strand 2 is a perfect complement to Strand 1, peeling it from the crystal anchor strand and releasing dsRNA into solution as shown in **C.**

To distinguish anchor-mediated binding from nonspecific accumulation within the solvent channels, we compared a miR-34a capture-enabled CC1 to a non-tailed CC1 in the presence of FAM-labeled single-stranded miR-34a (Fig. S4.4). Both crystal types initially exhibited strong fluorescence, indicating that free solution ssRNA readily diffuses into and accumulates within the porous lattice. However, upon transfer to a wash solution and equilibration with the surrounding buffer, only the tailed CC1 crystals retained fluorescence. In contrast, fluorescence associated with the non-tailed crystals decayed to background levels. These observations indicate that while electrostatic interactions can accumulate ssRNA within the crystal pores, stable retention of miR-34a requires sequence-specific hybridization to the anchor tail.

Having established specific capture of the first strand of miR-34a, we next sought to demonstrate triggered release of double-stranded miR-34a from the crystalline scaffold. Because seven nucleotides of the bound ssRNA extend beyond the anchor sequence, this design presents

an accessible toehold for TMSD release. A crystal was loaded with fluorescently labeled strand 1 of miR-34a by soaking in a 100 μ M bath (Fig. 4.5). Extended time points show an edge loading effect, indicating specific binding as the ssRNA species diffuses through the crystal pores (Fig. S4.5).

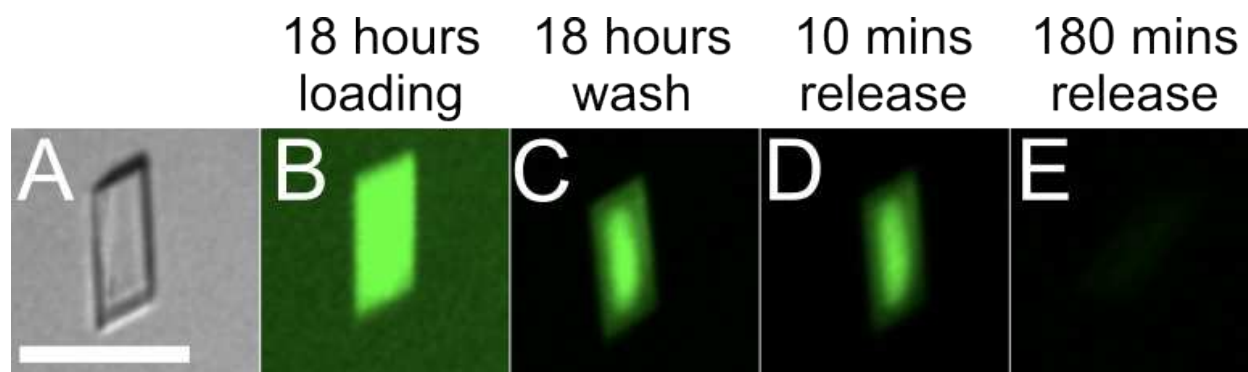


Figure 4.7: Capture and release of mir-34a. Scale bar is 100 μ m. **A.** CC1+10 containing a 15-nt mir-34a capture tail under differential interference contrast imaging. **B.** Incubation with 100 μ M strand 1 of mir-34a tagged with a fluorescein molecule. Imaging excitation with a 488 nm laser and emission with a 515 nm filter. After 18 hours the crystal appears saturated with the guest ssRNA. **C.** The crystal was looped into a fresh wash solution and left to equilibrate for 18 hours. **D.** Then the crystal was looped into a 100 μ M bath of strand 2 mir-34a. This is a ssRNA that is a perfect complement to strand 1 bound within the crystal. As strand 2 hybridizes with strand 1, strand 1 is peeled off the anchor strand of the crystal. The fluorescently tagged strand 1-strand 2 duplex then diffuses out of the crystal as it equilibrates with the surrounding solution as seen in **E.**

The crystal was then transferred to wash buffer and allowed to equilibrate for 18 hours. Upon addition of the perfectly complementary strand 2 of miR-34a to the surrounding solution, fluorescence progressively cleared from the crystal, with near-complete loss of signal observed after approximately two hours (Fig. 4.5 E). More time points showing the gradual release are presented in Fig. S4.6. In contrast, addition of a scrambled version of strand 2 to similarly loaded and equilibrated crystals produced no detectable change in fluorescence intensity up to 72 hours after the scramble strand was introduced (Fig. S4.7). This sequence dependent response

demonstrates that the release of double stranded mir-34a is triggerable and controllable, a very desirable aspect for a nanoscale device.

4.3 Structural analysis of capture-enabled co-crystals by X-ray diffraction

4.3.1 Benefits of XRD studies

The periodicity of CC1 is a defining feature that distinguishes it from other nucleic acid capture systems and enables high-precision structural characterization by X-ray diffraction. Because the 15-nt tails are incorporated uniformly throughout the lattice, this structural feature is repeated with crystallographic regularity. As a result, XRD provides a powerful means of assessing how crystals are assembling as we incorporate new, complex modalities such as the capture tails. Described here are XRD data of the DNA and RNA capture crystals with 15-nt unpaired ssDNA tails. The data presented here confirms the space group of both DNA and RNA capture device crystals to be the expected space group 5 (C 121/I 121). We report empty structures with no guest oligos loaded, as well as a DNA capture crystal loaded with Strand 1 and an RNA capture crystal loaded with strand 1 mir34a.

4.3.2 Tailed crystals without nucleic acid guests

Achieving the expected crystal packing with long ssDNA tails was a significant development of the CC1 platform. As discussed in Chapter 2, one copy of RepE54 can bind ssDNA while it is bound to cognate dsDNA, as part of its biological transcription factor function. Poly-T content in ssDNA was shown to be preferentially bound by wild type RepE by Wegrzyn, *et al*, 2023. (44). By mutating an aromatic surface-accessible amino acid (tyrosine 184) to a negatively charged amino acid (glutamic acid) we have previously shown that we can prevent

ssDNA binding with short, all cytosine tails and form CC1 crystals in space group 5.

Importantly, poly-T tails using wild type RepE54 gave new, unexpected packing arrangements of CC1 building blocks (though still crystalline assemblies), and Y184E mutant poly-T crystals did not grow. The DNA capture device anchor strand has 7 thymine out of 15 nucleotides. Despite this high thymine content, we achieved the correct crystal packing at a resolution of 3.81 Å, demonstrating that the Y184E mutation is successful at deterring aberrant packing driven by thymine-content in the anchor strand. The RNA capture device has a lower thymine content, only one nucleotide, and diffracted to a similar resolution (3.88 Å). In both crystal structures, only the beginnings of the anchor strand were able to be modelled, with two ssDNA nucleotides apparent in the DNA device and three nucleotides in the RNA device. This is an expected result, as the ssDNA tail likely has a heterogeneous population of poses, as it is loose in the solvent channel pores. Refinements of structures was done using Phenix, with final structural stats in Table S4.1 and S4.2 (46).

4.3.3 Tailed crystals with nucleic guests

Crystals of each capture modality were soaked with their nucleic acid guest at a concentration of 100 µM (for the DNA device this was strand 1 and for the RNA device this was miR-34a, as seen in Fig. 4.1B). XRD on these loaded crystals confirmed the expected space group, number 5 (I 121/C 121). The loaded DNA device crystal data sets were not suitable for refinement beyond confirming the expected crystal packing, due to poor resolution (> 6 Å). The guest loaded RNA capture crystal diffracted to a resolution of 4.5 Å, suitable to compare how well resolved the 15-nt tail was, both without and with the nucleic acid guest. The first four nucleotides of ssDNA (and no bases of the complementary guest strand) were able to be

modelled when the crystal was soaked with guest ssRNA (Fig. S4.8). This is in line with the XRD results of the Seeman and Mao original DNA device paper (5). This indicates that the anchor tails, both with and without guests loaded, are quite labile in the lattice. Future studies optimizing the loading conditions of guest strands and improving the resolution could observe more density corresponding to duplex features with higher occupancy of guest elements.

4.4 Discussion and Conclusions

4.4.1 Porous co-crystals as programmable nucleic acid reservoirs

The results presented in this chapter establish CC1 as a programmable nucleic acid reservoir crystalline material. The combination of sequence-encoded capture, environmental stability, and triggerable release enables controlled handling of nucleic acid cargo within a rigid, porous crystalline framework. Because capture elements are organized throughout the lattice, CC1 supports high-density, spatially uniform loading post-crystallization. The crystallographic order and predictable packing arrangement of CC1+10 extends to these tailed variants and is an example of the level of organization that crystalline materials provide.

A key feature of this system is its modularity. The DNA component of the building block provides a straightforward route to sequence-programmable installation of capture handles without altering the underlying lattice architecture. This design principle permits tuning of cargo identity and release mechanism. As demonstrated here for both DNA and RNA guests, nucleic acid cargo can be captured from solution, retained under wash conditions, and selectively released via sequence-specific strand displacement, establishing CC1 as a reusable and reconfigurable storage platform.

4.4.2 miR-34a capture and release as a step toward functional cargo

The crosslinking that enables CC1+10 to retain its structure and organization outside of its growth conditions innovates on the operation of nucleic acid-based devices previously demonstrated in porous DNA-based crystals (5). We have previously demonstrated that ligation of the CC1 platform is tunable, a desirable trait for a drug delivery vehicle (1, 2). Here we have shown how the stable lattice of CC1 is able to bind ssRNA from solution, protect it through wash steps, and then release dsRNA on a trigger. This behavior is conceptually related to other solid or particulate nucleic acid reservoirs, including lipid nanoparticles, polymeric nanoparticles, hydrogels, and inorganic nanocarriers, that can encapsulate or complex nucleic acids to improve stability, retention, and triggered or sustained release (47-50). However, CC1 differs from many of these platforms by combining nucleic acid storage with an ordered, sequence-programmable crystalline lattice that can be adapted through the encoded capture strand. These properties position protein–DNA co-crystals as a promising class of crystalline biomaterials for future nucleic-acid-handling and delivery applications.

4.4.5 Limits and unresolved questions

Structural studies and confocal microscopy can suggest that guest molecules are loading into CC1 with high occupancy, but to directly quantify the loading and release efficiency more kinetic studies are required. We are developing a custom cuvette that sequesters a crystal from the light path of a fluorescence reader, while keeping the crystal in solution contact (Fig. S4.9). This allows us to take time course measurements of the change in fluorescence of a solution containing a crystal. Finding the equilibrium concentrations at different starting concentrations

and how quickly these equilibria are reached will allow us to quantify CC1's capacity for loading and release for fluorescently tagged guest molecules.

In this work we show that ligation affords the ability to move crystals from their growth solutions, that contain building blocks at an equilibrium that sustains the crystal's formation, and move them through loading and wash solutions. Further experimentation in varied solutions will be necessary to realize function for this platform in diverse applications.

4.6 Future directions

The ability to use porous co-crystals as nucleic acid reservoirs has implications in nanotechnology where the precise spatial arrangement and sequence specificity are desirable for nanometer scale installation of guest molecules (51-53). As demonstrated in our previous report, scaffold assisted structural determination is possible using the CC1+10 platform, and observing the structure of RNA molecules held steady by anchor strands in CC1 would be of interest to structural biologists (4). Expanding the library of solutions where nucleic acid guests can load, be protected, and released will increase the likelihood that this technology can be adapted to biomedical applications, as well as cytocompatibility studies of CC1 in nanoparticle size.

4.7 Materials and Methods

4.7.1 Protein cloning, expression, purification

Cloning and expression of the protein Replication initiator RepE54 was previously described by Komori *et al.* (54). In hopes of disrupting any favorable intermolecular contacts at the ssDNA binding cleft identified in pdb 8AAN we added a negatively charged side chain. Variant RepE54 (Y184E) was cloned into the pETDuet plasmid (Novagen, Sigma-Aldrich # 71146-3) using Gibson cloning (Hi-Fi assembly New England Biolabs #E2621L). The protein was overexpressed with a T7 promoter in *E. coli* BL21(DE3) competent cells (New England Biolabs #C2527H). Upon addition of 0.5 mM IPTG, the cells were incubated at 25 °C for 20 h. The cell pellets were sonicated in lysis buffer (1X PBS, 300 mM NaCl, 25 mM imidazole, pH 7.4) and applied to HisTrap (HisPur™ Ni-NTA resin, ThermoFisher #88222) equilibrated with HisTrap buffer (1X PBS, 300 mM NaCl, 25 mM imidazole, pH 7.4). The protein was eluted with 100 mM imidazole in HisTrap buffer. Fractions containing the desired protein were pooled and further purified at the CSU Histone Source facility with HiLoad Superdex 200 PG column (Cytiva #28989335). RepE54 variant (Y184E) was concentrated to 15 mg/mL in CC1 storage buffer (100 mM sodium citrate pH 6.2, 100 mM KCl, 10 mM MgCa₂, and 10% glycerol), flash frozen in liquid nitrogen, and stored at −80°C.

4.7.2 Duplex annealing and co-crystal complexing

Individual oligomers were synthesized and HPLC purified by Integrated DNA Technologies. The DNA oligomer sequences used for co-crystallization are listed in Tables S3. Each oligomer contained the 19-bp iteron sequence TGTGACAAATTGCCCTCAG, but typically also contained the core 21-bp DNA block used in PDB:1REP from Komori *et al.* (54):

CCTGTGACAAATTGCCCTCAG (omitting a dangling 3' T). In Table S3, the standard 21-bp footprint is shaded in gray, and (purposeful) deviations are shaded in cyan. Flanking variable DNA sequences depended on the expansion scheme and desired sequence. The oligomers were resuspended in a standardized buffer (henceforth CC1 oligo buffer: 50 mM Tris HCl, 100 mM KCl, pH 6.7) and combined in equal molar ratio (1:1) with the complementary strand. The strands were annealed by heating to 94 °C for 2 min and slowly cooling to room temperature over approximately 60 minutes. The final concentration of all CC1 dsDNA blocks following duplexing was 4 mM. The scaffold protein and DNA were mixed in a 1.2:1 ratio (440:364 μM, 55 μL total) and were incubated on ice for 30 minutes prior to crystallization via sitting drop vapor diffusion.

4.7.3 Crystallization

Crystallization trials used CrysChem sitting-drop plates (Hampton Research HR3-159), with 400 μL reservoir solution per well and 4 μL crystallization drops prepared by mixing 2 μL each of the protein–DNA complex and reservoir solution directly on the platform; drops were mixed by gentle pipetting. Plates were sealed immediately after setup and incubated at 25 °C under static conditions. Crystallization conditions for the porous lattices were 10-80 mM Magnesium Acetate, 0.3-1.8 M Lithium Sulfate, and 50 mM MES pH 6.5 (or occasionally 50 mM sodium cacodylate pH 6.5). A detailed list of individual crystals, PDB codes, and corresponding growth conditions are reported in Table S2. Crystals grew to full size in a range of 24 hours to 30 days (most often in 1-2 days). Crystal images were taken with a Moticam X5 Plus camera attached to a Motic SMZ-168 stereozoom microscope.

4.7.4 Crosslinking

Co-crystals were washed in conditions similar to their growth conditions (0.3-2.0 M Lithium sulfate, 20-50 mM Magnesium acetate, and 50 mM MES, pH 6.5). The co-crystals were washed in 200 μ L on 9-well glass plates (Hampton) to remove additional protein and DNA monomers. Note: For EDC ligation it is important to avoid buffers with 1°/2° amines or carboxylic acids. 1-Ethyl-3(3-dimethylaminopropyl)carbodiimide (EDC) (Advanced Chemtech CAS#:25952-53-8) was resuspended in the crosslinking solution (1.0 M Lithium sulfate, 50 mM Magnesium acetate, and 50 mM MES, pH 6.0) to final concentration values of 30-60 mg/mL and used immediately. The co-crystals were crosslinked in a 200 μ L EDC solution volume overnight (16-24 hrs). The crosslinking reaction was ended by looping the co-crystals back into the wash solution.

4.7.5 Confocal Microscopy

Crystals loaded with fluorescently labeled nucleic acids were imaged using a Nikon Eclipse Ti Yokagawa CSU-X1 spinning-disk confocal microscope equipped with an Andor iXon Ultra 879U EMCCD camera. Fluorescein signal was collected using 488 nm laser excitation and a 515 nm emission filter. Cy5 signal was collected using 640 nm laser excitation and a 660 nm emission filter. Cy3 signal was collected using 561 nm laser excitation and 570 nm emission filter. Image acquisition and processing were performed using Nikon NIS-Elements software.

4.7.6 Guest Loading

Crosslinked co-crystals containing DNA sequences for guest nucleic acid capture were washed at room temperature to remove components that could interfere with guest nucleic acid binding and to equilibrate the crystals to the guest loading soak solution (1.2 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, 50 mM MES pH 7.1). Equilibrated crystals were moved to 9 μ L fresh buffer solution in a micro-Insert 4 Well in μ -Dish (Ibidi Cat.No:80406). 1 μ L of 1 mM guest nucleic acid was added on top

of the well, for a final guest concentration of 100 μ M. The co-crystals sitting in guest loading solutions were sealed for 24-72 hours to allow guest species to diffuse and reach binding equilibrium within the crystal lattice.

4.7.7 X-ray Diffraction collection and refinement

Crystals were flash frozen in liquid nitrogen after briefly transferring the crystal to a cryo-protectant solution (a 1:1 mixture of the reservoir solution and 50% glycerol stock for a final concentration of 25% glycerol). Single crystal X-Ray diffraction data was collected at APS beamline 24-IDE (Dectris EIGER2 detector) from 0 to 180° with an omega delta of 0.2° and an exposure of 0.5 s or 0.25 s. Data was processed with XDS (55) and refined using PHENIX and COOT (46, 56). The first porous CC1+10 dataset (PDB 9Z08) was refined using newly generated RFree flags. All subsequent isomorphous structures were refined using the same RFree flags to preserve a fixed cross-validation benchmark across datasets. Decisions about truncating the structure factor data were made on a case-by-case basis, jointly considering the following thresholds $CC1/2 > 0.3$, mean $\langle I/\sigma(I) \rangle > 1.0$, and completeness $> 85\%$ as target guidelines for the highest-resolution shell. Most datasets satisfy all three thresholds in the highest-resolution shell.

CHAPTER FOUR BIBLIOGRAPHY

1. Orun, A. R., Dmytriw, S., Vajapayajula, A., & Snow, C. D. (2022). Stabilizing DNA–Protein Co-Crystals via Intra-Crystal Chemical Ligation of the DNA. *Crystals*, 12(1), Article 1. <https://doi.org/10.3390/cryst12010049>
2. Orun, A. R., Slaughter, C. K., Shields, E. T., Vajapayajula, A., Jones, S., Shrestha, R., & Snow, C. D. (2024). Tuning Chemical DNA Ligation within DNA Crystals and Protein–DNA Cocystals. *ACS Nanoscience Au*. <https://doi.org/10.1021/acsnanoscienceau.4c00013>
3. Orun, A. R., Shields, E. T., Dmytriw, S., Vajapayajula, A., Slaughter, C. K., & Snow, C. D. (2023). Modular Protein–DNA Cocystals as Precise, Programmable Assembly Scaffolds. *ACS Nano*, 17(14), 13110–13120. <https://doi.org/10.1021/acsnano.2c07282>
4. Shields, E. T., Slaughter, C. K., Mekkaoui, F., Magna, E. N., Shepherd, C., Lukeman, P. S., Spratt, D. E., & Snow, C. D. (2026). Modular Scaffold Crystals for Programmable Installation and Structural Observation of DNA-Binding Proteins (p. 2026.03.04.709581). *bioRxiv*. <https://doi.org/10.64898/2026.03.04.709581>
5. Hao, Y., Kristiansen, M., Sha, R., Birktoft, J. J., Hernandez, C., Mao, C., & Seeman, N. C. (2017). A device that operates within a self-assembled 3D DNA crystal. *Nature Chemistry*, 9(8), 824–827. <https://doi.org/10.1038/nchem.2745>
6. Woloszyn, K., Vecchioni, S., Ohayon, Y. P., Lu, B., Ma, Y., Huang, Q., Zhu, E., Chernovolenco, D., Markus, T., Jonoska, N., Mao, C., Seeman, N. C., & Sha, R. (2022). Augmented DNA Nanoarchitectures: A Structural Library of 3D Self-Assembling Tensegrity Triangle Variants. *Advanced Materials*, 34(49), 2206876. <https://doi.org/10.1002/adma.202206876>
7. Mao, C., Birktoft, J. J., Chen, Y., Seeman, N. C., Constantinou, P. E., Wang, T., Sha, R., Ginell, S. L., & Zheng, J. (2009). From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature (London)*, 461(7260), 74–77. <https://doi.org/10.1038/nature08274>
8. Chen, X. (2012). Expanding the rule set of DNA circuitry with associative toehold activation. *Journal of the American Chemical Society*, 134(1), 263–271. <https://doi.org/10.1021/ja206690a>
9. Zhang, D. Y., & Winfree, E. (2009). Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *Journal of the American Chemical Society*, 131(47), 17303–17314. <https://doi.org/10.1021/ja906987s>
10. Tarasov, V., Jung, P., Verdoodt, B., Lodygin, D., Epanchintsev, A., Menssen, A., Meister, G., & Hermeking, H. (2007). Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. *Cell Cycle*, 6(13), 1586–1593. <https://doi.org/10.4161/cc.6.13.4436>
11. Chang, T.-C., Wentzel, E. A., Kent, O. A., Ramachandran, K., Mullendore, M., Lee, K. H., Feldmann, G., Yamakuchi, M., Ferlito, M., Lowenstein, C. J., Arking, D. E., Beer, M. A., Maitra, A., & Mendell, J. T. (2007). Transactivation of miR-34a by p53 broadly

- influences gene expression and promotes apoptosis. *Molecular Cell*, 26(5), 745–752. <https://doi.org/10.1016/j.molcel.2007.05.010>
12. He, L., He, X., Lowe, S. W., & Hannon, G. J. (2007). microRNAs join the p53 network—Another piece in the tumour-suppression puzzle. *Nature Reviews. Cancer*, 7(11), 819–822. <https://doi.org/10.1038/nrc2232>
 13. Hermeking, H. (2010). The miR-34 family in cancer and apoptosis. *Cell Death and Differentiation*, 17(2), 193–199. <https://doi.org/10.1038/cdd.2009.56>
 14. Maroof, H., Salajegheh, A., Smith, R. A., & Lam, A. K.-Y. (2014). Role of microRNA-34 family in cancer with particular reference to cancer angiogenesis. *Experimental and Molecular Pathology*, 97(2), 298–304. <https://doi.org/10.1016/j.yexmp.2014.08.002>
 15. Zhang, L., Liao, Y., & Tang, L. (2019). MicroRNA-34 family: A potential tumor suppressor and therapeutic candidate in cancer. *Journal of Experimental & Clinical Cancer Research*, 38(1), 53. <https://doi.org/10.1186/s13046-019-1059-5>
 16. Jiang, L., & Hermeking, H. (2017). miR-34a and miR-34b/c Suppress Intestinal Tumorigenesis. *Cancer Research*, 77(10), 2746–2758. <https://doi.org/10.1158/0008-5472.CAN-16-2183>
 17. Sun, H., Tian, J., Xian, W., Xie, T., & Yang, X. (2015). miR-34a inhibits proliferation and invasion of bladder cancer cells by targeting orphan nuclear receptor HNF4G. *Disease Markers*, 2015, 879254. <https://doi.org/10.1155/2015/879254>
 18. Siemens, H., Jackstadt, R., Hüntten, S., Kaller, M., Menssen, A., Götz, U., & Hermeking, H. (2011). miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell Cycle (Georgetown, Tex.)*, 10(24), 4256–4271. <https://doi.org/10.4161/cc.10.24.18552>
 19. Beg, M. S., Brenner, A. J., Sachdev, J., Borad, M., Kang, Y.-K., Stoudemire, J., Smith, S., Bader, A. G., Kim, S., & Hong, D. S. (2017). Phase I study of MRX34, a liposomal miR-34a mimic, administered twice weekly in patients with advanced solid tumors. *Investigational New Drugs*, 35(2), 180–188. <https://doi.org/10.1007/s10637-016-0407-y>
 20. Hong, D. S., Kang, Y.-K., Borad, M., Sachdev, J., Ejadi, S., Lim, H. Y., Brenner, A. J., Park, K., Lee, J.-L., Kim, T.-Y., Shin, S., Becerra, C. R., Falchook, G., Stoudemire, J., Martin, D., Kelnar, K., Peltier, H., Bonato, V., Bader, A. G., ... Beg, M. S. (2020). Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. *British Journal of Cancer*, 122(11), 1630–1637. <https://doi.org/10.1038/s41416-020-0802-1>
 21. Mirna Therapeutics, Inc. (with Cancer Prevention Research Institute of Texas). (2016). A Multicenter Phase I Study of MRX34, MicroRNA miR-RX34 Liposomal Injection (Clinical Trial Registration NCT01829971). [clinicaltrials.gov](https://clinicaltrials.gov/study/NCT01829971). <https://clinicaltrials.gov/study/NCT01829971>
 22. Segal, M., & Slack, F. J. (2020). Challenges identifying efficacious miRNA therapeutics for cancer. *Expert Opinion on Drug Discovery*, 15(9), 987–992. <https://doi.org/10.1080/17460441.2020.1765770>

23. Tian, H., Cheng, L., Liang, Y., Lei, H., Qin, M., Li, X., & Ren, Y. (2024). MicroRNA therapeutic delivery strategies: A review. *Journal of Drug Delivery Science and Technology*, 93, 105430. <https://doi.org/10.1016/j.jddst.2024.105430>
24. Seyhan, A. A. (2024). Trials and Tribulations of MicroRNA Therapeutics. *International Journal of Molecular Sciences*, 25(3). <https://doi.org/10.3390/ijms25031469>
25. Diener, C., Keller, A., & Meese, E. (2022). Emerging concepts of miRNA therapeutics: From cells to clinic. *Trends in Genetics: TIG*, 38(6), 613–626. <https://doi.org/10.1016/j.tig.2022.02.006>
26. Iacomino, G. (2023). miRNAs: The Road from Bench to Bedside. *Genes*, 14(2), Article 2. <https://doi.org/10.3390/genes14020314>
27. Stenvang, J., Petri, A., Lindow, M., Obad, S., & Kauppinen, S. (2012). Inhibition of microRNA function by anti-miR oligonucleotides. *Silence*, 3, 1. <https://doi.org/10.1186/1758-907X-3-1>
28. Lima, J. F., Cerqueira, L., Figueiredo, C., Oliveira, C., & Azevedo, N. F. (2018). Anti-miRNA oligonucleotides: A comprehensive guide for design. *RNA Biology*, 15(3), 338–352. <https://doi.org/10.1080/15476286.2018.1445959>
29. O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Frontiers in Endocrinology*, 9. <https://www.frontiersin.org/articles/10.3389/fendo.2018.00402>
30. Dexheimer, P. J., & Cochella, L. (2020). MicroRNAs: From Mechanism to Organism. *Frontiers in Cell and Developmental Biology*, 8. <https://www.frontiersin.org/articles/10.3389/fcell.2020.00409>
31. Wang, P., Zhou, Y., & Richards, A. M. (2021). Effective tools for RNA-derived therapeutics: siRNA interference or miRNA mimicry. *Theranostics*, 11(18), 8771–8796. <https://doi.org/10.7150/thno.62642>
32. Li, Y., Duo, Y., Bi, J., Zeng, X., Mei, L., Bao, S., He, L., Shan, A., Zhang, Y., & Yu, X. (2018). Targeted delivery of anti-miR-155 by functionalized mesoporous silica nanoparticles for colorectal cancer therapy. *International Journal of Nanomedicine*, 13, 1241–1256. <https://doi.org/10.2147/IJN.S158290>
33. Medley, J. C., Panzade, G., & Zinovyeva, A. Y. (2021). microRNA strand selection: Unwinding the rules. *Wiley Interdisciplinary Reviews. RNA*, 12(3), e1627. <https://doi.org/10.1002/wrna.1627>
34. Wu, Y., Jin, M., Peng, C., Wang, G. A., & Li, F. (2025). Kinetics and Activation Strategies in Toehold-Mediated and Toehold-Free DNA Strand Displacement. *Biosensors*, 15(10), 683. <https://doi.org/10.3390/bios15100683>
35. G. Smith, F., P. Goertz, J., Jurinović, K., M. Stevens, M., & E. Ouldrige, T. (2024). Strong sequence-dependence in RNA/DNA hybrid strand displacement kinetics. *Nanoscale*, 16(37), 17624–17637. <https://doi.org/10.1039/D4NR00542B>

36. Liu, H., Hong, F., Smith, F., Goertz, J., Ouldrige, T., Stevens, M. M., Yan, H., & Šulc, P. (2021). Kinetics of RNA and RNA:DNA Hybrid Strand Displacement. *ACS Synthetic Biology*, 10(11), 3066–3073. <https://doi.org/10.1021/acssynbio.1c00336>
37. Ratajczyk, E. J., Bath, J., Šulc, P., Doye, J. P. K., Louis, A. A., & Turberfield, A. J. (2025). Controlling DNA–RNA strand displacement kinetics with base distribution. *Proceedings of the National Academy of Sciences*, 122(23), e2416988122. <https://doi.org/10.1073/pnas.2416988122>
38. Srinivas, N., Ouldrige, T. E., Šulc, P., Schaeffer, J. M., Yurke, B., Louis, A. A., Doye, J. P. K., & Winfree, E. (2013). On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Research*, 41(22), 10641–10658. <https://doi.org/10.1093/nar/gkt801>
39. Šulc, P., Ouldrige, T. E., Romano, F., Doye, J. P. K., & Louis, A. A. (2015). Modelling Toehold-Mediated RNA Strand Displacement. *Biophysical Journal*, 108(5), 1238–1247. <https://doi.org/10.1016/j.bpj.2015.01.023>
40. Zhang, D. Y., & Seelig, G. (2011). Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chemistry*, 3(2), 103–113. <https://doi.org/10.1038/nchem.957>
41. Jurinović, K., Mitra, M., Mukherjee, R., & Ouldrige, T. E. (2026). Design of DNA strand displacement reactions. *Current Opinion in Biotechnology*, 97, 103396. <https://doi.org/10.1016/j.copbio.2025.103396>
42. Wu, Z., Xie, X., Li, P., Zhao, J., Huang, L., & Zhou, X. (2013). Equilibrium Strand Exchange Promoted by DNA Conformational Switching. *Scientific Reports*, 3, 1121. <https://doi.org/10.1038/srep01121>
43. Kornienko, I. V., Aramova, O. Yu., Tishchenko, A. A., Rudoy, D. V., & Chikindas, M. L. (2024). RNA Stability: A Review of the Role of Structural Features and Environmental Conditions. *Molecules*, 29(24), 5978. <https://doi.org/10.3390/molecules29245978>
44. Wegrzyn, K., Oliwa, M., Nowacka, M., Zabrocka, E., Bury, K., Purzycki, P., Czaplowska, P., Pipka, J., Giraldo, R., & Konieczny, I. (2023). Rep protein accommodates together dsDNA and ssDNA which enables a loop-back mechanism to plasmid DNA replication initiation. *Nucleic Acids Research*, 51(19), 10551–10567. <https://doi.org/10.1093/nar/gkad740>
45. Wang, T., Sha, R., Birktoft, J., Zheng, J., Mao, C., & Seeman, N. C. (2010). A DNA Crystal Designed to Contain Two Molecules per Asymmetric Unit. *Journal of the American Chemical Society*, 132(44), 15471–15473. <https://doi.org/10.1021/ja104833t>
46. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., ... Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica*

Section D: Structural Biology, 75(10), 861–877.

<https://doi.org/10.1107/S2059798319011471>

47. Mendes, B. B., Conniot, J., Avital, A., Yao, D., Jiang, X., Zhou, X., Sharf-Pauker, N., Xiao, Y., Adir, O., Liang, H., Shi, J., Schroeder, A., & Conde, J. (2022). Nanodelivery of nucleic acids. *Nature Reviews Methods Primers*, 2(1), 24.
<https://doi.org/10.1038/s43586-022-00104-y>
48. Piotrowski-Daspiet, A. S., Kauffman, A. C., Bracaglia, L. G., & Saltzman, W. M. (2020). Polymeric Vehicles for Nucleic Acid Delivery. *Advanced Drug Delivery Reviews*, 156, 119–132. <https://doi.org/10.1016/j.addr.2020.06.014>
49. El Moukhtari, S. H., Garbayo, E., Amundarain, A., Pascual-Gil, S., Carrasco-León, A., Prosper, F., Agirre, X., & Blanco-Prieto, M. J. (2023). Lipid nanoparticles for siRNA delivery in cancer treatment. *Journal of Controlled Release*, 361, 130–146.
<https://doi.org/10.1016/j.jconrel.2023.07.054>
50. Sarett, S. M., Nelson, C. E., & Duvall, C. L. (2015). Technologies for Controlled, Local Delivery of siRNA. *Journal of Controlled Release : Official Journal of the Controlled Release Society*, 218, 94–113. <https://doi.org/10.1016/j.jconrel.2015.09.066>
51. Zhang, F., Jiang, S., Wu, S. *et al.* Complex wireframe DNA origami nanostructures with multi-arm junction vertices. *Nature Nanotech* 10, 779–784 (2015). <https://doi-org.ezproxy2.library.colostate.edu/10.1038/nnano.2015.162>
52. Teller, C., & Willner, I. (2010). Organizing protein–DNA hybrids as nanostructures with programmed functionalities. *Trends in Biotechnology*, 28(12), 619–628.
<https://doi.org/10.1016/j.tibtech.2010.09.005>
53. Seeman, N. C. (2003). DNA in a material world. *Nature*, 421(6921), 427–431.
<https://doi.org/10.1038/nature01406>
54. Komori, H., Sasai, N., Matsunaga, F., Wada, C., & Miki, K. (1999). Crystallization and Preliminary X-Ray Diffraction Studies of a Replication Initiator Protein (RepE54) of the Mini-F Plasmid Complexed with Iteron DNA. *Journal of Biochemistry*, 125(1), 24–26.
<https://doi.org/10.1093/oxfordjournals.jbchem.a022262>
55. Kabsch, W. (2010). XDS. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 125–132. <https://doi.org/10.1107/S09074444909047337>
56. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, 66(4), 486–501.
<https://doi.org/10.1107/S09074444910007493>

CHAPTER 5: BATCH MICROCRYSTALLIZATION

5.1 Introduction

5.1.1 Motivation for switching to microcrystallization

Vapor-diffusion crystallization exploits differences in precipitant concentration between the crystallization drop and the reservoir, creating salt and osmotic gradients that gradually dehydrate the drop and drive nucleation (1, 2). In this format, the timing and extent of nucleation are governed by the rate at which these gradients develop, as well as by the identities and concentrations of precipitants and building blocks, in our case, protein-DNA complexes.

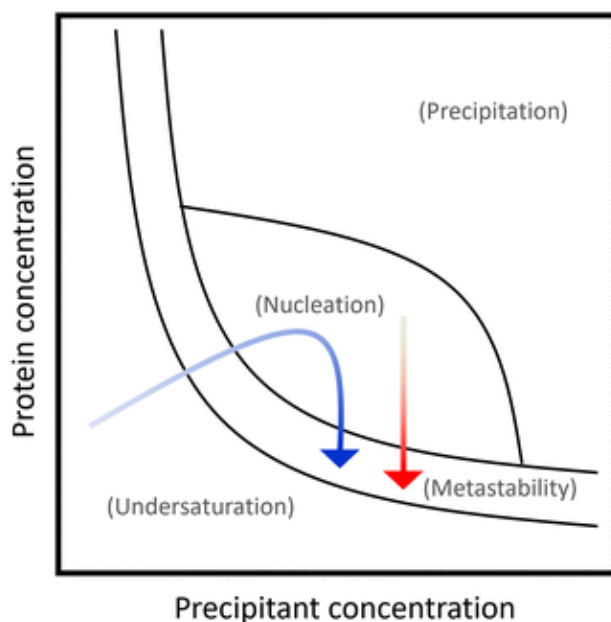


Figure 5.1: Phase diagram. The trajectory, over time, of protein and precipitant concentration for vapour diffusion (blue arrow) and batch crystallization (red arrow). From “Porous protein crystals: synthesis and applications” A. A. Jones and C. D. Snow, Chem. Commun., 2024, 60, 5790 DOI: 10.1039/D4CC00183D. CC BY (1)

As a result, nucleation is typically delayed until suitable supersaturation conditions are reached, after which a limited number of nuclei form and grow over time. When well optimized, vapor diffusion sitting drop experiments can yield large, well-ordered crystals that are ideally suited for X-ray diffraction. For the CC1 platform in particular, crystal growth by vapor diffusion

can be comparatively rapid, with crystals often appearing within days rather than weeks. As crystal growth proceeds in this format, however, the number of discrete crystals per well generally decreases as growth becomes increasingly focused on a small number of dominant nuclei (1-4).

In contrast, microcrystallization seeks to initiate nucleation immediately by combining protein, nucleic acid, and precipitant at concentrations designed to place the system directly within the nucleation zone of the phase diagram (Fig. 5.1) (2). Rather than relying on time-dependent dehydration to reach supersaturation, batch microcrystallization begins with conditions that favor rapid and simultaneous nucleation. This shift results in the formation of many crystals per well, that are significantly smaller than those produced by vapor diffusion. By minimizing time for growth, reducing crystal size, and increasing the numbers of crystals formed, microcrystallization provides an ideal method for batch style crystal growth (5, 6).

These differences in growth behavior make microcrystallization particularly attractive for applications that require bulk production of crystalline biomaterial rather than isolated, diffraction-quality crystals. Accordingly, transitioning from vapor-diffusion crystal growth to batch microcrystallization represents a necessary step in adapting protein-DNA co-crystals to downstream applications. However, producing small crystals alone is not sufficient for materials-oriented applications. The width of the particle size distribution is also a critical design parameter because broad or poorly controlled distributions can lead to inconsistent handling, filtration, sedimentation, diffusion, dosing, and stability outcomes. Therefore, an important goal of batch microcrystallization is not only to reduce the average crystal size, but also to narrow the distribution around an application-relevant target size range. In this work, batch uniformity is

evaluated primarily by microscopy and DLS, with future optimization aimed at reducing polydispersity through improved control of nucleation, growth time, washing, and filtration.

5.1.2 Role of chemical crosslinking in microcrystal stabilization

As established in earlier chapters, CC1 crystals as grown are vulnerable to dissolution or structural degradation when removed from their mother liquor. This intrinsic fragility poses a significant limitation for any application that requires crystal handling, buffer exchange, or exposure to environments that differ from the original crystallization conditions. Chemical crosslinking provides a practical strategy for reinforcing the crystal lattice and preserving crystal integrity under such challenges.

As shown in chapter 3, Carbodiimide-mediated crosslinking using EDC is particularly well suited for CC1 crystals, as the reaction can be carried out in solutions that closely mimic the original mother liquor, including comparably high lithium sulfate and magnesium acetate concentrations and similar pH conditions. Previous work from our laboratory demonstrated that EDC-crosslinked ipCC1 crystals grown by vapor diffusion, typically large crystals on the order of $\sim 100\ \mu\text{m}$ in length, survive exposure to biomimetic environments such as lysosomal mimic solutions, stomach acid mimic, and bovine blood serum (7). In those studies, crosslinking reactions were performed in optimized buffer systems tailored to the ipCC1 crystallization conditions and yielded high ligation efficiencies (8). In Chapter 3 of this work, we examined how moving from ipCC1 crosslinking protocols to CC1 crosslinking effects ligation yields.

In the present work, this crosslinking strategy is extended to batches of CC1 microcrystals produced by microbatch under-oil crystallization. Compared to larger vapor-diffusion crystals, microcrystals present additional stabilization challenges due to their higher surface-area-to-volume ratios and increased susceptibility to dissolution during handling

and solvent exchange. Applying EDC crosslinking using the established CC1 protocols enables covalent reinforcement of these microcrystals without requiring extensive modification of the growth or treatment conditions. Here we show that EDC-crosslinked CC1 microcrystals retain their structural integrity when transferred into deionized water, bovine blood serum, and low-pH environments that rapidly destabilize uncrosslinked controls. As such, crosslinking represents a critical enabling step in adapting CC1 microcrystals for downstream stability testing and materials-oriented applications (9).

5.2 Development of crystal growth and crosslinking conditions

5.2.1 Batch microcrystallization conditions

Initial crystallization of CC1 was established using sitting-drop vapor diffusion, with pedestal drops composed of a 1:1 mixture of protein-DNA complex and reservoir solution. The reservoir condition typically consisted of 1-1.8 M Li_2SO_4 , 20-50 mM magnesium acetate, and 50 mM MES at pH 6.5. Under these conditions, crystal growth was driven by gradual equilibration of the drop against the reservoir, yielding large, well-formed co-crystals suitable for structural characterization.

To generate smaller crystals in a reproducible batch format, this condition was subsequently utilized for microbatch crystallization under oil. In contrast to vapor diffusion, microbatch crystallization establishes the final crystallization environment immediately within the drop, allowing more direct control over the concentrations experienced during nucleation and early growth. Following standard approaches for converting vapor-diffusion conditions to batch crystallization, the salt concentration was screened beginning at reduced strength relative to the original reservoir while preserving the same overall chemical components. Protein-DNA complex concentration was also varied to modulate nucleation density. Under-oil microbatch

drops reproducibly yielded dense populations of micron-scale co-crystals, with substantially higher particle numbers per drop than observed in vapor-diffusion experiments (Fig. 5.2).

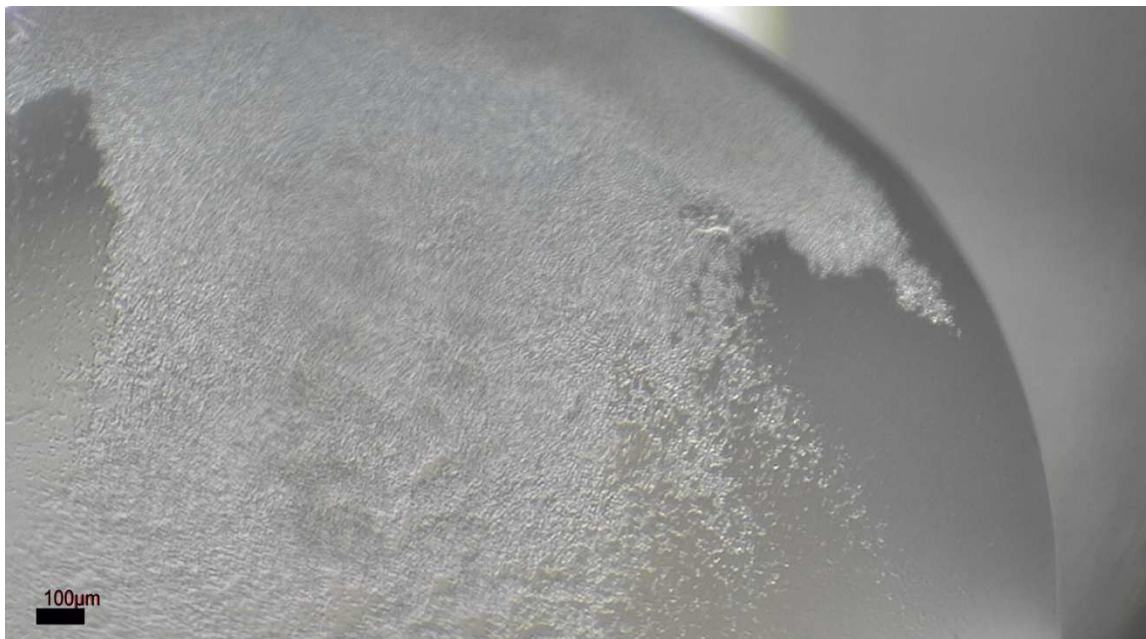


Figure 5.2: A field of CC1 microcrystals generated by microbatch under oil crystallization. The scale bar is 100 μm . The sample contains a dense population of optically clear, faceted particles with sharp edges and uniform crystalline morphology, consistent with microcrystals rather than amorphous protein precipitate. No diffuse, granular, or cloud-like material characteristic of nonspecific precipitation is apparent in the surrounding drop.

The under-oil microbatch format provided several practical advantages for this system. Because drops are under oil and therefore cannot evaporate at an appreciable rate, final concentrations are highly reproducible, and batch-to-batch variability is reduced. The format also facilitates the generation of bulk quantities of microcrystalline material within a single plate, maximizing building blocks to usable crystal yields. Importantly, the resulting microcrystals retained the characteristic morphology observed for CC1 crystals grown by vapor diffusion, suggesting that the crystal identity was preserved despite the change in growth format.

Based on these observations, a set of working microbatch conditions was selected that consistently produced high densities of microcrystals with minimal amorphous precipitation. At

this stage, condition selection was based primarily on visual assessment of crystal density, morphology, and gross size heterogeneity. However, for future biomaterials applications, the target is not simply high microcrystal yield, but a narrow and reproducible particle size distribution. Future optimization should therefore quantify distribution width directly, for example by reporting DLS polydispersity index, coefficient of variation, or distribution span across independently prepared batches.

Starting with a protein-DNA complex at 9.4 mg/mL after complexing RepE54 and DNA variant at a 1:1.2 molar ratio we screened final drop concentrations from 3.5-5.5 mg/mL, with microcrystals observed at all these concentrations and no obvious trend in size/quality dependent on ending complex concentration. A salt stock was made for screening precipitant concentrations while maintaining conditions comparable to vapor diffusion. 100% salt stock consisted of 1.3 M Li₂SO₄, 65 mM MgAc, and 50 mM MES pH 6.5, and final drop salt concentrations were screened from 20-45%, with water used to dilute. Similarly to the complex concentration gradient, all the salt percentages screened were fruitful and no obvious correlation was observed between size/quality and final salt concentration. Final well volumes ranged from 2-4 μ L, with smaller drop sizes nucleating more quickly. All drops were covered with 14 μ L of heavy oil, and microcrystals were observed on the order of hours to days.

5.2.2 Chemical crosslinking of batch-produced microcrystals

As established in our previous report, chemical crosslinking with EDC is sufficient to allow CC1 to persist in solutions outside of mother liquor (7-9). In this work we have shown that EDC ligation yields are tunable dependent on the number of sticky base overhangs and phosphate location (3' vs 5'). This control over dissolution timing is a desirable characteristic for a biomaterial. Here we report a protocol for crosslinking microbatches of CC1 with EDC.

EDC crosslinking buffer consisted of 1 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, and 50 mM MES pH 6.0. This buffer is comparable to mother liquor growth conditions, with a lower pH to assist in the reactivity of the phosphate group. The buffer was supplemented with EDC dissolved to a final concentration of 30 mg/mL, and 200 μL of this was added to a well on a separate glass plate. Crystals were pipetted out of the well (slowly - to prevent mechanical damage to the microcrystals), avoiding the oil drop, and transferred to the crosslinking bath. The microbatch was left in the crosslinking bath for 12 hours. After 12 hours, some protein precipitation was observed above the settled microcrystal batch, but the opaque cloud remained visually distinguishable from the crystals. The microcrystals settled to the bottom of the well, while the precipitation floated above. The crosslinking bath was pipetted off, removing the precipitated protein but leaving the microcrystals in a minimal amount of remaining buffer. A new solution of 30 mg/mL EDC crosslinking bath was added on top and left for 12 hours. After two doses of EDC, crystals were ready for stability testing and batch characterization.

5.3. Batch characterization and stability testing

5.3.1 Batch characterization

Dynamic light scattering (DLS) measures temporal fluctuations in scattered light from Brownian motion to estimate a hydrodynamic particle diameter. The most common model utilized in DLS data analysis assumes that the sample consists of a distribution of spheres and reports this distribution accordingly. For non-spherical particles such as our microcrystals, this value represents an average rather than a true physical dimension and is biased toward larger or more slowly diffusing species. Consequently, DLS measurements of microcrystals are interpreted with this in mind when assessing batch size distributions and consistency (10-12).

To prepare microcrystals for DLS we took a crosslinked batch and washed the crystals to remove any free building blocks, protein precipitant, and other debris that could skew size distribution measurements. Microcrystals were pipetted out of the crosslinking bath into a microcentrifuge tube with deionized water, then pelleted at 2000 x g for 10 minutes. This allows the crystals to pellet, with any protein precipitant remaining in the supernatant above the crystal pellet. The deionized water was gently pipetted off the crystals, then replaced with fresh water to resuspend the crystals. This process was repeated five times to ensure all non-crystalline material was removed.

We used filtration to selectively size the particles. Because DLS is strongly biased toward larger particles, even minor populations of oversized crystals or aggregates can dominate the scattering signal, making prefiltration useful before measurement. In practice, filter selection must also account for non-ideal membrane retention behavior and clogging, since pore-size ratings do not act as perfectly sharp size cutoffs. Accordingly, a pore size (1.5 μm) somewhat larger than the target particle population (sub-micron) was chosen to reduce loss of borderline particles while still excluding larger contaminants.

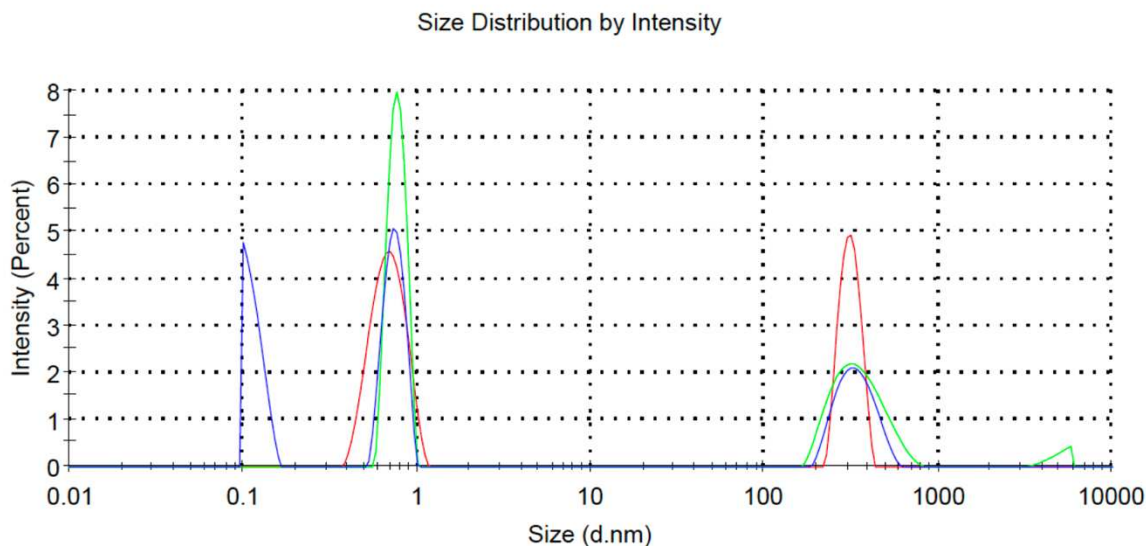


Figure 5.3 Batch characterization using dynamic light scattering. Three separate measurements are displayed in three different colors (red, blue, green). All three measurements show a peak below 1 nm (smaller than a single crystal building block) and a peak at an average diameter of $310 \text{ nm} \pm 40 \text{ nm}$.

Size selection reduces the influence of rare large particles or aggregates on DLS measurements while also demonstrating that microcrystal batches can be enriched for defined size ranges by selective filtration. DLS analysis of the filtered batch showed a primary crystal-associated population centered at 310 nm with an apparent standard deviation of $\pm 40 \text{ nm}$. However, the instrument-reported polydispersity index (PDI) was 0.411, indicating that the overall sample remained polydisperse despite filtration. This elevated PDI is likely influenced by the secondary peak below 1 nm, that is smaller than an intact CC1 crystal building block and is therefore more consistent with soluble species, small contaminants, or scattering artifacts than with microcrystals. Thus, the $\pm 40 \text{ nm}$ value describes the apparent width of the primary size population, whereas the PDI reflects the broader heterogeneity of the full measured sample. Future optimization should aim to reduce both the width of the crystal-associated population and

the contribution of secondary peaks, with a near-term target of $PDI \leq 0.2$ and a more stringent long-term target approaching $PDI \leq 0.1$ for highly uniform formulations. (13-15).

An aliquot of crystals from this batch were saved and imaged using scanning electron microscopy (SEM). SEM allows us to visualize crystals too small to be resolved via light microscopy, and confirms that crosslinked, washed, and filtered crystals maintained their structural integrity throughout sample prep (16). The aliquot was pipetted onto an aluminum stub, allowed to dehydrate, then coated with gold prior to SEM imaging. Coinciding with the DLS results, SEM imaging revealed that there was remaining salt in the washed microcrystal batches that obscured protein-DNA microcrystals from imaging due to salt crystal formation during dehydration prior to SEM imaging. Diluting the sample 1:100 with deionized water reduced the salt crystal population but made the protein-DNA microcrystal population sparse in the image field. A protein-DNA crystal of dimensions $0.86 \mu\text{m} \times 1.6 \mu\text{m}$ was observed to be well formed and structurally sound, with facet morphology analogous to large, vapor-diffusion grown CC1 crystals (Fig. 5.4). As wash protocols for CC1 microcrystals develop, DLS and SEM become even more viable methods for characterizing batches and ensuring wash and filtration are not mechanically disrupting the crystal structure.

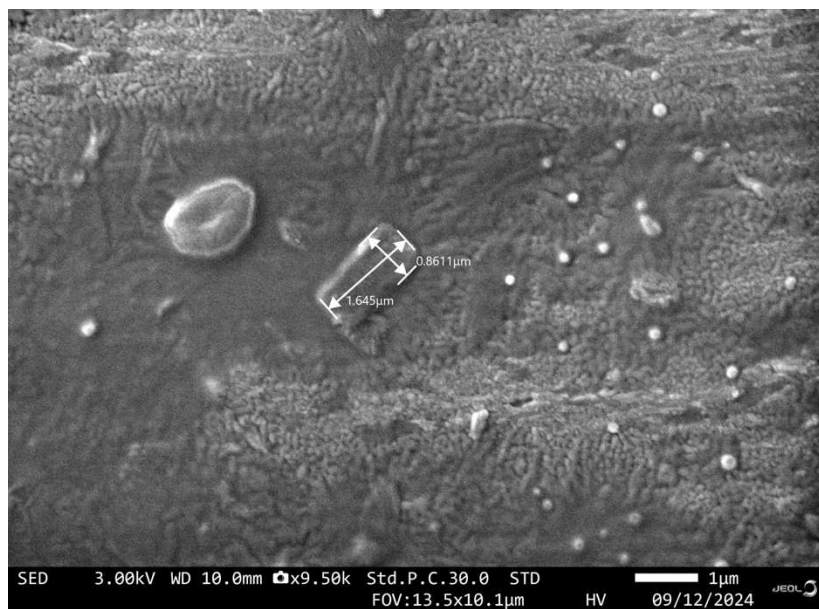


Figure 5.4 SEM image of a protein-DNA microcrystal after crosslinking, washing, filtration, and dilution in deionized water. The dimensions of the crystal are listed, with the facets of the crystal visible and analogous to large protein-DNA co-crystals.

5.3.2 Stability of microcrystals under biomimetic challenge conditions

For microcrystalline biomaterials to be viable beyond crystallization itself, it is necessary to evaluate whether these particles can persist outside of their native crystallization buffer. The previous section demonstrates the stability of CC1 microcrystals after crosslinking in deionized water, a solution that would typically dissolve noncovalently link microcrystals. Environments relevant to translational and materials-oriented applications such as acidic conditions and blood serum present chemical and biological challenges (17, 18). Accordingly, microcrystal stability under such conditions represents a key metric for assessing the suitability of our microcrystalline system for storage, handling, and delivery-oriented applications. To observe crystal persistence in these solutions we used a combination of white light and fluorescence microscopy. For blood serum tests, microcrystals from the same batch as the DLS and SEM characterization above were used. For pH trials microcrystals with a fluorescein appended to each building block were used, allowing us to track persistence using fluorescence microscopy. Fig. 5.5 shows white light

microscopy imaging of these fluorescein-tagged crystals, with a variety of crystal size distributions across the plate.

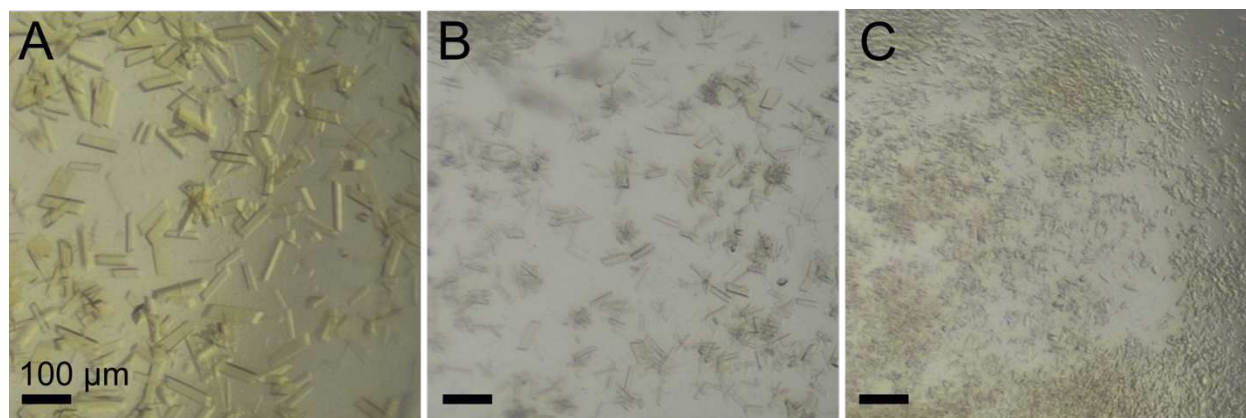


Figure 5.5: Fluorescein-tagged microcrystals grown via microbatch under oil. No clear trend was observed in the differences between salt and protein-DNA complex concentrations leading to the size differences between A, B, and C wells, similar to what was observed with non-tagged microcrystals. **A.** Crystals are well resolved via white light microscopy with dimensions greater than 100 μm . **B.** Crystals are smaller, but still resolvable with dimensions $\sim 50 \mu\text{m}$ across. **C.** Crystals are too small to be measured accurately, with a granular appearance indicative of microcrystals and distinct from protein precipitation.

When considering what biomimetic solutions would be relevant for drug delivery microcrystal, low pH environments are highly relevant. Fasted gastric mimics (a solution mimicking stomach acid in the absence of food) are typically ~ 1.6 pH HCl, with slightly higher pH (up to 1.9) simulating the gastric chamber after eating (18, 19). Here we show the difference in survival time of microcrystal CC1 tagged with fluorescein in pH 2 HCl. Crystals were crosslinked and washed using the protocol described above, and crystals not subjected to crosslinking were treated as a negative control. Initial images were taken with excitation at 488 nm and detection at 515 nm in deionized water for crosslinked crystals, and dilute mother liquor for non crosslinked (Figure 5.6 A and C), showing the field of microcrystals distinctly. A solution of pH 2 HCl was pipetted on top of crystals (5:1 acid:wash) and survival of the crystals was recorded via video. After 30 seconds the uncrosslinked crystals were mostly dissolved, as

apparent by the disappearance of the fluorescent signal as the crystal building blocks come apart and diffuse away (Fig. 5.6 C). In comparison, after 5 minutes crosslinked crystals were still apparent with a strong fluorescent signal persisting (Fig.5.6 D).

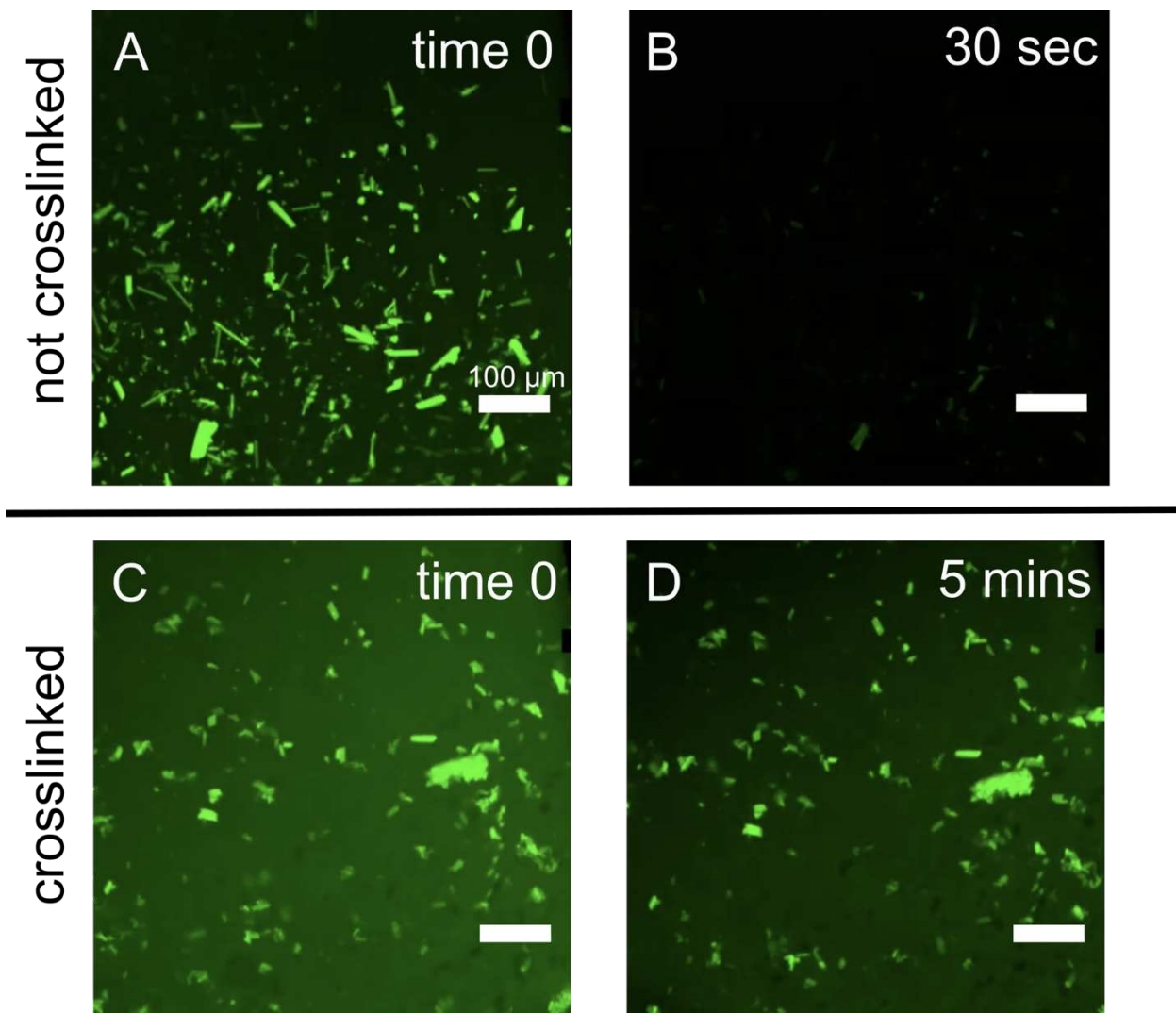


Figure 5.6 Crosslinked vs non crosslinked crystals in pH 2 HCl solution. A. Crystals at time 0 in 0.9 M Li_2SO_4 , 15 mM $\text{Mg}(\text{OAc})_2$, and 15 mM MES pH 6.0 were imaged using a 488 nm excitation and 515 nm emission filter. B. After addition (5:1 acid:wash volume) of pH 2 HCl crystals rapidly dissolved, with fluorescent building blocks diffusing away and becoming indistinguishable from background fluorescence. C. Crosslinked crystals at time 0 in deionized water. D. After addition (5:1 acid:wash volume) of pH 2 HCl, crystals remained intact with a strong fluorescent signal.

Blood serum presents a particularly relevant challenge for microcrystals, as the circulatory system is a common route for the delivery and transport of therapeutic materials (19). While serum is not a complete mimic of whole blood because it lacks blood cells and clotting factors, it provides a protein-rich, physiologically relevant medium containing soluble proteins, salts, and biomolecular components that can adsorb to particle surfaces, alter colloidal behavior, and promote destabilization or aggregation. (17, 18, 20). Proteases, binding proteins, and ionic fluctuations within serum can promote crystal dissolution, especially for particles stabilized primarily by noncovalent interactions (21, 22). As such, evaluating microcrystal stability both crosslinked and non-crosslinked in serum provides an important assessment of whether these materials can withstand physiologically relevant conditions encountered during potential delivery or circulation.

We utilized microcrystals lacking a fluorescent tag from the same batch used for DLS characterization to evaluate serum stability by white-light microscopy. Non-crosslinked crystals were first transferred into a mother liquor mimic and imaged prior to serum exposure. Upon addition of bovine blood serum, these crystals dissolved within minutes, indicating rapid destabilization in a complex biological environment. In contrast, crosslinked microcrystals in deionized water prior to serum exposure remained intact after the addition of bovine blood serum, with no observable dissolution over two weeks (Fig. 5.7). This side-by-side comparison highlights the critical role of chemical crosslinking in preserving microcrystal integrity under serum challenge conditions.

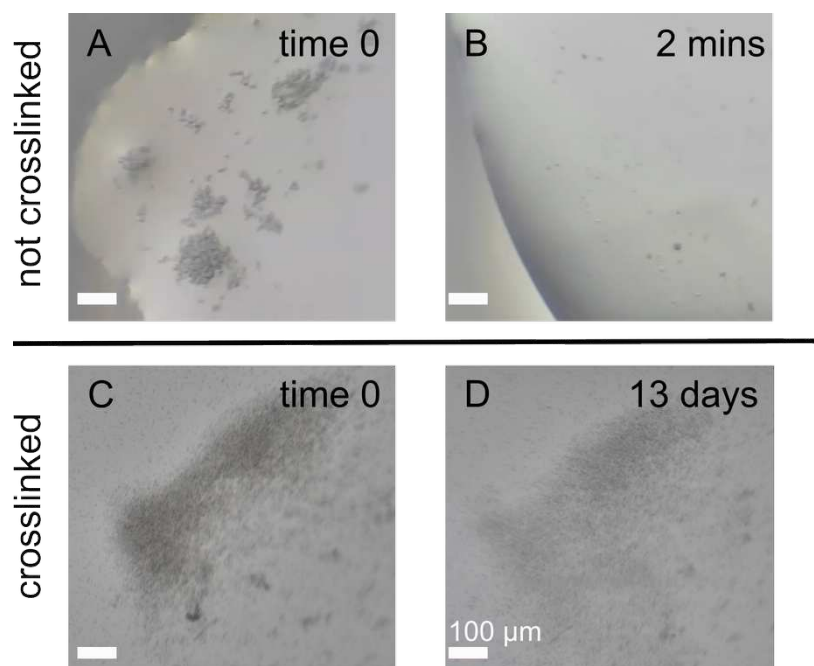


Figure 5.7: Crosslinked vs non crosslinked crystals in bovine blood serum. **A.** Crystals at time 0 in 0.9 M Li_2SO_4 , 15 mM $\text{Mg}(\text{OAc})_2$, and 15 mM MES pH 6.0 **B.** After addition of bovine blood serum crystals rapidly dissolved. **C.** Crosslinked crystals at time 0 in deionized water. **D.** After addition of bovine blood serum, crystals remained intact up to two weeks.

5.4 Discussion and conclusion

5.4.1 Translating co-crystals to materials with biomedical potential

Here we demonstrate that CC1 protein–DNA co-crystals can be reproducibly transitioned from a vapor-diffusion crystallization format to a batch microcrystalline material while retaining crystal identity and structural integrity. By adapting established crystallization conditions to an under-oil microbatch format, dense populations of micron-scale crystals were generated in a reproducible and scalable manner. These results establish a practical framework for producing CC1 in quantities and sizes that are compatible with materials-oriented characterization and downstream testing.

Chemical crosslinking is an essential enabling step for stabilizing CC1 microcrystals outside of their growth solution. EDC crosslinking preserved microcrystal integrity under a range

of destabilizing conditions, including exposure to deionized water, acidic environments, and bovine blood serum. In contrast, non-crosslinked microcrystals rapidly dissolved under identical challenges. This underscores the central role of covalent lattice reinforcement in converting protein-DNA co-crystals from biomolecular assemblies into mechanically robust microcrystalline materials.

5.4.2 Barriers remaining to wide adoption of co-crystals

Despite the advances demonstrated in this chapter, several challenges remain between the growth of stabilized CC1 microcrystals and their practical adaptation for use in biomedical applications. A major remaining target is improved control over distribution width. The filtered batch characterized here had an instrument-reported DLS PDI of 0.411, indicating that the full sample remained polydisperse even though a primary crystal-associated population was enriched near 310 nm. Future process development should aim to reduce the PDI toward ≤ 0.2 as an initial formulation target and ultimately toward ≤ 0.1 for applications requiring highly uniform particle behavior. While batch microcrystallization and chemical crosslinking enable the production of robust microcrystalline material, control over particle size distributions and physical heterogeneity remains limited, that may complicate reproducibility and application-specific optimization. In addition, although crosslinking confers substantial resistance to dilution, pH extremes, and serum exposure, the breadth of crosslinking tunability and cytocompatibility have yet to be fully explored (23). Scalability of handling and processing strategies, including washing, filtration, and storage protocols that preserve crystal integrity while minimizing loss or aggregation will need to be addressed. Considering these challenges will be essential for transitioning protein-DNA microcrystals to translational applications.

CHAPTER 5 BIBLIOGRAPHY

1. McPherson, A., & Gavira, J. A. (2014). Introduction to protein crystallization. *Acta Crystallographica Section F: Structural Biology Communications*, 70(1), 2–20. <https://doi.org/10.1107/S2053230X13033141>
2. Krauss, I. R., Merlino, A., Vergara, A., & Sica, F. (2013). An Overview of Biological Macromolecule Crystallization. *International Journal of Molecular Sciences*, 14(6), 11643–11691. <https://doi.org/10.3390/ijms140611643>
3. Arthur Jones, A., & D. Snow, C. (2024). Porous protein crystals: Synthesis and applications. *Chemical Communications*, 60(45), 5790–5803. <https://doi.org/10.1039/D4CC00183D>
4. McPherson, A., & Cudney, B. (2014). Optimization of crystallization conditions for biological macromolecules. *Acta Crystallographica Section F: Structural Biology Communications*, 70(11), 1445–1467. <https://doi.org/10.1107/S2053230X14019670>
5. Chayen, N. E. (1998). Comparative Studies of Protein Crystallization by Vapour-Diffusion and Microbatch Techniques. *Acta Crystallographica Section D*, 54(1), 8–15. <https://doi.org/10.1107/S0907444997005374>
6. Budziszewski, G. R., Stojanoff, V., & Bowman, S. E. J. (2025). Preparing for successful protein crystallization experiments. *Acta Crystallographica Section F: Structural Biology Communications*, 81(7), 272–280. <https://doi.org/10.1107/S2053230X25004650>
7. Orun, A. R., Slaughter, C. K., Shields, E. T., Vajapayajula, A., Jones, S., Shrestha, R., & Snow, C. D. (2024). Tuning Chemical DNA Ligation within DNA Crystals and Protein–DNA Cocystals. *ACS Nanoscience Au*. <https://doi.org/10.1021/acsnanoscienceau.4c00013>
8. Orun, A. R., Dmytriw, S., Vajapayajula, A., & Snow, C. D. (2022). Stabilizing DNA–Protein Co-Crystals via Intra-Crystal Chemical Ligation of the DNA. *Crystals*, 12(1), Article 1. <https://doi.org/10.3390/cryst12010049>
9. Shields, E. T., Slaughter, C. K., Mekkaoui, F., Magna, E. N., Shepherd, C., Lukeman, P. S., Spratt, D. E., & Snow, C. D. (2026). *Modular Scaffold Crystals for Programmable Installation and Structural Observation of DNA-Binding Proteins* (p. 2026.03.04.709581). bioRxiv. <https://doi.org/10.64898/2026.03.04.709581>
10. de Wijn, R., Rollet, K., Engilberge, S., McEwen, A. G., Hennig, O., Betat, H., Mörl, M., Riobé, F., Maury, O., Girard, E., Bénas, P., Lorber, B., & Sauter, C. (2020). Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering. *Crystals*, 10(2), 65. <https://doi.org/10.3390/cryst10020065>
11. Cheng, R. K. (2020). Towards an Optimal Sample Delivery Method for Serial Crystallography at XFEL. *Crystals*, 10(3), 215. <https://doi.org/10.3390/cryst10030215>
12. Joudeh, N., & Linke, D. (2022). Nanoparticle classification, physicochemical properties, characterization, and applications: A comprehensive review for biologists. *Journal of Nanobiotechnology*, 20, 262. <https://doi.org/10.1186/s12951-022-01477-8>

13. Clayton, K. N., Salameh, J. W., Wereley, S. T., & Kinzer-Ursem, T. L. (2016). Physical characterization of nanoparticle size and surface modification using particle scattering diffusometry. *Biomicrofluidics*, 10(5), 054107. <https://doi.org/10.1063/1.4962992>
14. K. Filippov, S., Khusnutdinov, R., Murmiliuk, A., Inam, W., Ya. Zakharova, L., Zhang, H., & V. Khutoryanskiy, V. (2023). Dynamic light scattering and transmission electron microscopy in drug delivery: A roadmap for correct characterization of nanoparticles and interpretation of results. *Materials Horizons*, 10(12), 5354–5370. <https://doi.org/10.1039/D3MH00717K>
15. Manaia, E. B., Abuçafy, M. P., Chiari-Andréo, B. G., Silva, B. L., Oshiro Junior, J. A., & Chiavacci, L. A. (2017). Physicochemical characterization of drug nanocarriers. *International Journal of Nanomedicine*, 12, 4991–5011. <https://doi.org/10.2147/IJN.S133832>
16. Beale, E. V., Warren, A. J., Trincão, J., Beilsten-Edmands, J., Crawshaw, A. D., Sutton, G., Stuart, D., & Evans, G. (2020). Scanning electron microscopy as a method for sample visualization in protein X-ray crystallography. *IUCrJ*, 7(3), 500–508. <https://doi.org/10.1107/S2052252520003875>
17. Klein, S. (2010). The Use of Biorelevant Dissolution Media to Forecast the *In Vivo* Performance of a Drug. *The AAPS Journal*, 12(3), 397–406. <https://doi.org/10.1208/s12248-010-9203-3>
18. Wang, Z., Lou, H., Dening, T. J., & Hageman, M. J. (2023). Biorelevant Dissolution Method Considerations for the Appropriate Evaluation of Amorphous Solid Dispersions: Are Two Stages Necessary? *Journal of Pharmaceutical Sciences*, 112(4), 1089–1107. <https://doi.org/10.1016/j.xphs.2022.12.008>
19. Chenthamara, D., Subramaniam, S., Ramakrishnan, S. G., Krishnaswamy, S., Essa, M. M., Lin, F.-H., & Qoronfleh, M. W. (2019). Therapeutic efficacy of nanoparticles and routes of administration. *Biomaterials Research*, 23(1), 20. <https://doi.org/10.1186/s40824-019-0166-x>
20. Kim, W., Ly, N. K., He, Y., Li, Y., Yuan, Z., & Yeo, Y. (2023). Protein corona: Friend or foe? Co-opting serum proteins for nanoparticle delivery. *Advanced Drug Delivery Reviews*, 192, 114635. <https://doi.org/10.1016/j.addr.2022.114635>
21. Wang, X., Zhang, S., Xu, Y., Zhao, X., & Guo, X. (2018). Ionic Strength-Responsive Binding between Nanoparticles and Proteins. *Langmuir*, 34(28), 8264–8273. <https://doi.org/10.1021/acs.langmuir.8b00944>
22. Böttger, R., Hoffmann, R., & Knappe, D. (2017). Differential stability of therapeutic peptides with different proteolytic cleavage sites in blood, plasma and serum. *PLOS ONE*, 12(6), e0178943. <https://doi.org/10.1371/journal.pone.0178943>
23. Hartje, L. F., Bui, H. T., Andales, D. A., James, S. P., Huber, T. R., & Snow, C. D. (2018). Characterizing the Cytocompatibility of Various Cross-Linking Chemistries for the Production of Biostable Large-Pore Protein Crystal Materials. *ACS Biomaterials Science & Engineering*, 4(3), 826–831. <https://doi.org/10.1021/acsbiomaterials.8b00023>

CHAPTER 6: SUMMARY OF FINDINGS AND FUTURE DIRECTIONS FOR PROTEIN-DNA CO-CRYSTALS AS FUNCTIONAL NANOPARTICLES

The work presented in this dissertation characterizes a biomaterial that combines the atomic precision of protein crystallography with the modular programmability of DNA nanotechnology. Across the preceding chapters, we show optimization of the protein-DNA co-crystal CC1 platform through probing the limits of modularity, exploring chemical and enzymatic ligation, and adapting CC1 for nucleic acid capture and release. This work collectively advances knowledge of the viability of CC1 in downstream applications. The result is a platform that is not only structurally well-defined but also capable of selective nucleic-acid capture, stabilization, and programmable release. These capabilities position CC1 as a versatile foundation for future work in structural biology, biosensing, and therapeutic delivery.

The first major contribution in **Chapter 2** of this dissertation was the demonstration that the DNA component of the co-crystal lattice can be extensively modified without disrupting crystallinity. By incorporating non-canonical nucleobases, alternative backbone chemistries, and engineered secondary structures, I showed that the lattice tolerates a plethora of chemical and structural diversity. This finding is significant because it establishes CC1 as a programmable scaffold in which DNA motifs of interest for different fields can be installed with spatial precision. We show that fluorescent tags for real time crystal observation via fluorescence microscopy are well-tolerated. We show, using the CC1 system, DNA-binding proteins bound to common DNA imperfections (base mismatch, bulge, deamination) and a DNA:RNA helix, that will be of particular interest to structural biologists who are looking to observe these motifs. The modularity with which we incorporated ssDNA tails suggests that the platform can be adapted to host a wide range of functional DNA elements, aptamers, catalytic motifs, sensors, and more,

opening the door to applications that require sequence-specific recognition or molecular sensing. The ability to engineer DNA within a rigid protein framework also provides a unique opportunity to explore how DNA structure behaves under crystallographic constraints, potentially revealing new principles of nucleic-acid organization.

Chapter 3 demonstrates further optimization of stabilization strategies to help overcome one of the central limitations of DNA-containing crystalline materials: their fragility in varied environments. One of the breakthrough applications of CC1 is in scaffold-assisted structural biology, where the protein-DNA lattice serves as a template for organizing DNA-binding proteins for structural observation. The innovation that has underpinned CC1's ability to function as this scaffold is the chemical ligation step - that fortifies CC1 outside of mother liquor and within solution conditions that allow guest protein binding. To expand the library of solutions in which CC1 can maintain structural integrity sufficient to allow high resolution X-ray diffraction data we have explored upping the yields of ligation. These studies show intracrystalline enzymatic ligation and then enzymatic, selective degradation of the lattice. A future co-crystal that is covalently reinforced at all three interfaces, protein-protein, protein-DNA, and DNA-DNA, would serve as a tunable, fully covalent, porous co-crystal with exceptional environmental resilience. The implications extend beyond structural support for scaffolded structural biology. By showing dynamic processes carried out by enzymes within the crystal, we advance the nanotechnology branch of CC1's applications. If the ligase and endonuclease activity demonstrated here are balanced so that building blocks of the crystal could be selectively replaced, we could imagine CC1 as a living biomaterial that could be reconstructed at will.

The contribution of **Chapter 4** of this dissertation was the demonstration that porous protein-DNA co-crystals can selectively capture nucleic acids from solution, store them within

the lattice, and release them in response to specific molecular triggers. Using toehold-mediated strand displacement, I showed that guest nucleic acids can be loaded into the pores and subsequently released with high specificity. This establishes the co-crystal as a responsive biomaterial capable of dynamic cargo exchange. The ability to load and unload nucleic acids in a controlled manner is a desirable trait in crystalline materials and suggests several future applications. In biosensing, the co-crystal could act as a molecular sieve that captures target sequences from complex environments and releases them only upon detection of a specific trigger. In therapeutic delivery, the lattice could protect fragile nucleic acids from degradation and release them in response to intracellular cues. In fundamental biophysics, the system provides a controlled environment for studying nucleic-acid interactions under confinement. For nanotechnology, the precise display of sequence specific anchors allows for anchoring of DNA nanotech, small molecules, and enzymes. This manuscript discusses how RepE54 has been shown to bind poly-T ssDNA simultaneous to its cognate dsDNA, as discussed in Wegryzn, *et al.*, 2023. In pursuit of nucleic acid capture, we explored whether this binding cleft could capture ssDNA once the target porous CC1 lattice was formed (Fig. 6.1). While the poly-T ssDNA appeared to bind specifically, we didn't achieve high enough occupancy to observe this binding via XRD. Still, combining this binding modality with the tailed crystals, one could dual anchor a nucleic acid guest. These could potentially be an RNA molecule whose secondary structure was of interest for scaffold assisted structural biology, or an aptamer that spans across the solvent channel pore. Future work exploring multi-anchor installations could make use of the work presented here.

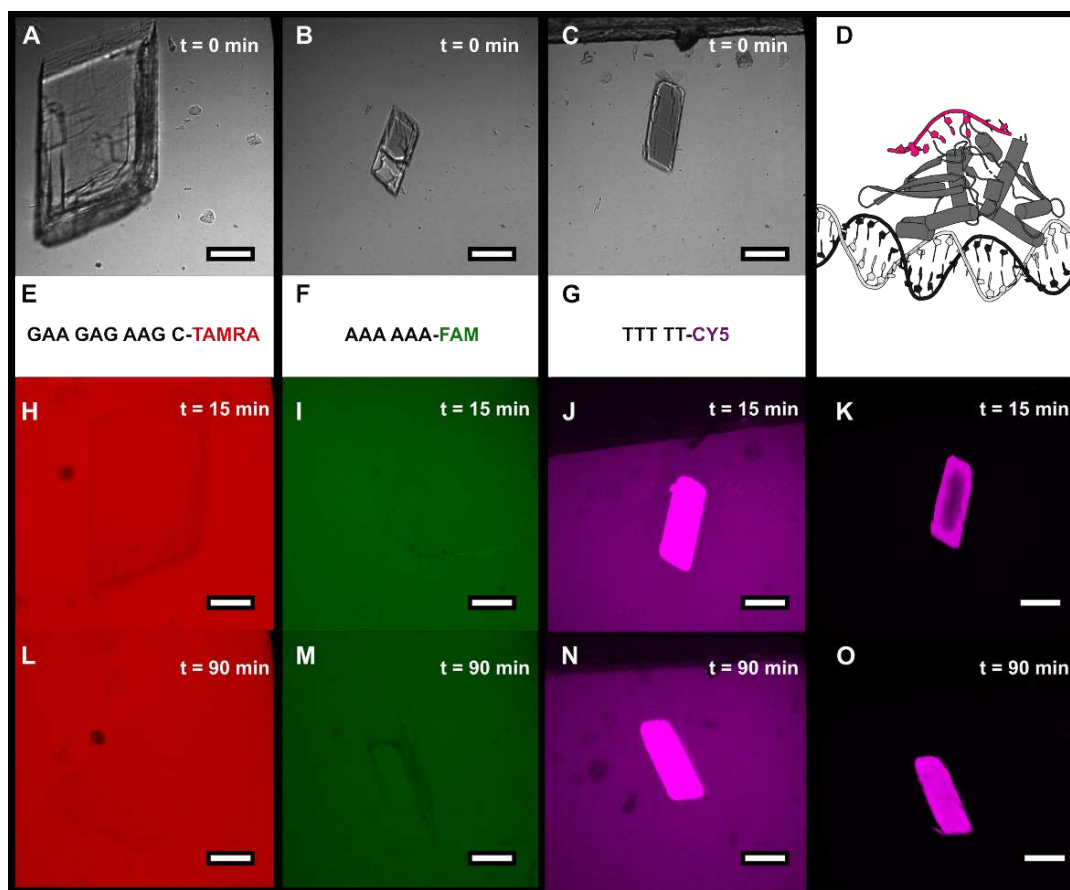


Figure 6.1: Poly-T tails specifically bind RepE54. Crystal structure from pdb 8AAN shown in **D**. All oligos were added a final concentration of 10 mM in 1.2 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, 50 mM MES pH 7.1. Sister crystals with no pendant tail, **A-C**., were soaked with various ssDNA with fluorophores (**E-G**). ssDNA with no thymine content shows no binding by RepE54 (**H,I,L,M**). Poly-T ssDNA tagged with Cy5 binds over 90 min (**J, N**). Edge-in loading shown with contrast adjusting (**K,O**).

Chapter 5 contributes to the development of a batch microcrystallization protocol and characterization of the stability of these batches post-chemical ligation. This report addresses practical challenges that limit the use of crystalline biomaterials outside of structural biology. By enabling reproducible production of co-crystals at scale and providing a straightforward method for characterizing these batches, these advances make the platform more accessible for translational research. As the field moves toward adaptation of CC1 for nanotech, scalable and reproducible manufacturing will be essential.

Looking forward, several promising research directions emerge from this work. One of the most compelling is the use of protein–DNA co-crystals as scaffolds for structural biology. Because the lattice diffracts to 3 Å resolution and contains programmable DNA binding sites, it offers a unique opportunity to organize nucleic-acid-binding proteins, RNAs, or small molecules within a pre-formed crystalline framework. This approach enables structural studies of molecules that are difficult to crystallize independently, such as transient complexes, and regulatory RNAs. Even more exciting is the possibility to observe subtle changes in the interaction between proteins and the CC1 scaffold in a bevy of different solution conditions, something impossible to do with traditional crystallography. The modularity of the DNA component means that binding sites can be installed with precision, allowing the scaffold to be adapted to a wide range of targets. As I move into postdoctoral work, I am particularly excited about the possibility of using this platform to capture RNA molecules for structural study, a particular sect of crystallography that has been desperate for innovation. The ability to position these structures within a rigid, high-resolution scaffold could reveal structural features that are otherwise inaccessible.

Another promising direction lies in biosensing and molecular sieving. The porous architecture of the co-crystal, combined with its programmable DNA elements, makes it an attractive candidate for sequence-specific capture of environmental nucleic acids, metabolites, or small molecules. The lattice could be engineered to respond to specific molecular cues, enabling controlled release or signal amplification. If a small molecule is installed on the DNA struts of CC1, we can bind aptamers within the large solvent channel pores. As demonstrated in Chapter 2, modified nucleotides containing bulky additions are well tolerated. Because the co-crystal is structurally well-defined, it also provides a unique environment for studying how molecules diffuse, bind, and react within confined spaces. These properties could be leveraged for

environmental monitoring, diagnostic assays, or catalytic applications.

A third direction involves therapeutic delivery. The enhanced stability achieved through ligation strategies suggests that protein-DNA co-crystals may eventually be suitable for *in vivo* use. Their ability to protect nucleic acids from degradation, combined with programmable release mechanisms, makes them attractive candidates for delivering therapeutic oligonucleotides, gene-editing components, or RNA-based drugs such as mir-34a. Before such applications can be realized, it will be essential to evaluate cytotoxicity, biodistribution, immune interactions, and clearance pathways. Surface modifications may be needed to improve biocompatibility or target specific tissues, of which the protein aspects of the co-crystal would be a good starting point. Nonetheless, the structural precision of the co-crystal offers a level of control that is difficult to achieve with amorphous nanoparticles.

Finally, advances in computational design, particularly tools capable of modeling protein–DNA interfaces and chemically modified nucleic acid, will accelerate the development of next-generation co-crystals. As structure prediction algorithms continue to improve, it will become possible to design new protein scaffolds with tailored pore geometries, engineer multi-component lattices with hierarchical organization, and incorporate synthetic nucleic acids with expanded chemical functionality. These tools will broaden the design space and enable increasingly sophisticated architectures with precisely engineered pore environments, guest-binding sites, and responsive behaviors. Additionally, with the advent of scaffold-assisted structural biology using the CC1 platform, generation of XRD structural information will become higher throughput. As the training set of determined structures grows, so too does the accuracy of prediction models.

In summary, this dissertation establishes protein-DNA co-crystals as a versatile and

engineerable biomaterial platform with broad potential impact. By demonstrating modular DNA engineering, lattice stabilization, programmable nucleic-acid capture and release, and scalable production, I have laid the groundwork for a new class of crystalline nanoparticles that combine structural precision with functional adaptability. The future of this platform is rich with possibility, and the work presented here represents only the beginning of what protein–DNA co-crystals may ultimately achieve.

APPENDIX I. Supplemental Information for Chapter 2: Crystalline Biomaterials for Site-

Specific Organization of Modified DNA

Supplementary Figure 2.1: Alternative crystal packing with short poly-T ssDNA tail.

Supplementary Figure 2.2: Expression of Y184E RepE54 mutant

Supplementary Figure 2.3: Electron density maps of 2-adenine bulge CC1+10

Supplementary Figure 2.4: CC1+10 with TAMRA appended to each building block

Supplementary Figure 2.5: Interfacial Mg(II) and PT bonds

Supplementary Figure 2.6: Overlay of QM pose with crystallographic model

Supplementary Table 2.1: XRD statistics for CC1+10 with dU:dT bond

Supplementary Table 2.2: XRD statistics for CC1+10 with a pendant fluorescein and G:G mismatch

Supplementary Table 2.3: XRD statistics for CC1+10 with 2-adenine bulge

Supplementary Table 2.4: XRD statistics for CC1+10 with a pendant TAMRA

Supplementary Table 2.5: XRD statistics for CC1+12 with 12 bases of RNA:DNA hybrid

Supplementary Table 2.6: XRD statistics for CC1+10 with phosphorothioate backbone modifications

Supplementary Table 2.7: XRD statistics for CC1+10 with central CCCC tail

Supplementary Table 2.8: XRD statistics for CC1+10 with central CCCC tail and FAM-AAGGG guest

Supplementary Table 2.9: XRD statistics for CC1+10 with end CCCC tail

Supplementary Table 2.10: XRD statistics for CC1+10 with end CCCC tail and FAM-AAGGG guest

Supplementary Table 2.11: XRD statistics for CC1+10 with end 6-nt poly-T tail

Supplementary Table 2.12: XRD statistics for CC1+10 with central 6-nt poly-T tail

Supplementary Table 2.13: XRD statistics for CC1+10 Y184E

Supplementary Table 2.14: Crystallization conditions of CC1+10 (or 12)

Supplementary Table 2.15: Protein sequences

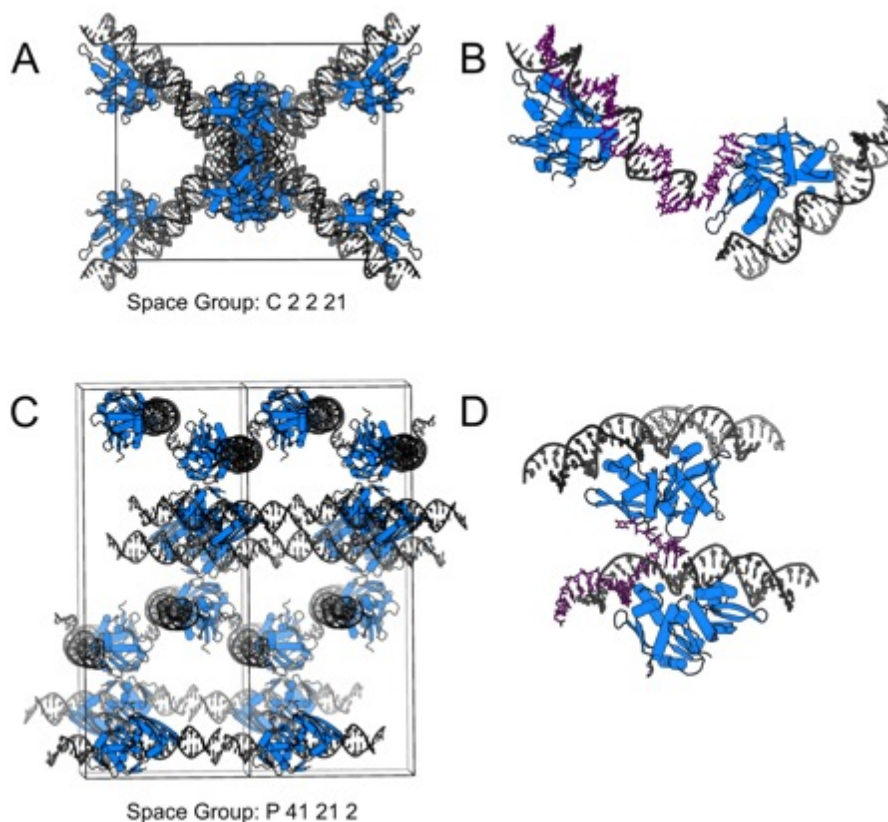
Supplementary Table 2.16: CC1+10 (or 12) oligonucleotide sequences

Supplementary Table 2.17: Oligonucleotide sequences that have yet to crystallize

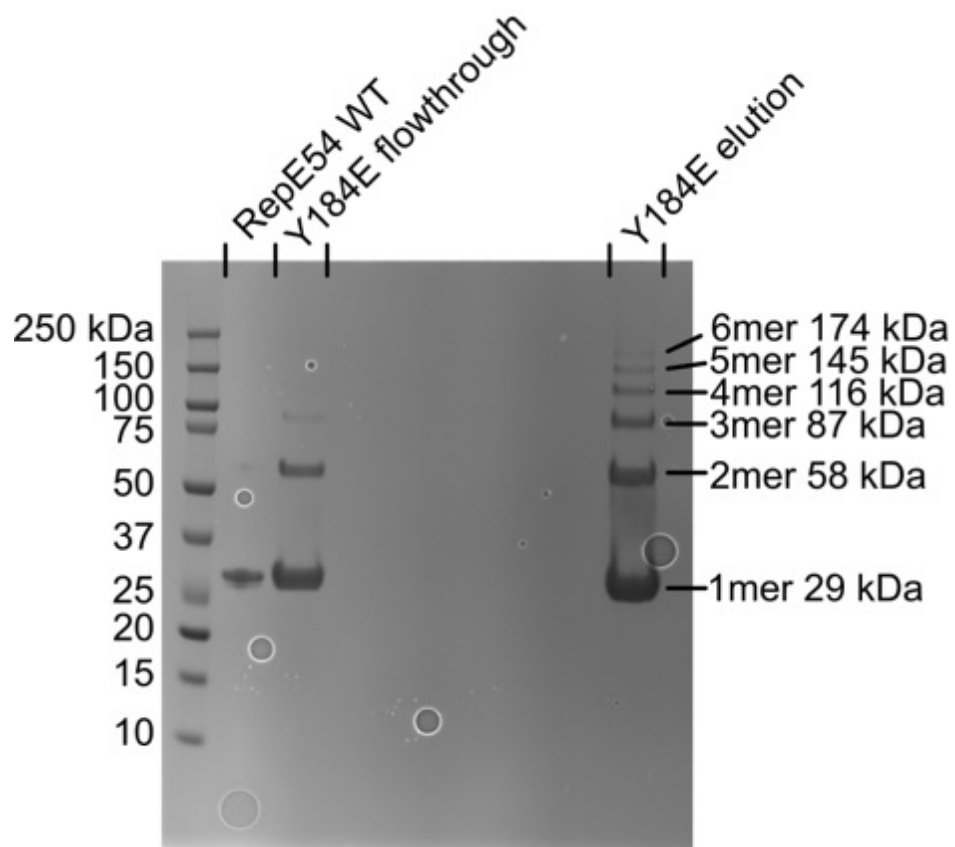
Supplementary Table 2.18: x3DNA analysis of helical geometry as compared to averages of CC1+10

Supplementary Table 2.19: x3DNA analysis of CC1+12 against references

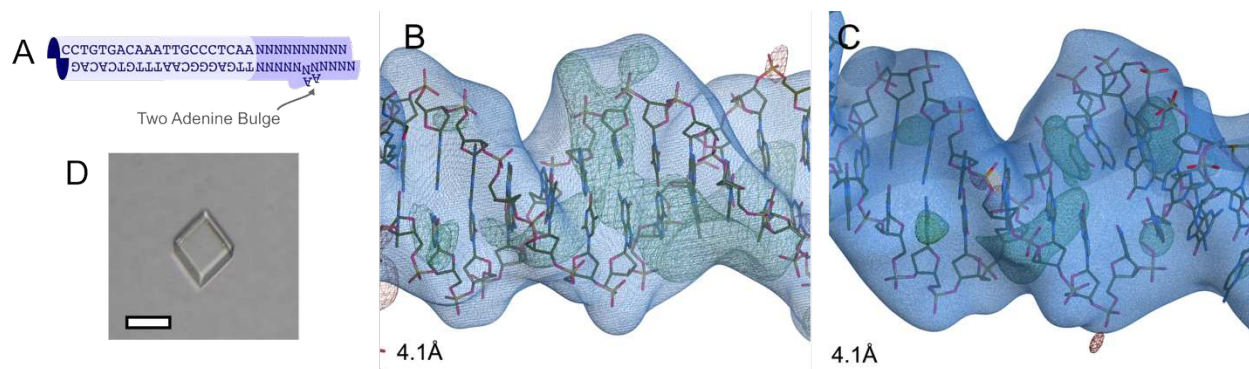
Supplementary Table 2.20: RMSD ranking of QM conformers and their energy scores



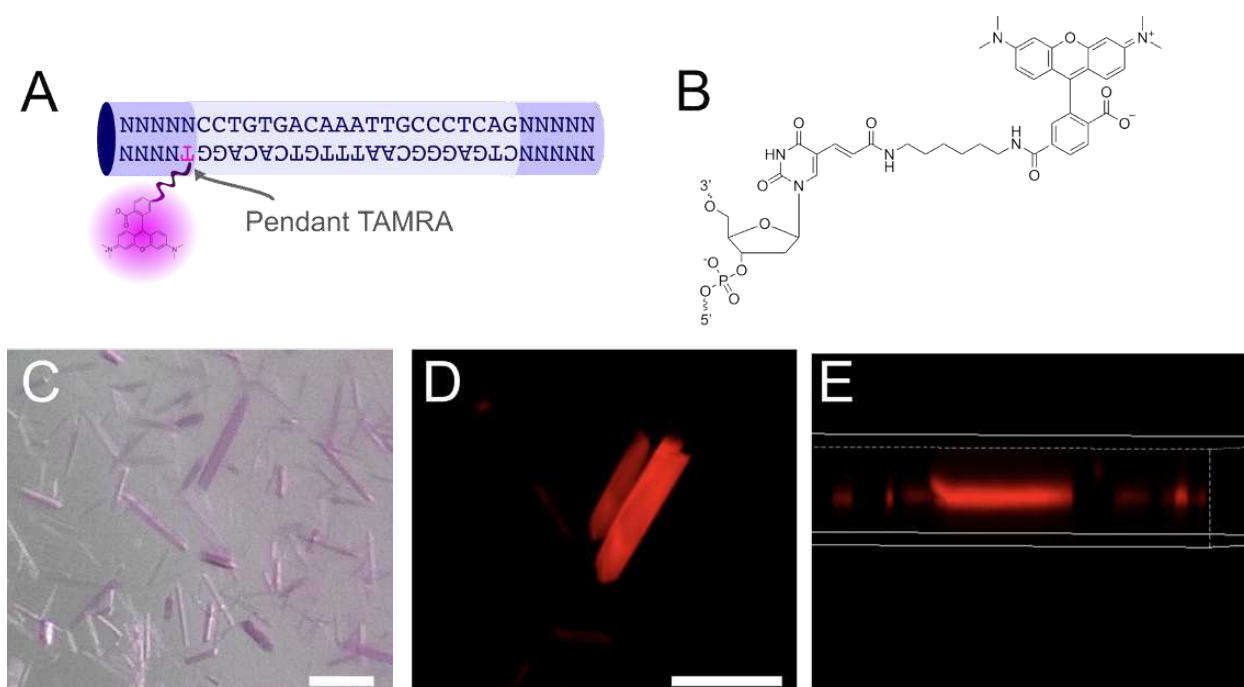
Supplementary Figure 2.1: Alternative crystal packing with short poly-T ssDNA tail. **A.** When a six nt poly-T tail is added to the end of an asymmetrically expanded building block, crystal packing shifts to allow poly-T:RepE54 binding (**B**, shown in magenta) while maintaining DNA:DNA stacking and one of the protein:protein contacts observed in CC1. **C.** When a six nt poly-T tail is added directly under RepE54's binding sequence, crystal packing shifts to allow poly-T:RepE54 binding (**D**, shown in magenta) while maintaining DNA:DNA stacking and one of the protein:protein contacts observed in CC1.



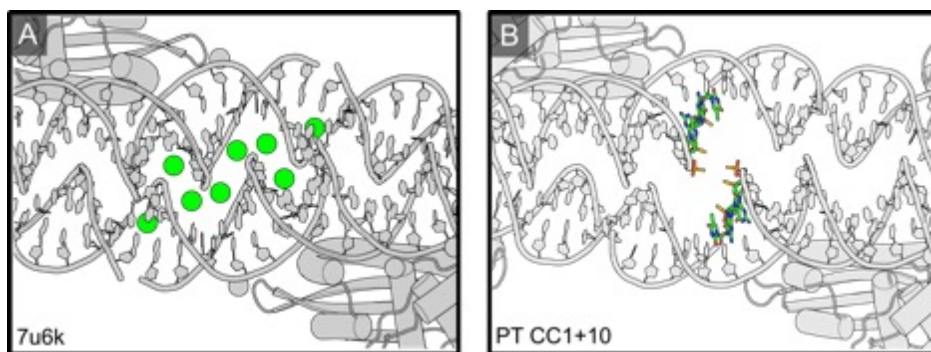
Supplementary Figure 2.2: Purification of Y184E mutant of RepE54. SDS-PAGE of RepE54 mutant Y184 heparin column purification. RepE54 WT in lane 1. Flowthrough of heparin column in lane 3. Lane 4 is the elution fraction showing strong oligomerization of Y184E despite the denaturing conditions of the SDS-PAGE, a phenomenon common when expressing WT RepE54 (18).



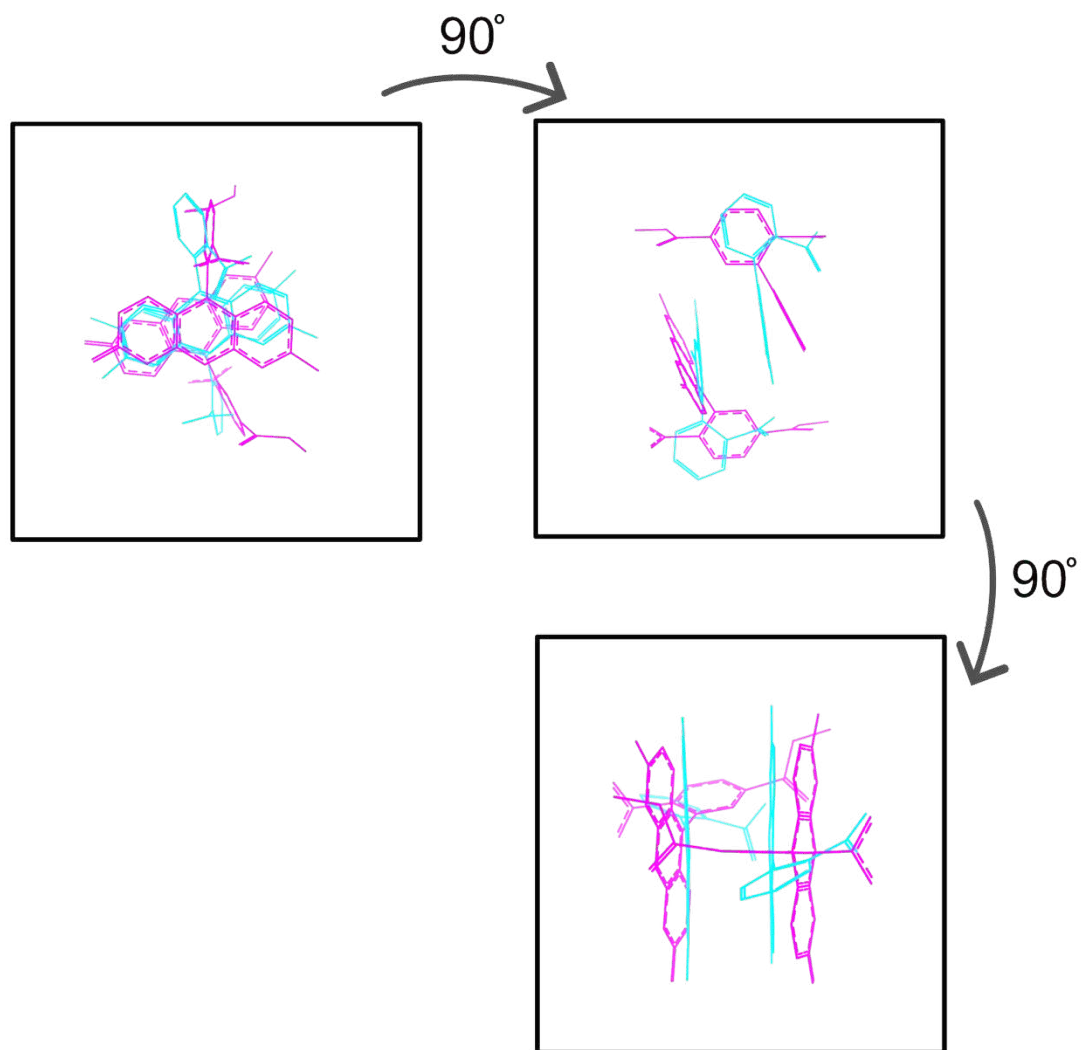
Supplementary Figure 2.3: 2-adenine bulge CC1⁺¹⁰. **A.** Schematic of DNA design, with the bulge placed within the asymmetric expansion region. **B.** and **C.** Composite 2Fo–Fc maps contoured at 1 σ and Fo–Fc maps contoured at 3 σ . **B.** Shows refinement with canonical asymmetric CC1⁺¹⁰, with unmodelled density positioned at the bulge inclusion. **C.** Refinement with 2-adenine bulge modelled in, fitting the density map more accurately than **B.** **D.** White light microscopy of 2-adenine bulge crystal. Scale bar is 100 μm .



Supplemental Figure 2.4: CC1⁺¹⁰ with TAMRA appended to each building block. A) CC1⁺¹⁰ building block showing symmetric expansion and pendant TAMRA B) Internal thymidine modification showing flexible linker and TAMRA moiety. C) White light microscopy, scale bar is 100 μm . D) 561 nm laser at 3% showing uniform fluorescence in one plane. E) Volume view of emission in all planes of the crystal shown in D. This phenomena of emission only in the lower planes is explored extensively in our previous work (18). Briefly, our confocal microscope excites fluors from the microscope slide up through the z-plane, and the high local concentration of fluors leads to extinction as the excitation plane advances through the crystal volume. This leads to bright emission along the planes closest to the microscope slide and along overhangs which have minimal extinction.



Supplementary Figure 2.5: Comparison of ipCC1 interfacial magnesium placement and phosphorothioate modifications. **A.** In dark grey are (pdb 7U6K) and symmetry mate, in green are magnesium ions. **B.** In white are PT modified CC1⁺¹⁰ and symmetry mate, shown in sticks and colored by element are the phosphorothioate modifications.



Supplementary Figure 2.6. Overlay of the crystallographic fluorescein dimer (gray) with the lowest-RMSD conformer from the AIMNET2-ranked conformational ensemble (teal; RMSD = 1.89 Å over the 50 shared heavy atoms of the two fluorescein cores), supporting the conformational plausibility of the X-ray-derived dimer pose.

Table S2.1. X-ray diffraction statistics for CC1⁺¹⁰ with dU:dT bond

PDB code	36WP
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	46.8 - 3.35 (3.44 - 3.35)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	73.79 127.95 136.57
α , β , γ (°)	90.00 91.10 90.00
Unique reflections	19669 (4007)
Multiplicity	3.4 (3.5)
Completeness (%)	99.2 (98.5)
Mean I/sigma(I)	3.5 (0.4)
Wilson B-factor	121.85
R-merge	0.136 (1.942)
R-meas	0.200 (610)
R-pim	0.107 (1.375)
CC1/2	0.999 (0.591)
Refinement	
Reflections used in refinement	18338 (1311)
Reflections used for Rfree	1831 (128)
R-work	0.2672 (0.4531)
R-free	0.2808 (0.4993)
Number of non-hydrogen atoms	3094
Macromolecules	3093
Ligands	1
Solvent	0
Protein residues	221
RMS (bonds) (Å)	0.008
RMS (angles) (°)	0.93
Ramachandran favored (%)	97.21
Ramachandran allowed (%)	33

Ramachandran outliers (%)	0.47
Rotamer outliers (%)	1.03
Clashscore	7.36
Average B-factor	141.07

Table S2.2 X-ray diffraction statistics for CC1⁺¹⁰ with a pendant fluorescein and G:G mismatch

PDB code	36NQ
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	39.8 - 3.25 (3.34 - 3.25)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	74.301 125.041 133.964
α , β , γ (°)	90 90.05 90
Unique reflections	54514 (522)
Multiplicity	3.4 (3.6)
Completeness (%)	96.94 (94.82)
Mean I/sigma(I)	4.6 (0.5)
Wilson B-factor	133.88
R-merge	0.085 (071)
R-meas	0.130 (998)
R-pim	0.070 (1.570)
CC1/2	0.996 (0.524)
Refinement	
Reflections used in refinement	18782 (1410)
Reflections used for Rfree	1876 (128)
R-work	0.2471 (0.4150)
R-free	0.2928 (0.4439)
Number of non-hydrogen atoms	2993
Macromolecules	2954
Ligands	39
Solvent	0
Protein residues	223
RMS (bonds) (Å)	0.017
RMS (angles) (°)	1.42
Ramachandran favored (%)	98.62
Ramachandran allowed (%)	0.92

Ramachandran outliers (%)	0.46
Rotamer outliers (%)	0.63
Clashscore	8.15
Average B-factor	238.28

Table S2.3. X-ray diffraction statistics for CC1⁺¹⁰ with 2-adenine bulge

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	70.55 - 4.1 (4.32 - 4.1)
Space group	C121
Unit cell dimensions	
a, b, c (Å)	156.852 121.785 75.238
α , β , γ (°)	90 115.9 90
Unique reflections	12415 (1733)
Multiplicity	7.0 (7.1)
Completeness (%)	96.09 (83.52)
Mean I/sigma(I)	4.32 (0.14)
Wilson B-factor	266.27
R-merge	0.1714 (1.249)
R-meas	0.1856 (1.339)
R-pim	0.07016 (0.4816)
CC1/2	0.997 (0.0876)
Refinement	
Reflections used in refinement	9686 (1191)
Reflections used for Rfree	985 (119)
R-work	0.2987 (0.6292)
R-free	0.3092 (0.6669)
Number of non-hydrogen atoms	2930
Macromolecules	2929
Ligands	1
Solvent	0
Protein residues	220
RMS (bonds) (Å)	0.025
RMS (angles) (°)	0.97
Ramachandran favored (%)	97.66
Ramachandran allowed (%)	1.87

Ramachandran outliers (%)	0.47
Rotamer outliers (%)	0
Clashscore	14.32
Average B-factor	365.8

Table S2.4. X-ray diffraction statistics for CC1⁺¹⁰ with a pendant TAMRA

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	66.98 - 4.118 (4.53 - 4.12)
Space group	I222
Unit cell dimensions	
a, b, c (Å)	73.636 127.97 133.965
α , β , γ (°)	90 90 90
Unique reflections	5126 (1239)
Multiplicity	3.6 (3.8)
Completeness (%)	99.48 (99.36)
Mean I/sigma(I)	3.97 (1.25)
Wilson B-factor	126.56
R-merge	0.8992 (1.072)
R-meas	0.9367 (1.116)
R-pim	0.2588 (0.3057)
CC1/2	0.8 (0.258)
Refinement	
Reflections used in refinement	5125 (1239)
Reflections used for Rfree	515 (123)
R-work	0.2779 (0.3680)
R-free	0.2913 (0.3746)
Number of non-hydrogen atoms	3112
Macromolecules	3111
Ligands	1
Solvent	0
Protein residues	221
RMS (bonds) (Å)	0.004
RMS (angles) (°)	0.70
Ramachandran favored (%)	97.62
Ramachandran allowed (%)	38

Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	9.45
Average B-factor	141.74

Table S2.5. X-ray diffraction statistics for CC1+12 with 12 bases of RNA:DNA hybrid

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	47.77 - 3.66 (3.79 - 3.66)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	73.9 13915 130.771
α , β , γ (°)	90 95.36 90
Unique reflections	13688 (3270)
Multiplicity	3.5 (3.5)
Completeness (%)	98.3 (98.9)
Mean I/sigma(I)	9.2 (1.0)
Wilson B-factor	176.41
R-merge	0.050 (1.460)
R-meas	0.078 (1.888)
R-pim	0.041 (0.995)
CC1/2	0.994 (0.378)
Refinement	
Reflections used in refinement	13635 (1372)
Reflections used for Rfree	1358 (140)
R-work	0.2786 (0.4195)
R-free	0.2867 (0.4228)
Number of non-hydrogen atoms	2824
Macromolecules	2822
Ligands	2
Solvent	0
Protein residues	208
RMS (bonds) (Å)	0.006
RMS (angles) (°)	0.84
Ramachandran favored (%)	96.46
Ramachandran allowed (%)	3.03

Ramachandran outliers (%)	0.51
Rotamer outliers (%)	1.68
Clashscore	7.38
Average B-factor	247.63

Table S2.6. X-ray diffraction statistics for CC1⁺¹⁰ with phosphorothioate backbone modifications

PDB code	36OW
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	46.02 - 3.14 (3.28 - 3.14)
Space group	I222
Unit cell dimensions	
a, b, c (Å)	73.215 128.63 131.758
α , β , γ (°)	90 90 90
Unique reflections	11160 (1370)
Multiplicity	6.7 (7.1)
Completeness (%)	99.47 (99.93)
Mean I/sigma(I)	3.8 (1.6)
Wilson B-factor	67.73
R-merge	0.328 (1.082)
R-meas	0.384 (1.260)
R-pim	0.148 (0.469)
CC1/2	0.960 (0.672)
Refinement	
Reflections used in refinement	11160 (1370)
Reflections used for Rfree	1116 (142)
R-work	0.2770 (0.3693)
R-free	0.3136 (0.4155)
Number of non-hydrogen atoms	2908
Macromolecules	2781
Ligands	127
Solvent	0
Protein residues	209
RMS (bonds) (Å)	0.003
RMS (angles) (°)	0.74
Ramachandran favored (%)	96.52
Ramachandran allowed (%)	3.48

Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	4.36
Average B-factor	87.70

Table S2.7. X-ray diffraction statistics for CC1⁺¹⁰ with central C4 tail

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	65.75 - 3.9 (4.08 - 3.9)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	74.681 130.518 131.499
α , β , γ (°)	90 90.45 90
Unique reflections	11549 (1691)
Multiplicity	1.0 (1.0)
Completeness (%)	96.87 (85.17)
Mean I/sigma(I)	32 (0.99)
Wilson B-factor	157.94
R-merge	0.282 (2080)
R-meas	0.309 (18.211)
R-pim	0.089 (4.872)
CC1/2	0.988 (0.517)
Refinement	
Reflections used in refinement	11217 (1218)
Reflections used for Rfree	1122 (134)
R-work	0.2521 (0.5832)
R-free	0.2991 (0.5778)
Number of non-hydrogen atoms	3081
Macromolecules	3079
Ligands	2
Solvent	0
Protein residues	226
RMS (bonds) (Å)	0.005
RMS (angles) (°)	0.81
Ramachandran favored (%)	96.85
Ramachandran allowed (%)	70

Ramachandran outliers (%)	0.45
Rotamer outliers (%)	0
Clashscore	13.71
Average B-factor	280.40

Table S2.8. X-ray diffraction statistics for CC1⁺¹⁰ with central C4 tail and FAM-AAGGG guest

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	67.69 - 4.15 (4.37 - 4.15)
Space group	C121
Unit cell dimensions	
a, b, c (Å)	154.097 126.7 73.816
α , β , γ (°)	90 118.54 90
Unique reflections	9409 (2665)
Multiplicity	6.8 (6.9)
Completeness (%)	98.53 (99.26)
Mean I/sigma(I)	10 (0)
Wilson B-factor	235.04
R-merge	0.087 (1.016)
R-meas	0.104 (1.201)
R-pim	0.041 (0.453)
CC1/2	0.972 (0.814)
Refinement	
Reflections used in refinement	9365 (1346)
Reflections used for Rfree	958 (130)
R-work	0.2432 (0.3986)
R-free	0.2512 (0.4635)
Number of non-hydrogen atoms	3081
Macromolecules	3079
Ligands	2
Solvent	0
Protein residues	226
RMS (bonds) (Å)	0.005
RMS (angles) (°)	0.78
Ramachandran favored (%)	97.30
Ramachandran allowed (%)	70

Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	7.32
Average B-factor	284.52

Table S2.9. X-ray diffraction statistics for CC1⁺¹⁰ with end C4 tail

PDB code	36OT
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	45.7 - 3.4 (3.5 - 3.4)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	73.24 117.288 144.284
α , β , γ (°)	90 93.89 90
Unique reflections	16527 (1413)
Multiplicity	3.5 (3.6)
Completeness (%)	98.1 (98.88)
Mean I/sigma(I)	5.5 (0.7)
Wilson B-factor	1351
R-merge	0.080 (1.639)
R-meas	0.124 (198)
R-pim	0.067 (1.149)
CC1/2	0.994 (0.267)
Refinement	
Reflections used in refinement	16527 (1413)
Reflections used for Rfree	1666 (146)
R-work	0.2535 (0.4205)
R-free	0.2874 (0.4546)
Number of non-hydrogen atoms	2919
Macromolecules	2917
Ligands	2
Solvent	0
Protein residues	221
RMS (bonds) (Å)	0.004
RMS (angles) (°)	0.70
Ramachandran favored (%)	95.81
Ramachandran allowed (%)	4.19

Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	9.57
Average B-factor	191.13

Table S2.10. X-ray diffraction statistics for CC1⁺¹⁰ with end C4 tail and FAM-AAGGG guest

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	71.4 - 4.6 (4.98 - 4.6)
Space group	C121
Unit cell dimensions	
a, b, c (Å)	154.762 123.085 73.496
α , β , γ (°)	90 112.72 90
Unique reflections	13020 (2573)
Multiplicity	7.0 (6.9)
Completeness (%)	98.96 (97.37)
Mean I/sigma(I)	4.28 (0.30)
Wilson B-factor	243.42
R-merge	0.3149 (3.589)
R-meas	0.343 (3.877)
R-pim	0.1335 (1.453)
CC1/2	0.875 (0.19)
Refinement	
Reflections used in refinement	6979 (1369)
Reflections used for Rfree	702 (135)
R-work	0.2129 (0.2930)
R-free	0.2512 (0.3501)
Number of non-hydrogen atoms	3011
Macromolecules	3010
Ligands	1
Solvent	0
Protein residues	217
RMS (bonds) (Å)	0.004
RMS (angles) (°)	0.66
Ramachandran favored (%)	98.10
Ramachandran allowed (%)	1.90

Ramachandran outliers (%)	0
Rotamer outliers (%)	0.59
Clashscore	7.35
Average B-factor	262.79

Table S2.11. X-ray diffraction statistics for CC1⁺¹⁰ with end 6-nt poly-T tail

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	49.1 - 4.0 (4.21 - 4.0)
Space group	C 2 2 21
Unit cell dimensions	
a, b, c (Å)	12572 15712 128.219
α , β , γ (°)	90 90 90
Unique reflections	21142 (3767)
Multiplicity	13.6 (14.4)
Completeness (%)	99.8 (99.4)
Mean I/sigma(I)	7.9 (0.8)
Wilson B-factor	193.09
R-merge	0.168 (3.619)
R-meas	0.181 (3.878)
R-pim	0.049 (1.015)
CC1/2	0.998 (0.694)
Refinement	
Reflections used in refinement	10477 (1473)
Reflections used for Rfree	997 (141)
R-work	0.2696 (0.3430)
R-free	0.3127 (0.4468)
Number of non-hydrogen atoms	2950
Macromolecules	2949
Ligands	1
Solvent	0
Protein residues	216
RMS (bonds) (Å)	0.005
RMS (angles) (°)	0.78
Ramachandran favored (%)	95.71
Ramachandran allowed (%)	3.33

Ramachandran outliers (%)	0.95
Rotamer outliers (%)	0
Clashscore	147
Average B-factor	259.68

Table S2.12 X-ray diffraction statistics for CC1⁺¹⁰ with central 6-nt poly-T tail

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	49.24 - 4.5 (4.85 - 4.5)
Space group	P 41 21 2
Unit cell dimensions	
a, b, c (Å)	98.476 98.476 233.019
α , β , γ (°)	90 90 90
Unique reflections	12560 (3472)
Multiplicity	18 (13.0)
Completeness (%)	99.8 (99.4)
Mean I/sigma(I)	6.3 (1.1)
Wilson B-factor	2224
R-merge	0.221 (397)
R-meas	0.239 (556)
R-pim	0.068 (0.702)
CC1/2	0.999 (0.309)
Refinement	
Reflections used in refinement	7299 (1401)
Reflections used for Rfree	731 (140)
R-work	0.2530 (0.2742)
R-free	0.3102 (0.3632)
Number of non-hydrogen atoms	3158
Macromolecules	3157
Ligands	1
Solvent	0
Protein residues	221
RMS (bonds) (Å)	0.011
RMS (angles) (°)	1.43
Ramachandran favored (%)	95.35
Ramachandran allowed (%)	3.72

Ramachandran outliers (%)	0.93
Rotamer outliers (%)	0.53
Clashscore	25.25
Average B-factor	328.53

Table S2.13. X-ray diffraction statistics for CC1⁺¹⁰ Y184E

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	46.01 - 3.56 (3.69 - 3.56)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	73.329 129.131 131.221
α , β , γ (°)	90 91.83 90
Unique reflections	14077 (3418)
Multiplicity	3.6 (3.6)
Completeness (%)	95.8 (97.9)
Mean I/sigma(I)	4.5 (0.7)
Wilson B-factor	150.31
R-merge	0.097 (1.552)
R-meas	0.136 (048)
R-pim	0.072 (1.065)
CC1/2	0.997 (0.290)
Refinement	
Reflections used in refinement	14036 (1409)
Reflections used for Rfree	1383 (140)
R-work	0.2491 (0.3998)
R-free	0.2868 (0.4563)
Number of non-hydrogen atoms	2783
Macromolecules	2782
Ligands	1
Solvent	0
Protein residues	217
RMS (bonds) (Å)	0.005
RMS (angles) (°)	0.80
Ramachandran favored (%)	98.58
Ramachandran allowed (%)	1.42

Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	7.12
Average B-factor	219.50

Table S2.14. Crystallization conditions of CC1⁺¹⁰ (or 12)

PDB Code:	Crystallization Condition:
Abbreviations: (pdb code pending release)	MgAc ₂ = magnesium acetate, Mg(CH ₃ COO) ₂ ; Li ₂ SO ₄ = lithium sulfate; MES = 2-(N-morpholino)ethanesulfonic acid.
dU:dT	10 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5
TAMRA	20 mM MgAc ₂ , 1.4 M Li ₂ SO ₄ , 50 mM MES pH 6.5
PT	20 mM MgAc ₂ , 1.8 M Li ₂ SO ₄ , 50 mM MES pH 6.5
End poly-C	30 mM MgAc ₂ , 1.2 M Li ₂ SO ₄ , 50 mM MES pH 6.5
End poly-C with guest	30 mM MgAc ₂ , 1.2 M Li ₂ SO ₄ , 50 mM MES pH 6.5
Central poly-T	30 mM MgAc ₂ , 1.3 M Li ₂ SO ₄ , 50 mM MES pH 6.5
Y184E no tail	30 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5
FAM and mismatch	30 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5
Central poly-C with guest	40 mM MgAc ₂ , 1.2 M Li ₂ SO ₄ , 50 mM MES pH 6.5
End poly-T	40 mM MgAc ₂ , 1.2 M Li ₂ SO ₄ , 50 mM MES pH 6.5
Central poly-C	40 mM MgAc ₂ , 1.2 M Li ₂ SO ₄ , 50 mM MES pH 6.5
2 adenine bulge	40 mM MgAc ₂ , 1.5 M Li ₂ SO ₄ , 50 mM MES pH 6.5
RNA:DNA hybrid	40 mM MgAc ₂ , 1.8 M Li ₂ SO ₄ , 50 mM MES pH 6.5

Table S2.15. Protein sequences

Protein Name	Amino Acid Sequence
RepE54 Transcription Factor	MRGS HHHHHH GSMAETA VINHKKRKNSPRIVQSNDLTEAAYSLSR DQKRMLYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEAS KDIRQALKSFAGKEVVFYRPEEDAGDEKGYESFPWFIKPAHSPSRG LYSVHINPYLIPFFIGLQNRFTQFRLSETKEITNPYAMRLYESLCQYR KPDGSGIVSLKIDWIIERYQLPQSYQRMPDFRRRFLQVCVNEINSRT PMRLSYIEKKKGRQTTHIVFSFRDITSMTTG
RepE54 variant (Y184E) <i>Mutation underlined</i>	MRGS HHHHHH GSMAETA VINHKKRKNSPRIVQSNDLTEAAYSLSR DQKRMLYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEAS KDIRQALKSFAGKEVVFYRPEEDAGDEKGYESFPWFIKPAHSPSRG LYSVHINPYLIPFFIGLQNRFTQFRLSETKEITNPYAMRLYESLC <u>Q</u> ER KPDGSGIVSLKIDWIIERYQLPQSYQRMPDFRRRFLQVCVNEINSRT PMRLSYIEKKKGRQTTHIVFSFRDITSMTTG

Hexahistag in blue, mutation underlined.

Table S2.16. CC1⁺¹⁰ (or 12) oligonucleotide sequences

PDB Code / Purpose	CC1 ⁺¹⁰ (or 12) DNA Sequences
<i>dU:dT</i> <u>Uracil underlined</u>	5' - pGCCCCGGTGGACCTGTGACAAATTGCCCTCAG - 3' 3' - GGGCCUCCTGGACACTGTTTAAACGGGAGTCCp - 5'
<i>2-adenine bulge</i> <u>Bulge underlined</u>	5' - pTCCGGGAACCGTCTTGAGGGCAATTTGTCACAG - 3' 3' - GCCC GGCAGAACTCCCGTTAAACAGTGTCTAG - 5'
<i>PT</i> <u>Phosphorothioate underlined</u>	5' - pGATAACCTGTGACAAATTGCCCTCAGCGCGG - 3' 3' - CCTATTGGACACTGTTTAAACGGGAGTCGCGCp - 5'
<i>End poly-C</i> <u>ssDNA tail underlined</u>	5' - pGACTGTGACAAATTGCCCTCAAGACGGCCCCG - 3' 3' - GACACTGTTTAAACGGGAGTCTGCGGGGCCTC4 - 5'
<i>Central poly-T</i> <u>ssDNA tail underlined</u> , <i>third strand italicized</i>	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCCG - 3' 3' - GCCTGGACACTGTTTAAACGGGAGTCGGGGCCp - 5' <u>TTTTTT</u>
<i>Y184E no tail</i>	5' - pGACTGTGACAAATTGCCCTCAAGACGGCCCCG - 3' 3' - GACACTGTTTAAACGGGAGTCTGCGGGGCC - 5'
<i>FAM and mismatch</i> <u>G:G mismatch underlined</u> , <u>fluorescein modification underlined</u> <u>and italicized</u>	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCCG - 3' 3' - GGGC <i>T</i> GGACACTGTTTAAACGGGAGTCGGGGCCp - 5'
<i>TAMRA tagged</i> <u>TAMRA modification underlined</u> <u>and italicized</u>	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCCG - 3' 3' - GGCC <i>T</i> GGACACTGTTTAAACGGGAGTCGGGGCCp - 5'
<i>End poly-T</i> <u>ssDNA tail underlined</u>	5' - pGACTGTGACAAATTGCCCTCAAGACGGCCCCG - 3' 3' - GACACTGTTTAAACGGGAGTCTGCGGGGCCTTTTTT - 5'
<i>Central poly-C</i> <u>ssDNA tail underlined</u> , <i>third strand italicized</i>	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCCG - 3' 3' - GCCTGGACACTGTTTAAACGGGAGTCGGGGCCp - 5' <u>C4</u>
<i>RNA:DNA hybrid</i> <u>RNA underlined</u> , <i>third strand italicized</i>	5' - pGAAGAGAAGCACCTGTGACAAATTGCCCTCAG 3' 3' - TTCTCTTCGTGGGACACTGTTTAAACGGGAGTCC 5'

Table S2.17. Sequences that did not crystallize

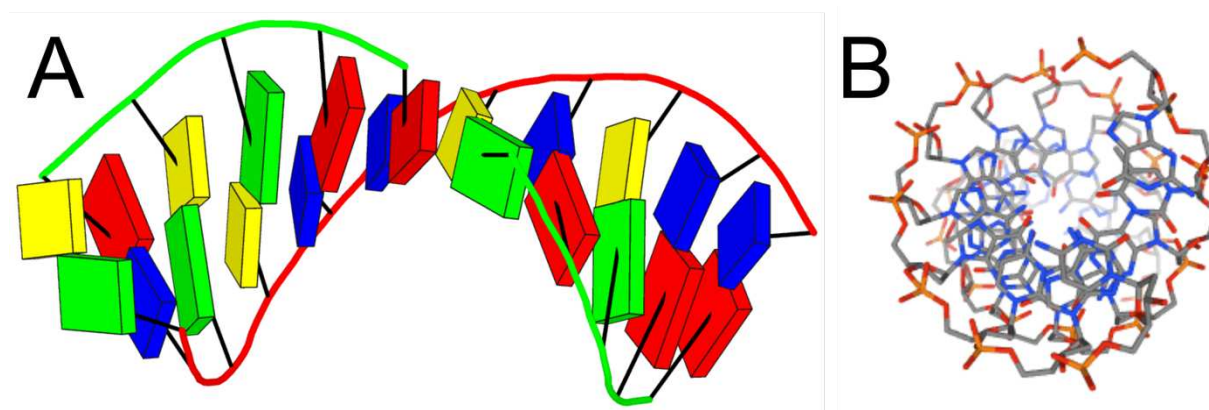
Crystal type	CC1 ⁺¹⁰ (or 12) DNA Sequences
<i>9 nt 1 sb RNA:DNA hybrid</i> <i><u>RNA underlined, third strand</u></i> <i><u>italicized</u></i>	5' - p <u>GCCGGCGGAC</u> CTGTGACAAATTGCCCTCAG - 3' 3' - GGCCGCCTGGACACTGTTTAACGGGAGTCCp - 5'
<i>10 nt 1 sb RNA:DNA hybrid</i> <i><u>RNA underlined, third strand</u></i> <i><u>italicized</u></i>	5' - p <u>GAAGAGAAGCC</u> CTGTGACAAATTGCCCTCAG - 3' 3' - TTCTCTTCGGGACACTGTTTAACGGGAGTCCp - 5'
<i>11 nt 1 sb RNA:DNA hybrid</i> <i><u>RNA underlined, third strand</u></i> <i><u>italicized</u></i>	5' - p <u>GAAGAGAAGCAC</u> CTGTGACAAATTGCCCTCAG - 3' 3' - TTCTCTTCGTGGACACTGTTTAACGGGAGTCCp - 5'
<i>12 nt 2 sb RNA:DNA hybrid</i> <i><u>RNA underlined, third strand</u></i> <i><u>italicized</u></i>	5' - p <u>GAAGAGAAGCAC</u> CCCTGTGACAAATTGCCCTCAG - 3' 3' - TCTCTTCGTGGGACACTGTTTAACGGGAGTCCT p - 5'
<i>13 nt 1 sb RNA:DNA hybrid</i> <i><u>RNA underlined, third strand</u></i> <i><u>italicized</u></i>	5' - p <u>GAAGAGAAGCAG</u> GCCTGTGACAAATTGCCCTCAG - 3' 3' - TTCTCTTCGTCCGGACACTGTTTAACGGGAGTCCp - 5'
<i>Altered RepE binding sequence</i> Sequence change <u>underlined</u>	5' - pGCCGGACCTGTGACT <u>TAATAG</u> CCCTCAGCCCG - 3' 3' - GGCCTGGACACTG <u>GATTAT</u> CGGGAGTCGGGCCp - 5'

Table S2.18. x3DNA analysis against average CC1⁺¹⁰

We used the program x3DNA to extract DNA geometry parameters including rise and twist. For our four new cases here, all had 1-sticky base overhangs, and thus there are two base-pair steps that are interfacial. For reference, see **A.** below for parameters (*e.g.* rise in magenta, and twist in blue) obtained from 8 CC1⁺¹⁰ models (PDB 9YZA, 9NCR, 9NB0, 9YZC, 9YZD, 9YZE, 9YZF, and 9YZG) that consist entirely of canonical base-pairs. twist and rise parameters. We had noted a tendency to slightly overwind DNA at the interface since the twist exceeds the average value for baseline B-DNA (SF 19).

Crystal type	Resolution (Å)	Average rise, non-interface (Å)	Rise sticky interface 1 (Å)	Rise sticky interface 2 (Å)	Average twist non-interface (°)	Twist sticky interface 1 (°)	Twist sticky interface 2 (°)
dU:dT	3.21	3.3 ± 0.1	3.3	3.3	34.6 ± 4.2	37.9	44.5
G:G	3.25	3.3 ± 0.4	3.32	3.1	35.3 ± 6.5	39.51	33.7
Bulge	4.4	3.2 ± 0.3	3.9	3.5	34.2 ± 4.5	29	43.9
PT	3.14	3.3 ± 0.2	3.3	3.8	34.6 ± 5	49	40.6

Table S2. 19: x3DNA analysis of CC1⁺¹² against pure DNA CC1⁺¹⁰ strut and RNA:DNA hybrid 479D. The 11 base pairs (A.) in the non-junction region of the RNA:DNA helix (B.) were analyzed by x3DNA and compared with B-form CC1 strut and RNA:DNA hybrid 479D values^{53,73}. The sticky base on the RNA strand and the preceding RNA:DNA step showed substantial disorder and were excluded from the calculations. Because this structure was solved at 3.66 Å resolution, only robust x3DNA parameters, including rise, twist, slide, and x-displacement, were used for comparison to avoid overinterpretation of the model. The CC1⁺¹² strut showed mixed characteristics between the two reference structures: the average rise was more consistent with CC1⁺¹⁰ geometry, whereas the twist, x-displacement and slide were more consistent with 479D geometry.



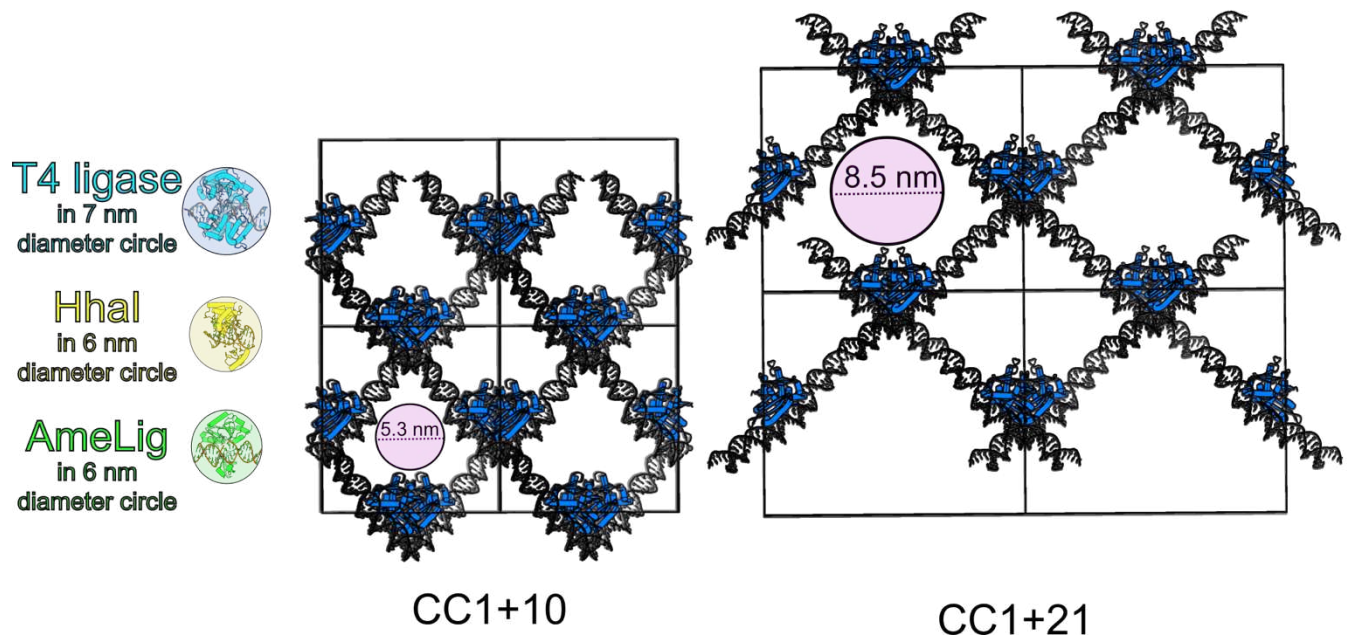
	Rise (Å)	Helical Twist (°)	Slide (Å)	X-displacement (Å)
479D	3.34 ± 0.17	32.41 ± 5.57	-1.43 ± 0.27	-3.80 ± 0.73
9Z55	3.28 ± 0.14	35.34 ± 5.32	0.48 ± 1.00	0.31 ± 1.88
11 bp RNA:DNA	3.25 ± 0.17	31.05 ± 4.56	-1.36 ± 0.53	-2.96 ± 1.36

Table S2.20. The five lowest-RMSD conformers from the 46-member AIMNET2-ranked ensemble relative to the crystallographic fluorescein dimer. Top conformer 44 overlay is shown in SF5.

Rank	Conformer	AIMNET2 Energy (Ha)	RMSD (Å)
1	44	-139.28009	1.89
2	34	-139.28171	1.94
3	45	-139.27968	2.19
4	16	-139.28469	2.29
5	39	-139.28145	2.38

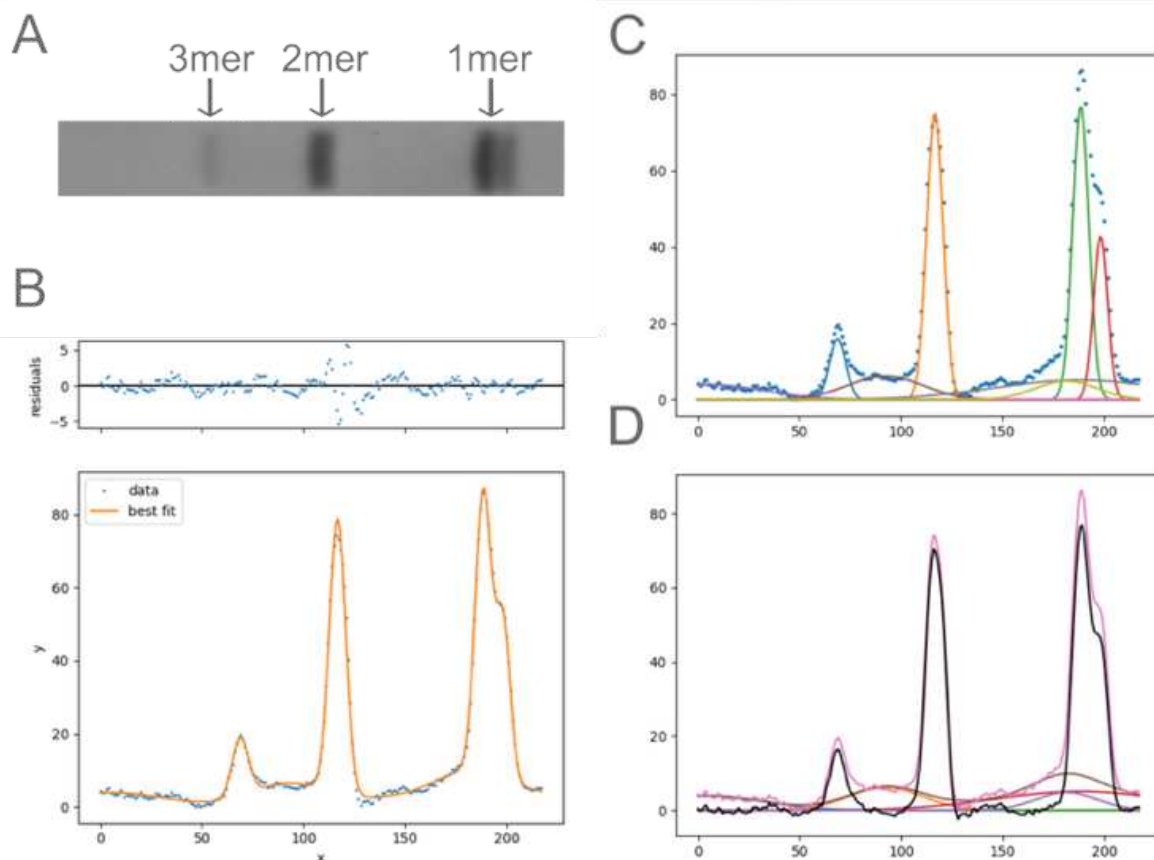
APPENDIX II. Supplemental Information for Chapter 3: Enzymatic and Chemical Ligation of the CC1 Platform

- Supplementary Figure 3.1:** Enzyme size comparison to CC1+10 and CC1+21 pores
- Supplementary Figure 3.2:** Densitometry fit for 0 SB 5'P in 2 doses 30 mg/mL EDC
- Supplementary Figure 3.3:** Densitometry fit for 1 SB 5'P in 2 doses 30 mg/mL EDC
- Supplementary Figure 3.4:** Densitometry fit for 1 SB 3'P in 2 doses 30 mg/mL EDC
- Supplementary Figure 3.5:** Densitometry fit for 2 SB 3'P in 2 doses 30 mg/mL EDC
- Supplementary Figure 3.6:** Densitometry fit for EDC Buffer 1
- Supplementary Figure 3.7:** Densitometry fit for EDC Buffer 2
- Supplementary Figure 3.8:** Densitometry fit for EDC Buffer 3
- Supplementary Figure 3.9:** Densitometry fit for EDC Buffer 4
- Supplementary Figure 3.10:** AmeLig expression SDS-PAGE
- Supplementary Figure 3.11:** Splint ligation oligos
- Supplementary Figure 3.12:** GuestScan registers for AmeLig in CC1+10
- Supplementary Figure 3.13:** GuestScan registers for AmeLig in CC1+21
- Supplementary Figure 3.14:** GuestScan registers for HhaI in CC1+10
- Supplementary Figure 3.15:** GuestScan registers for HhaI in CC1+21
- Supplementary Figure 3.16:** Densitometry fit for CC1+21 with EDC
- Supplementary Figure 3.17:** Densitometry fit for CC1+21 with AmeLig
- Supplementary Figure 3.18:** Trace labelled guest soaking with AmeLig and HhaI
- Supplementary Figure 3.19:** AmeLig ligation and HhaI digestion of free solution building blocks
- Supplementary Figure 3.20:** 9YZK ligation
- Supplementary Table 3.1:** Sequences of crystals used in ligation studies
- Supplementary Table 3.2:** XRD statistics of CC1+10 with HhaI binding site
- Supplementary Table 3.3:** Crystallization conditions
- Supplementary Table 3.4:** Protein sequences



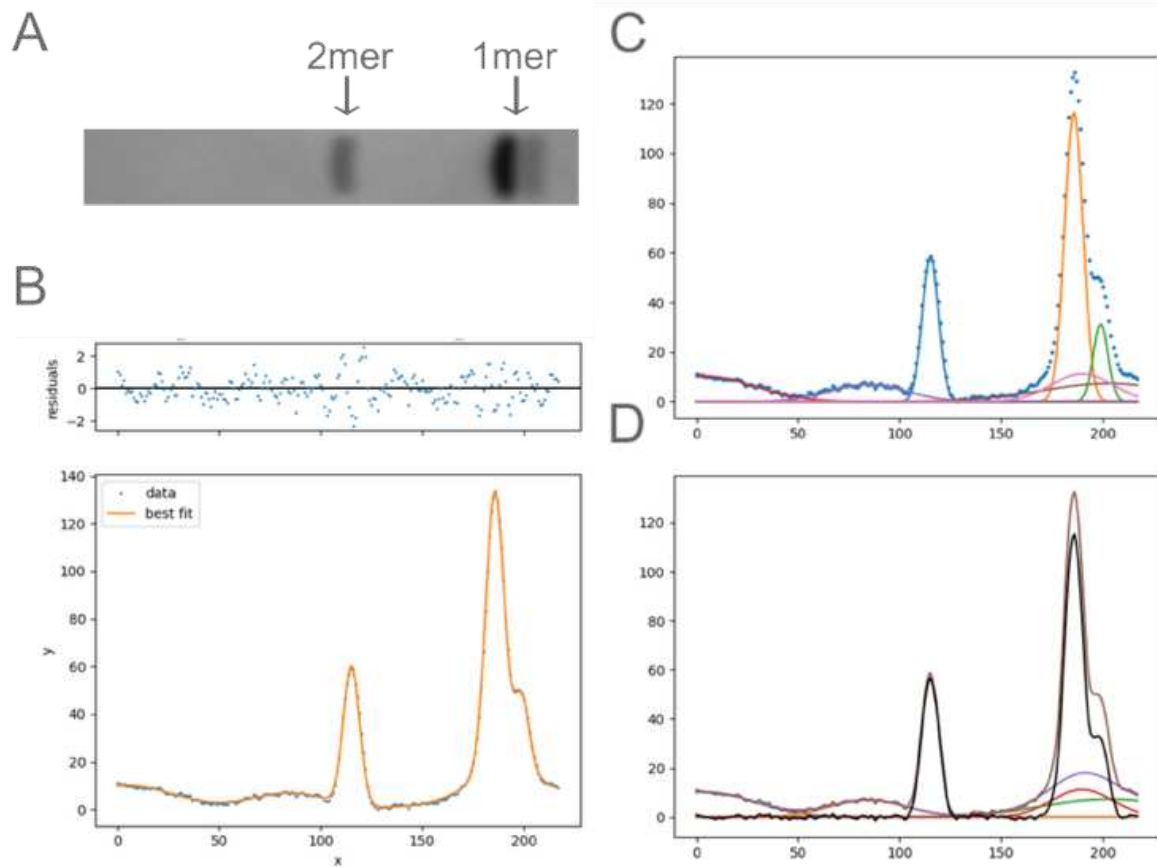
Supplementary Figure 3.1: Enzyme size comparison to CC1+10 and CC1+21 pores. The solvent-accessible channel in CC1+10 and CC1+21 were estimated MAP_CHANNELS, giving an apparent diameter of ~5.3 nm and ~8.5 nm (15). Reference circles representing T4 DNA ligase, HhaI, and AmeLig are shown with approximate diameters of 7, 6, and 6 nm, respectively. Although these diameters are larger than the CC1+10 pore circle, these values likely overestimate the steric barrier to diffusion because the enzymes are not rigid spheres. In solution, enzymes can sample flexible conformations and, when not bound to DNA, may present as smaller, increasing the likelihood that they can diffuse through the CC1+10 solvent channels. All three enzymes fit well within the CC1+21 estimated pore size.

CC1+10 0 sb 5' P 2 doses 30mg/mL EDC Lane 4



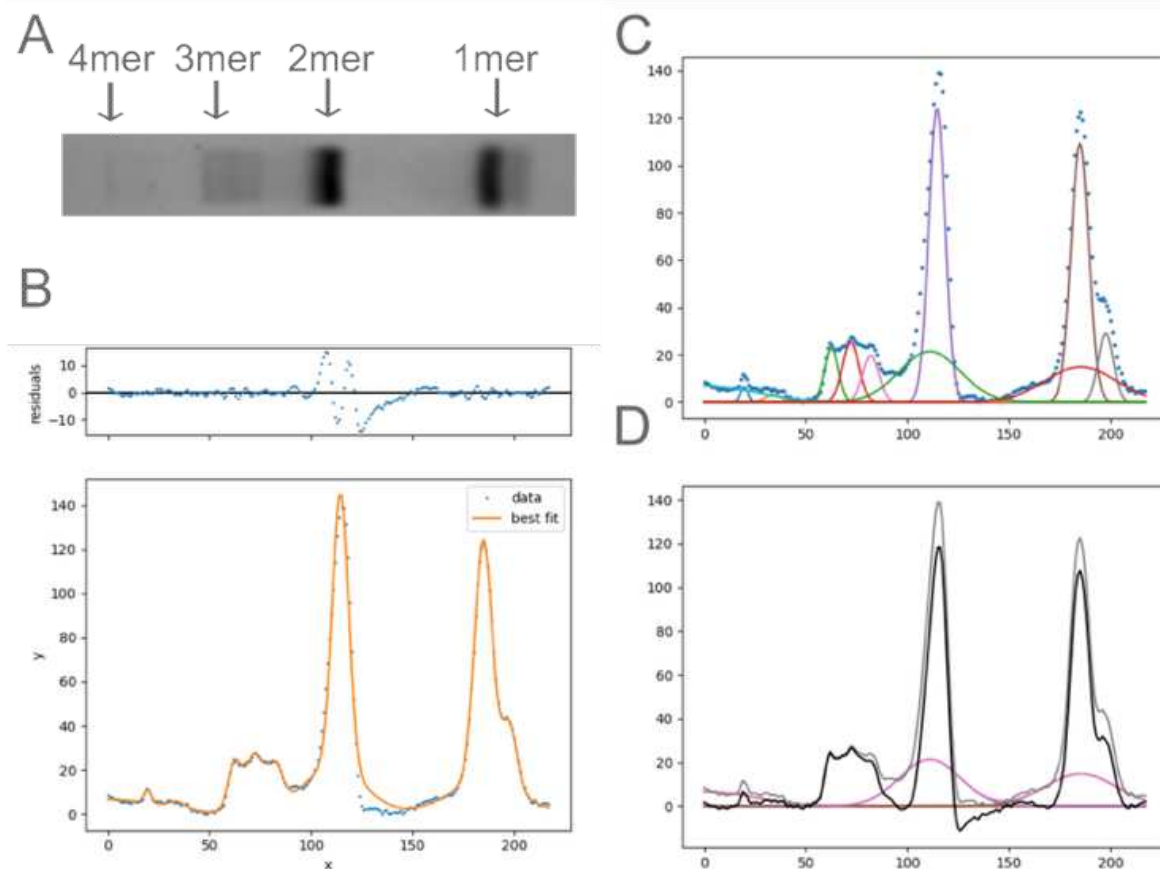
Supplementary Figure 3.2: Densitometry results and annotation for CC1+10 with 0 SB and 5'P. For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

CC1+10 1 sb 5' P 2 doses 30mg/mL EDC Lane 5



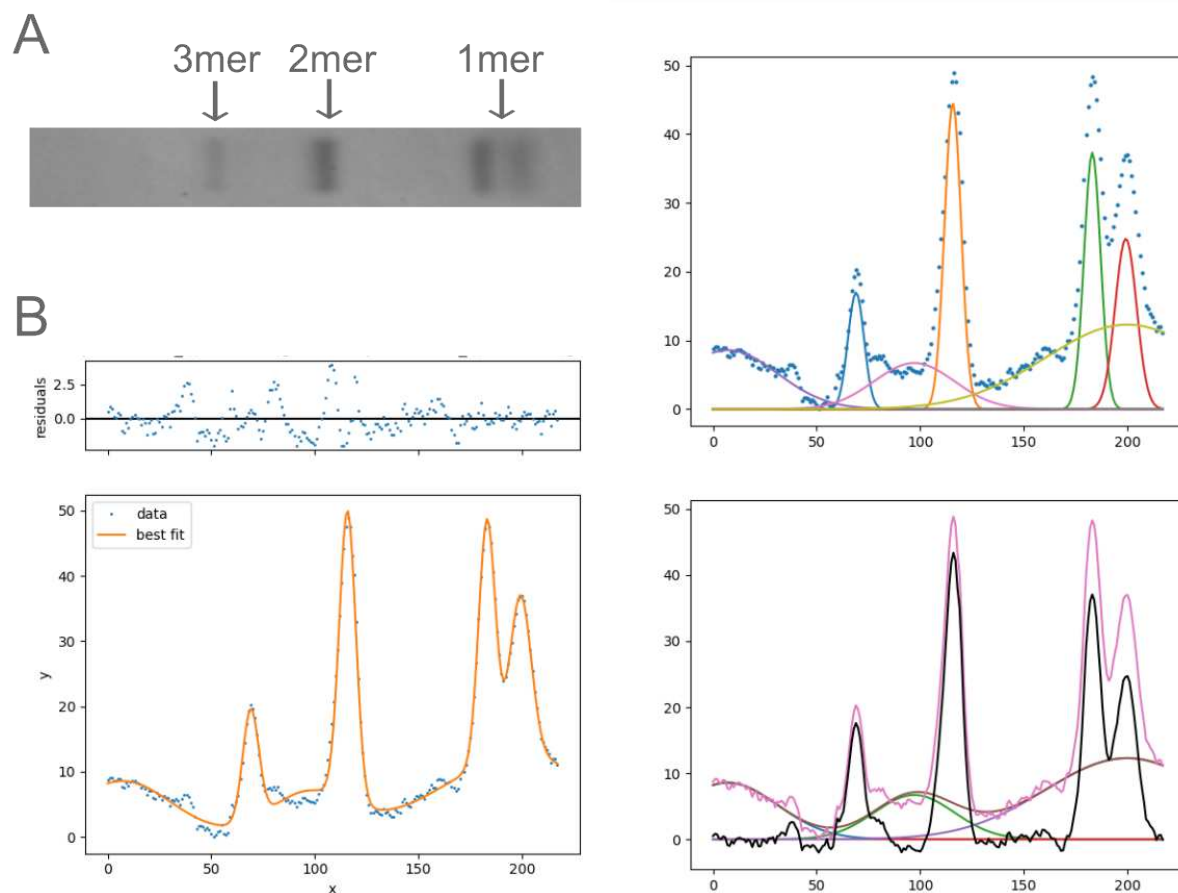
Supplementary Figure 3.3: Densitometry results and annotation for CC1+10 with 1 SB and 5'P. For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

CC1+10 1 sb 3' P 2 doses 30mg/mL EDC Lane 6



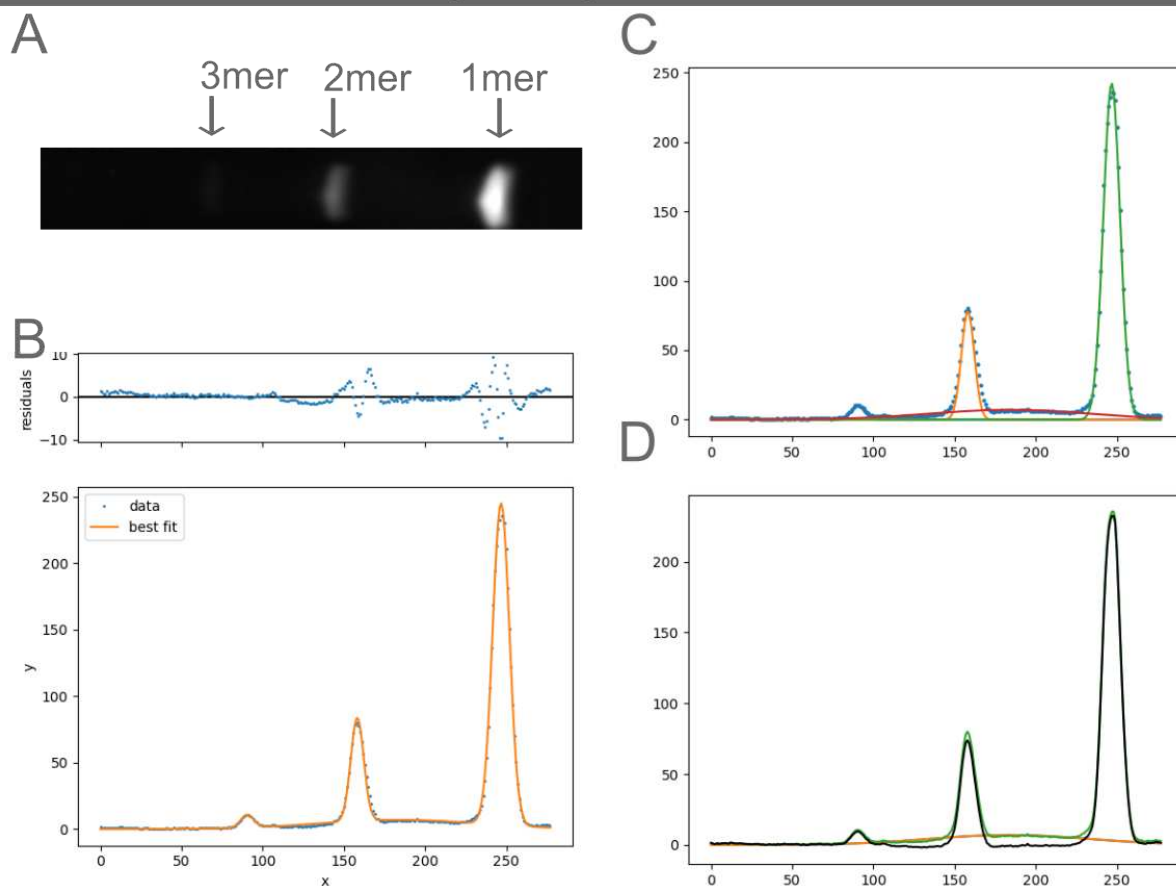
Supplementary Figure 3.4: Densitometry results and annotation for CC1+10 with 1 SB and 3'P. For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

CC1+10 2 sb 3' P 2 doses 30mg/mL EDC Lane 7



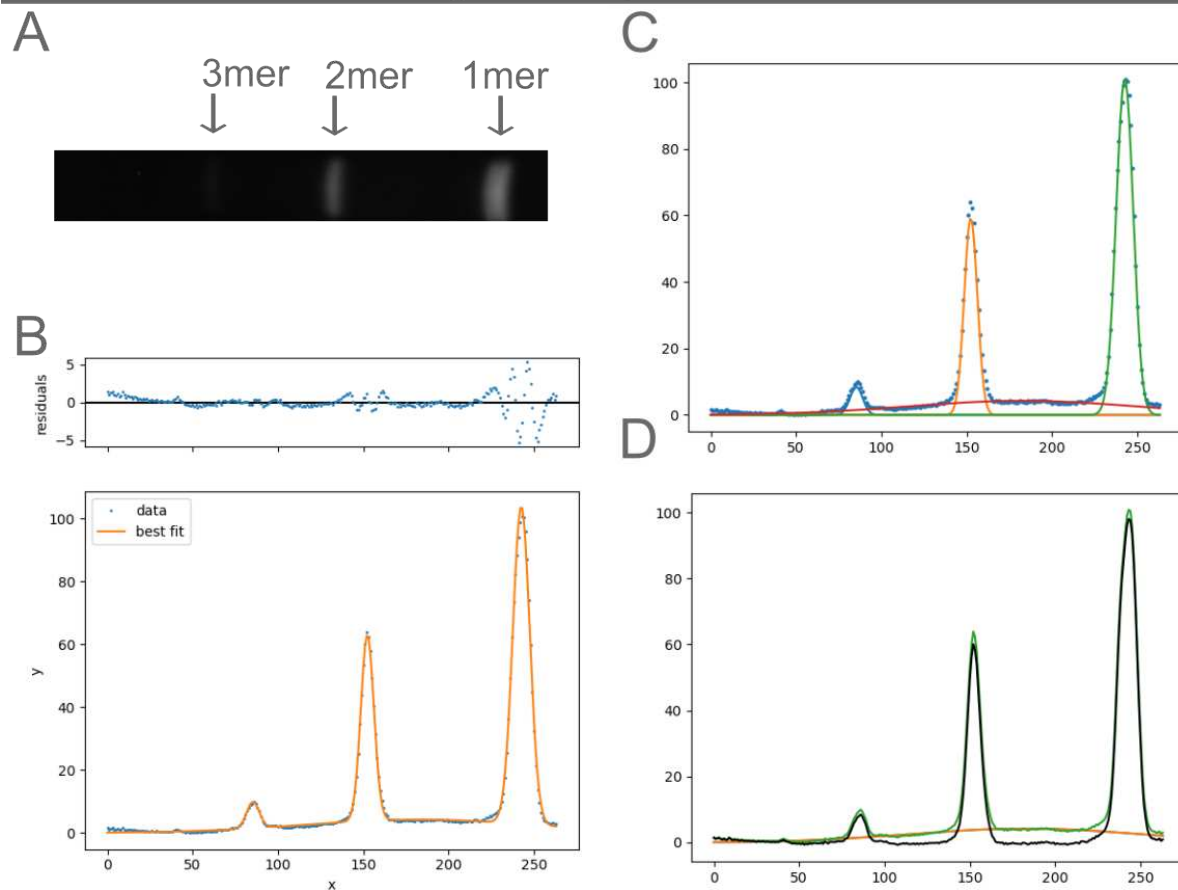
Supplementary Figure 3.5: Densitometry results and annotation for CC1+10 with 2 SB and 3'P. For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

CC1+10 Buffer 1 (NaCl)



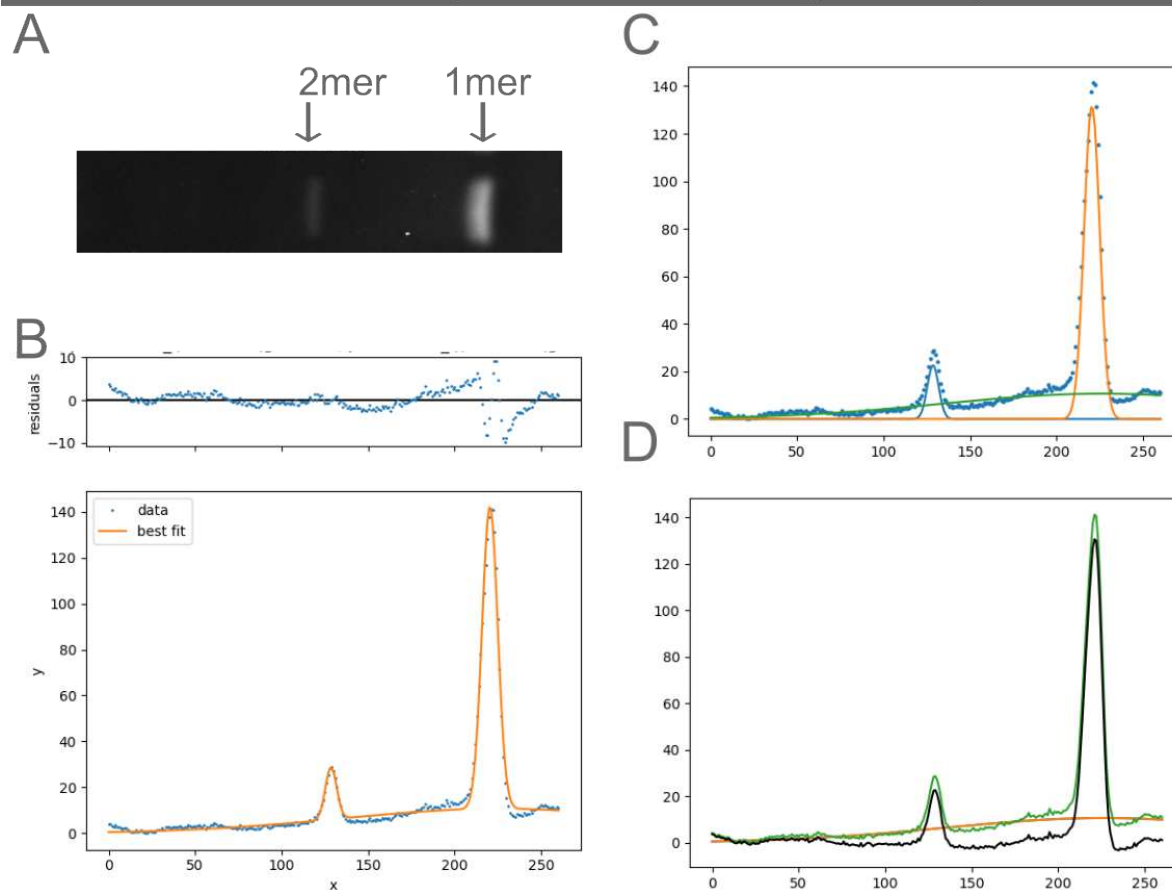
Supplementary Figure 3.6: Densitometry results and annotation for CC1+10 in crosslinking buffer 1 (1 M Li_2SO_4 , 50 mM NaCl, 50 mM MES pH 6.0). For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

CC1+10 Buffer 2 (MnCl₂)



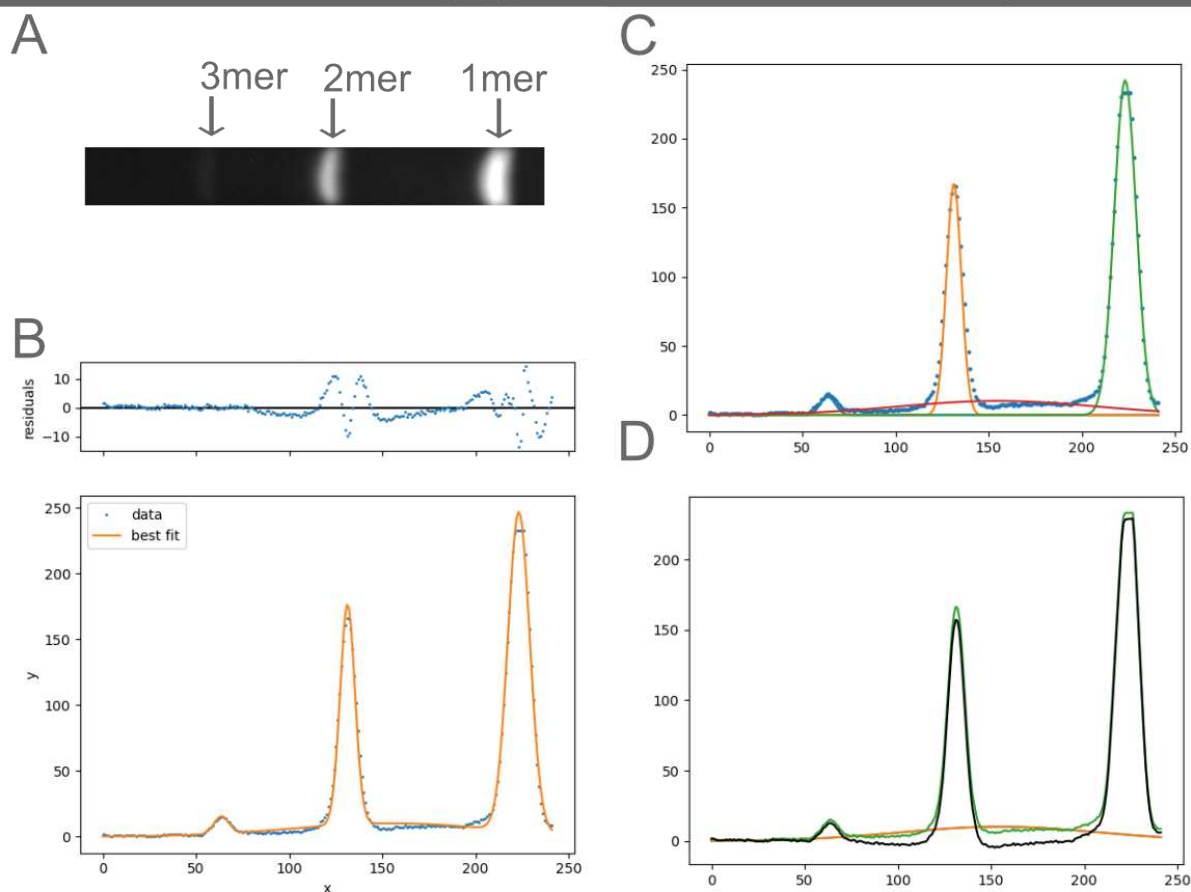
Supplementary Figure 3.7: Densitometry results and annotation for CC1+10 in crosslinking buffer 2 (1 M Li₂SO₄, 50 mM MnCl₂, 50 mM MES pH 6.0). For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

CC1+10 Buffer 3 (100 mM MES pH 6.0)

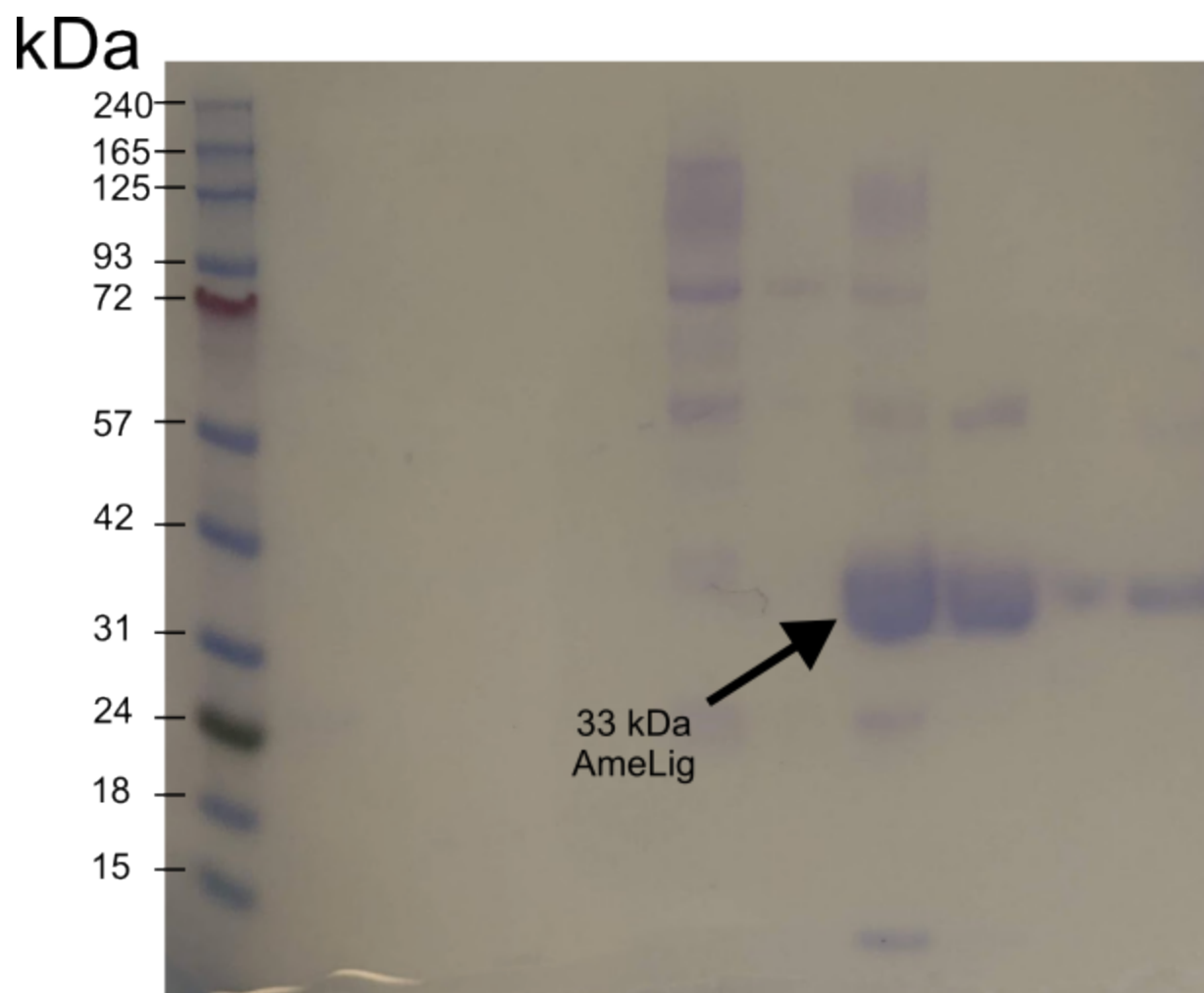


Supplementary Figure 3.8: Densitometry results and annotation for CC1+10 in crosslinking buffer 3 (1 M Li₂SO₄, 50 mM Mg(OAc)₂, 100 mM MES pH 6.0). For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

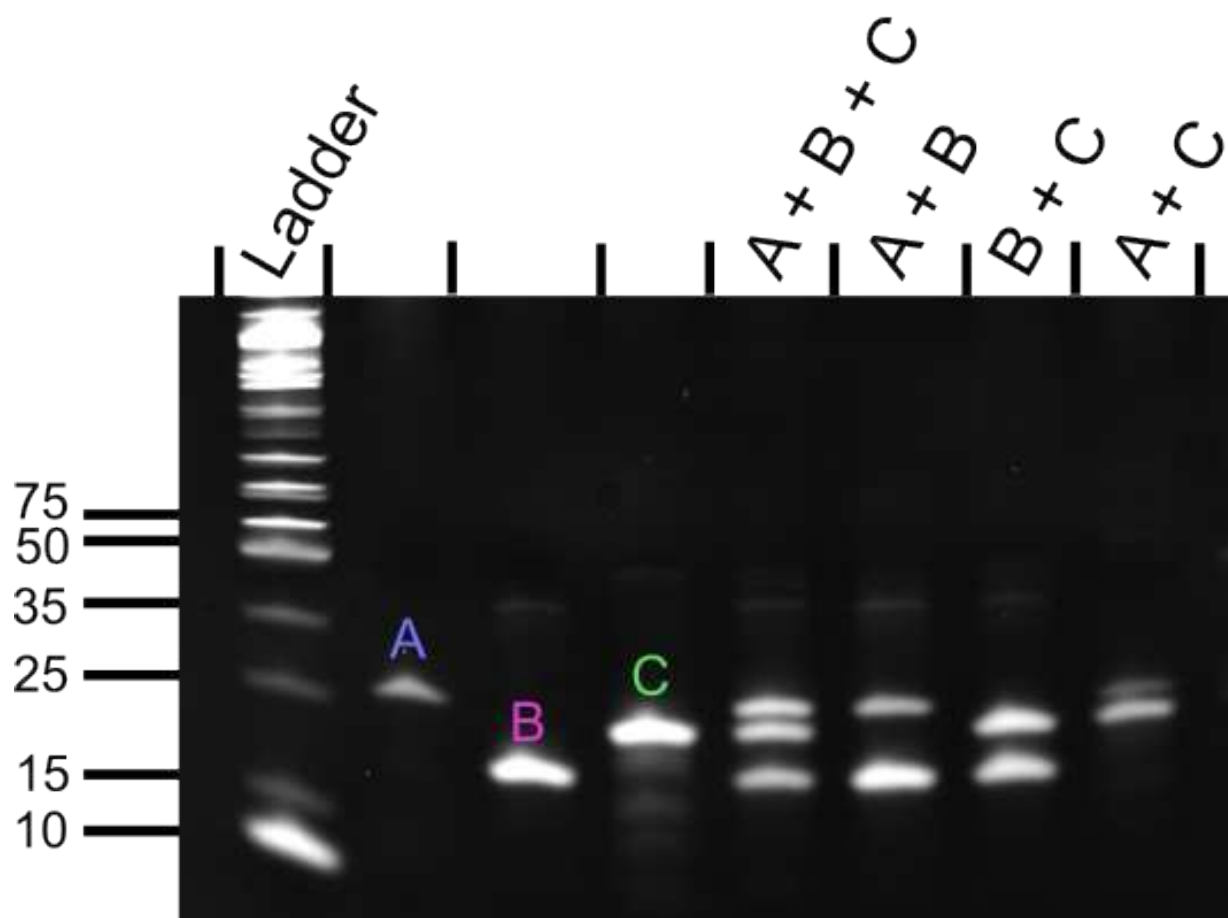
CC1+10 Buffer 4 (typical crosslink buffer)



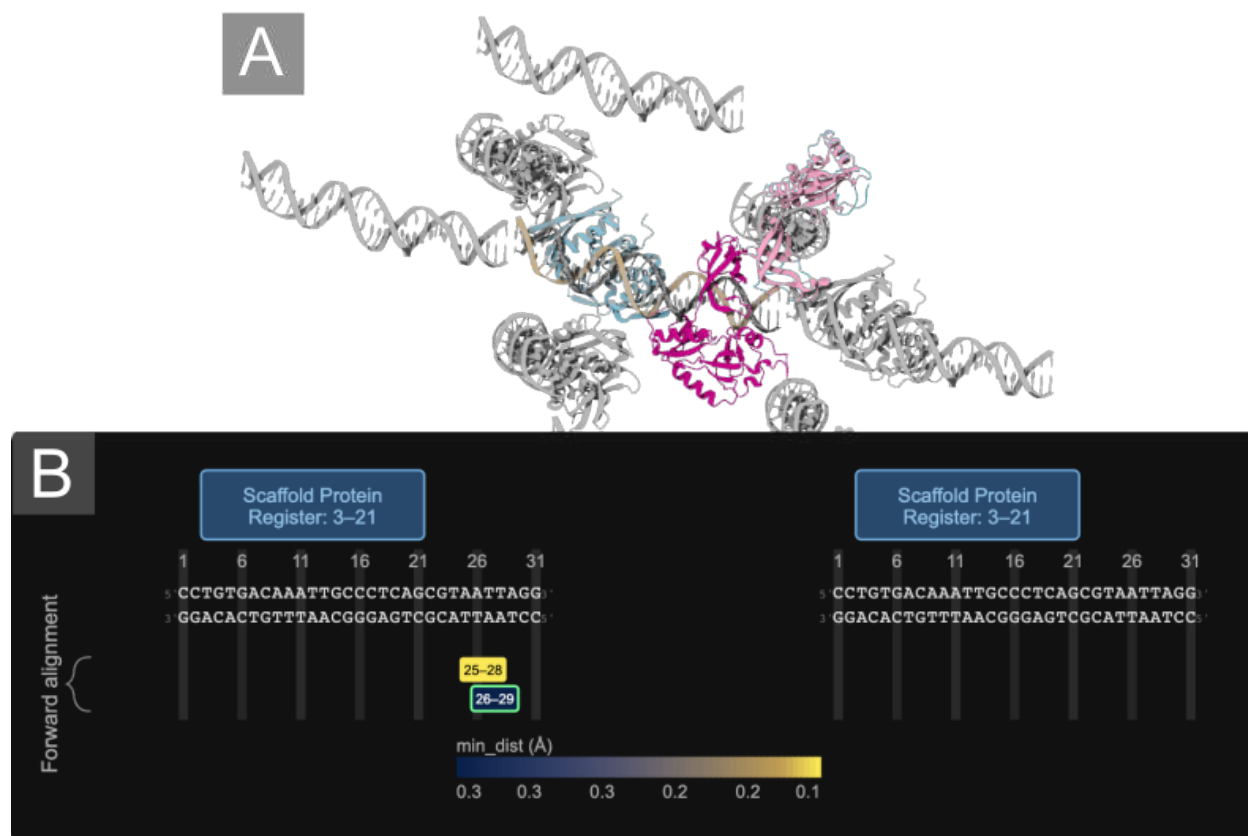
Supplementary Figure 3.9: Densitometry results and annotation for CC1+10 in crosslinking buffer 4 (1 M Li_2SO_4 , 50 mM $\text{Mg}(\text{OAc})_2$, 50 mM MES pH 6.0). For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.



Supplementary Figure S3.10: SDS-PAGE Heparin column purification of AmeLig expression. Lane 1 contains the protein molecular weight ladder, with marker sizes shown on the reference ladder at left. The remaining lanes show sequential heparin column elution fractions. A prominent band corresponding to the expected molecular weight of AmeLig, approximately 33 kDa, is observed in elution fractions 7–10, indicating enrichment of the target protein in these fractions.

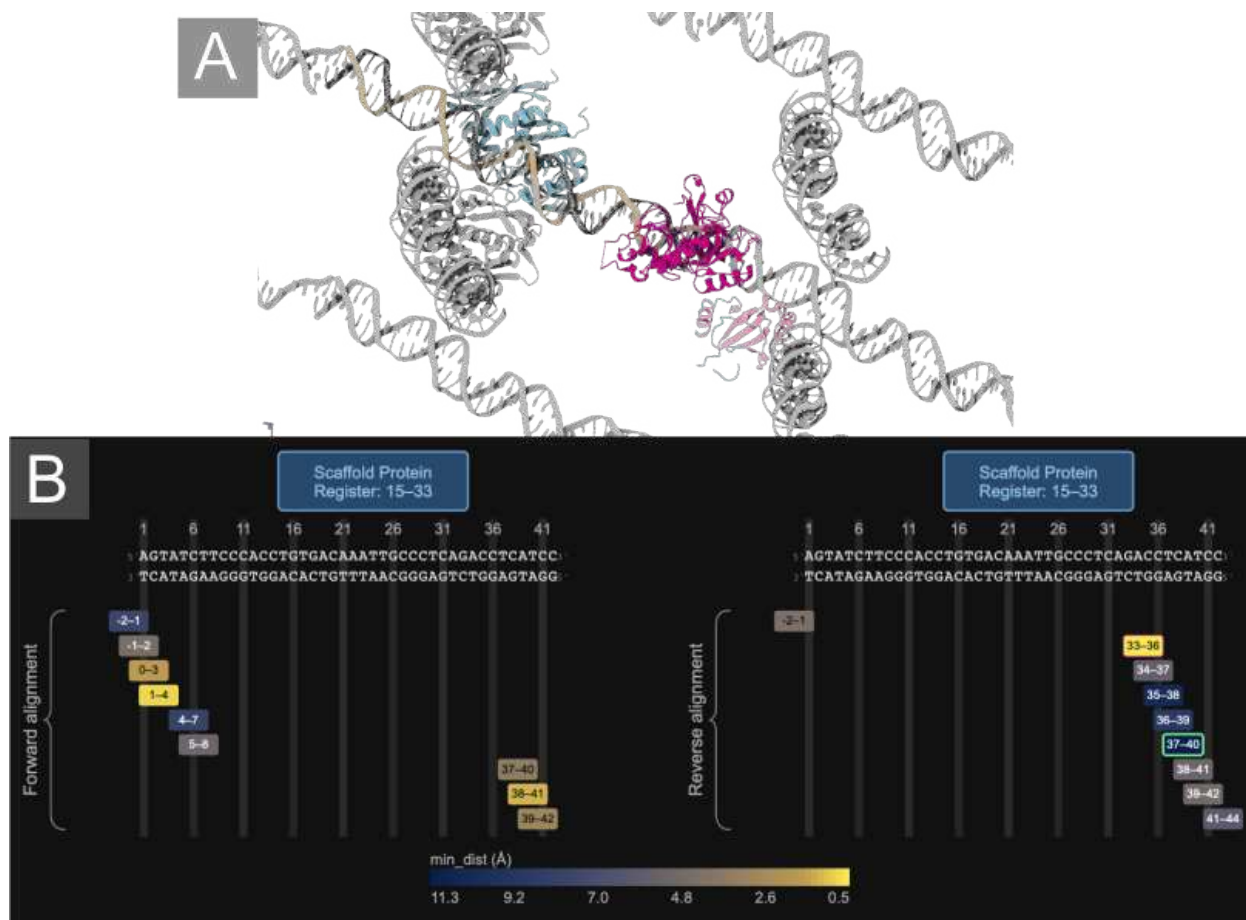


Supplementary Figure 3.11: Splint ligation oligos. First lane is a molecular weight ladder with values shown in nucleotides. The first three lanes are single oligos, A, B, C. The remaining lanes are combinations of oligos as noted above the lane.



Supplementary Figure 3.12: GuestScan pipeline results of AmeLig (pdb 6GDR) in the CC1+10 lattice. GuestScan is an open-source Python code to identify and classify candidate installation sites for DNA-binding domains (DBD) within the CC1+10 and CC1+21 scaffold crystals (<https://github.com/cdsnow/DBDscan>). A web interface is also available (<https://www.engr.colostate.edu/~cdasnow/guestscan.html>).

A. Scaffold protein RepE54 (light blue), Target guest site (magenta), Closest guest symmetry mate (pink).
B. This diagram shows possible binding registers that might be sterically feasible, colored by how close they get to other biomolecules within the lattice (excluding their dsDNA stack). The register with the green boundary is the currently selected model. The candidate registers with magenta boundaries overlap with the scaffold protein (RepE54) binding window that will presumably prevent freely changing the sequence in pursuit of guest installation.



Supplementary Figure 3.13: GuestScan pipeline results of AmeLig (pdb 6GDR) in the CC1+21 lattice. GuestScan is an open-source Python code to identify and classify candidate installation sites for DNA-binding domains (DBD) within the CC1+10 and CC1+21 scaffold crystals (<https://github.com/cdsnow/DBDscan>). A web interface is also available (<https://www.engr.colostate.edu/~cdasnow/guestscan.html>).

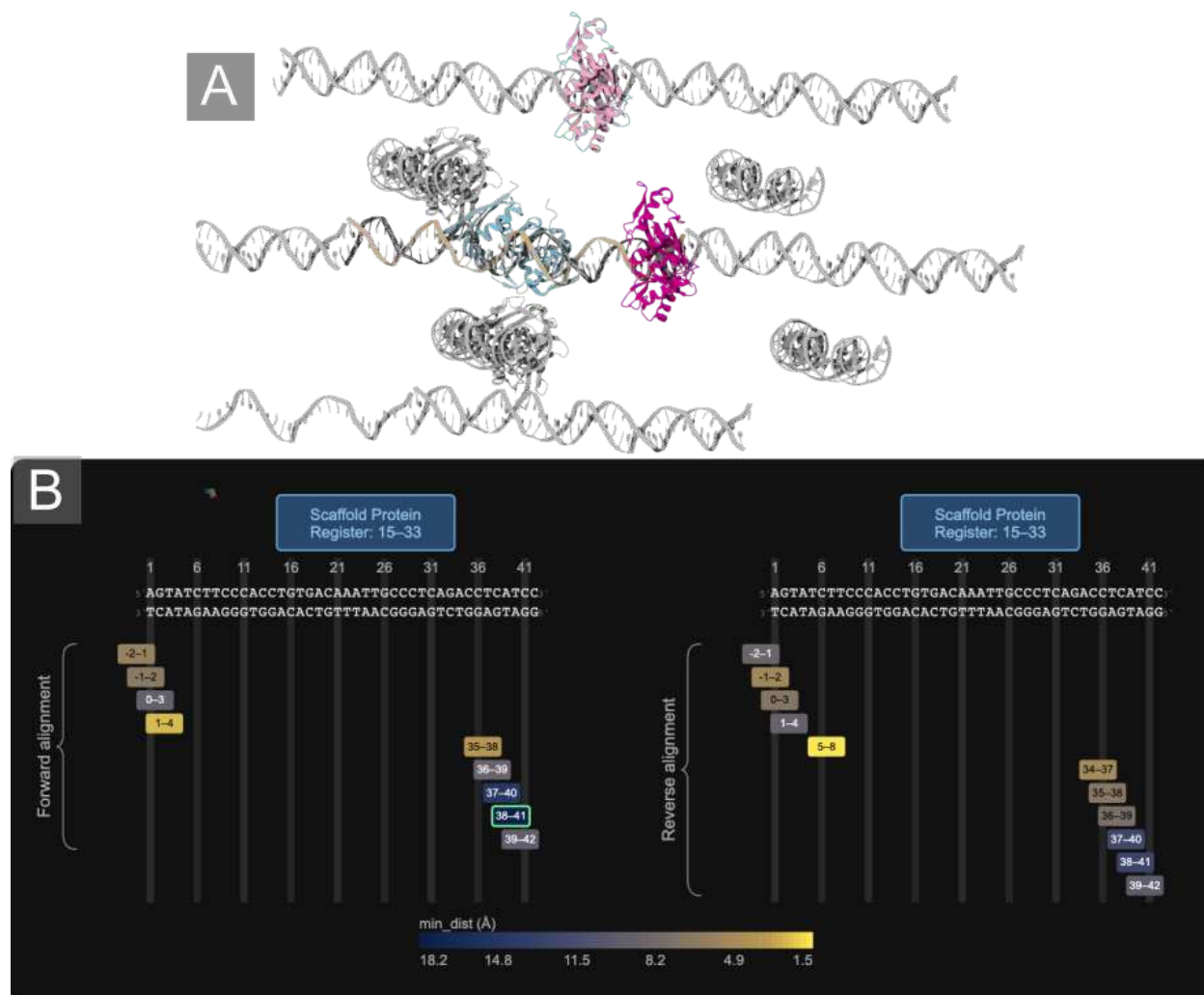
A. Scaffold protein RepE54 (light blue), Target guest site (magenta), Closest guest symmetry mate (pink).
B. This diagram shows possible binding registers that might be sterically feasible, colored by how close they get to other biomolecules within the lattice (excluding their dsDNA stack). The register with the green boundary is the currently selected model. The candidate registers with magenta boundaries overlap with the scaffold protein (RepE54) binding window that will presumably prevent freely changing the sequence in pursuit of guest installation.



Supplementary Figure 3.14: GuestScan pipeline results of HhaI (pdb 6GDR) in the CC1+10 lattice. GuestScan is an open-source Python code to identify and classify candidate installation sites for DNA-binding domains (DBD) within the CC1+10 and CC1+21 scaffold crystals (<https://github.com/cdsnow/DBDscan>). A web interface is also available (<https://www.engr.colostate.edu/~cdasnow/guestscan.html>).

A. Scaffold protein RepE54 (light blue), Target guest site (magenta), Closest guest symmetry mate (pink).

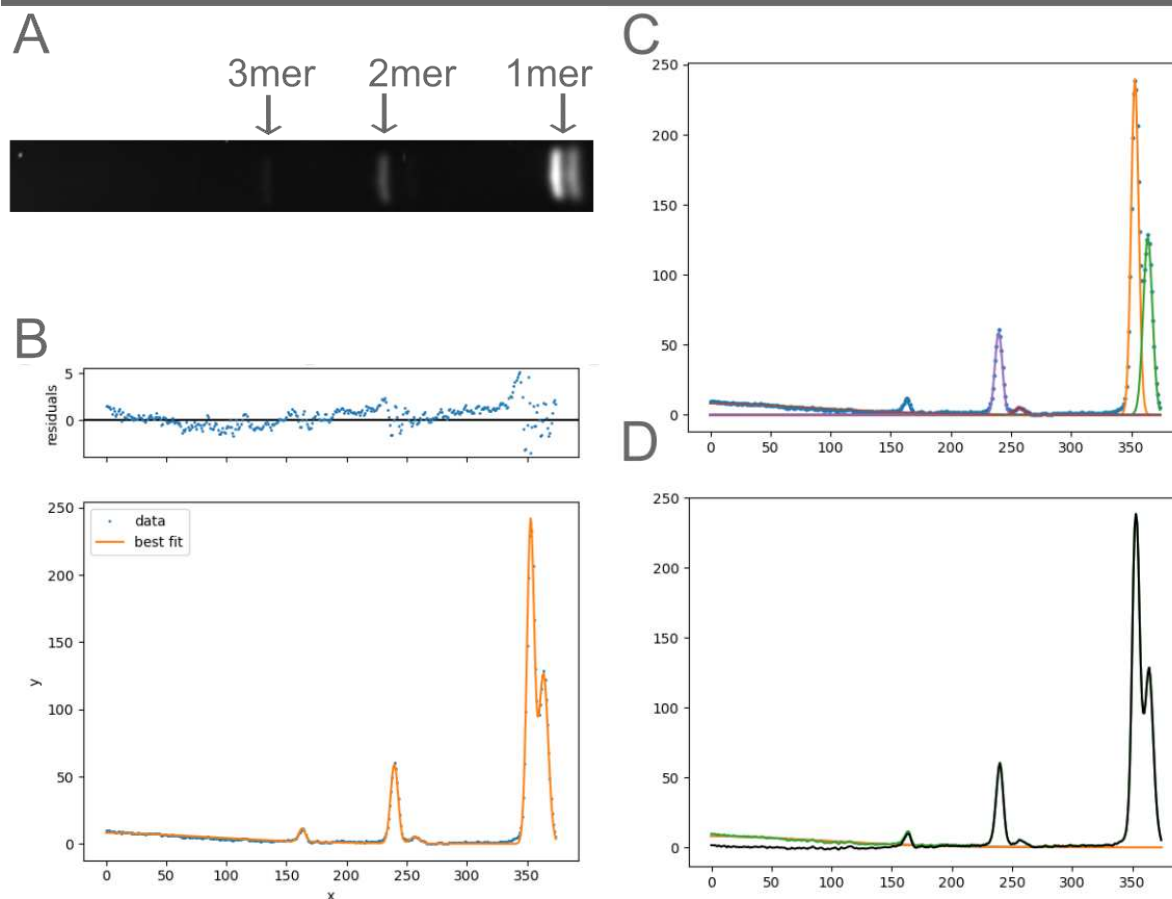
B. GuestScan returned no non-clashing registers. This diagram shows possible binding registers that might be sterically feasible, colored by how close they get to other biomolecules within the lattice (excluding their dsDNA stack). The register with the green boundary is the currently selected model. The candidate registers with magenta boundaries overlap with the scaffold protein (RepE54) binding window that will presumably prevent freely changing the sequence in pursuit of guest installation.



Supplementary Figure 3.15: GuestScan pipeline results of HhaI (pdb 6GDR) in the CC1+21 lattice. GuestScan is an open-source Python code to identify and classify candidate installation sites for DNA-binding domains (DBD) within the CC1+10 and CC1+21 scaffold crystals (<https://github.com/cdsnow/DBDscan>). A web interface is also available (<https://www.engr.colostate.edu/~cdasnow/guestscan.html>).

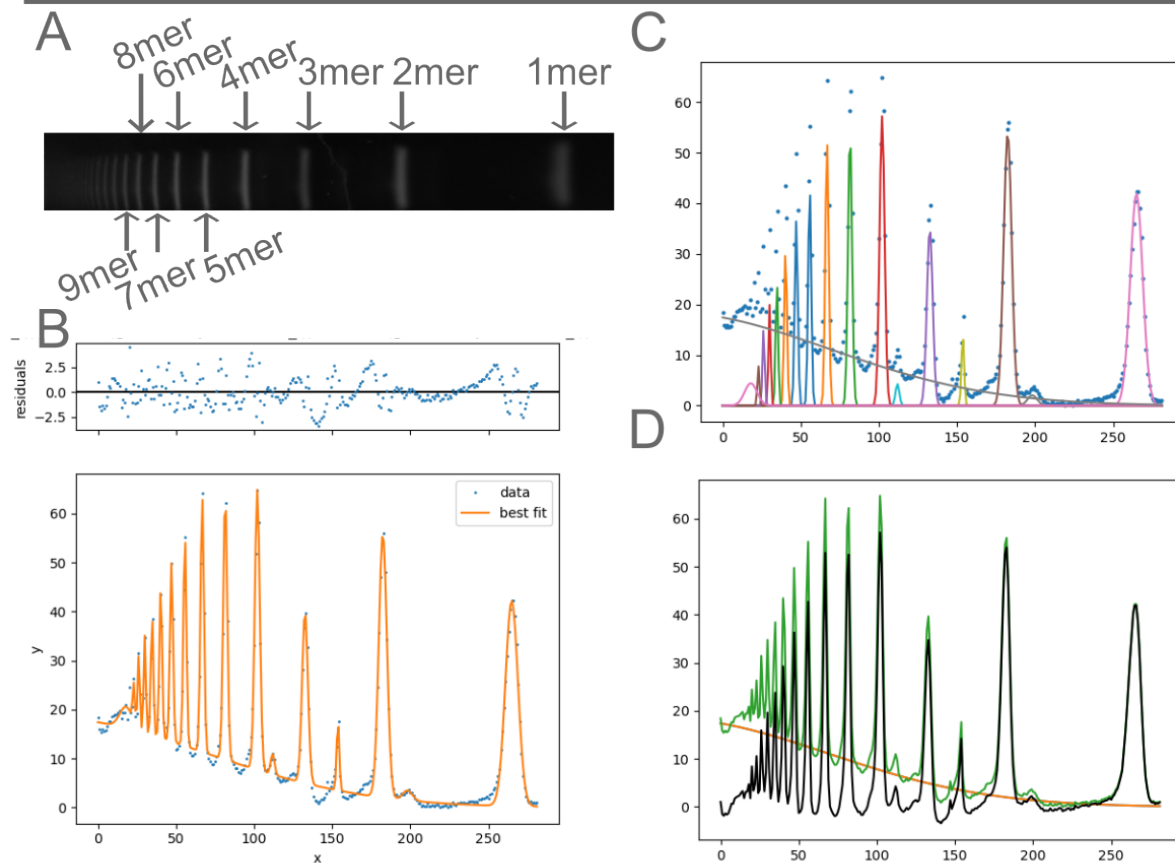
A. Scaffold protein RepE54 (light blue), Target guest site (magenta), Closest guest symmetry mate (pink). **B.** This diagram shows possible binding registers that might be sterically feasible, colored by how close they get to other biomolecules within the lattice (excluding their dsDNA stack). The register with the green boundary is the currently selected model. The candidate registers with magenta boundaries overlap with the scaffold protein (RepE54) binding window that will presumably prevent freely changing the sequence in pursuit of guest installation.

CC1+21 2 sb 5' P EDC treated

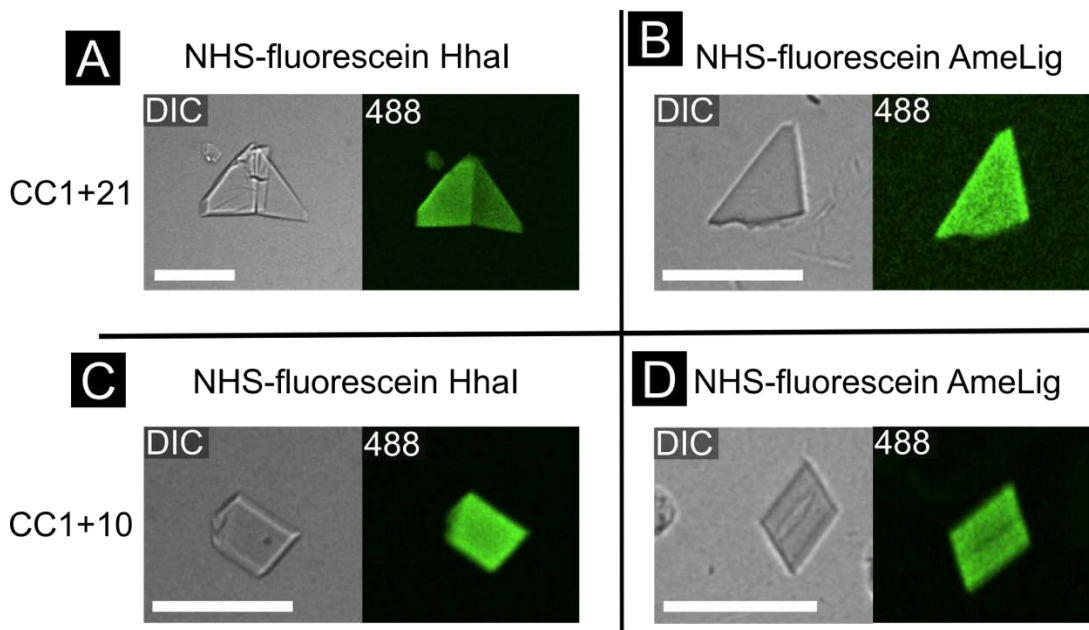


Supplementary Figure 3.16: Densitometry results and annotation for CC1+21 with 2 SB and 5'P after EDC ligation. For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

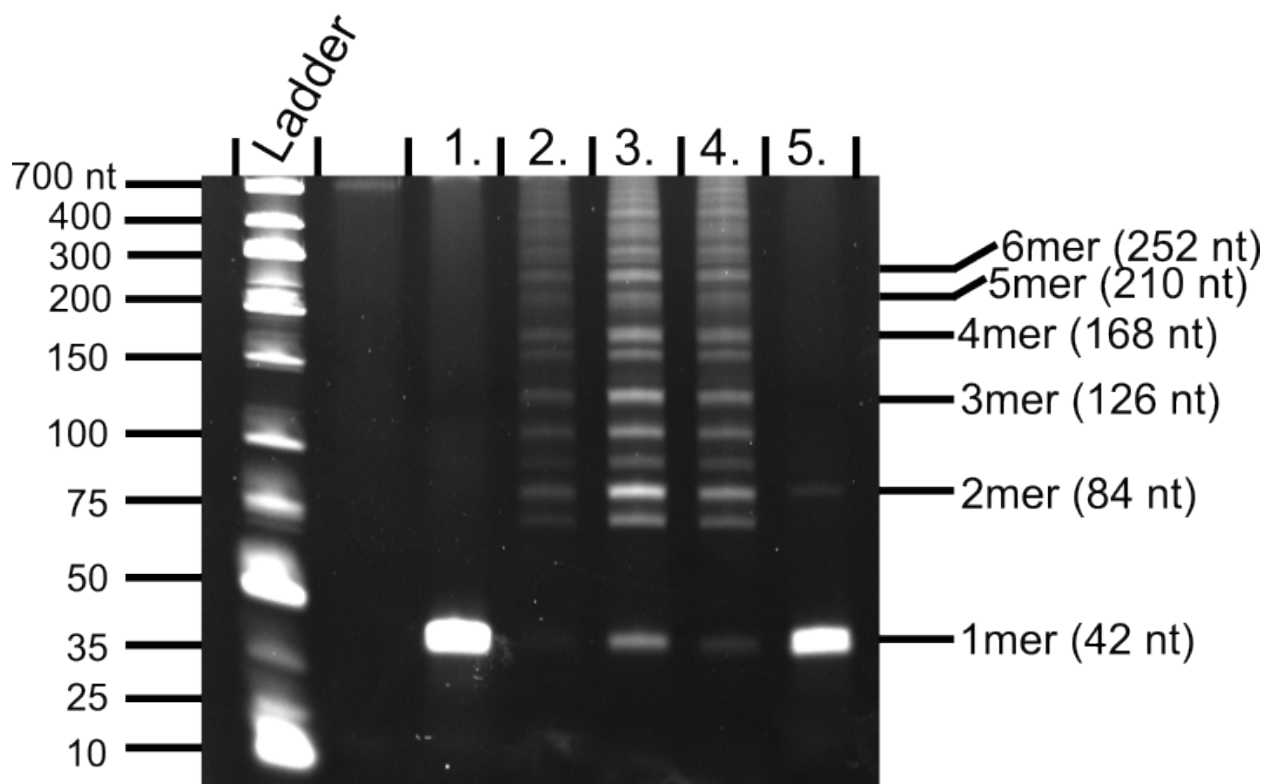
CC1+21 2 sb 5' P AmeLig treated



Supplementary Figure 3.17: Densitometry results and annotation for CC1+21 with 2 SB and 5'P after AmeLig ligation. For each lane, we conducted a multi-step densitometry analysis. All graph scale bars are pixel location from left to right along A (x-axis) and relative intensity (y-axis). **A.** The lane was manually excised from the gel using ImageJ software. **B.** The final fit (orange line) closely follows the raw intensity data (blue dots) that is the average intensity value across the lane (perpendicular left to right direction). The quality of the fit is also evident in the small residuals (above). **C.** The component gaussian functions inside the fit are shown with different colors. **D.** If we subtract the diffuse background (orange) from the full fit the adjusted curve is shown in black.

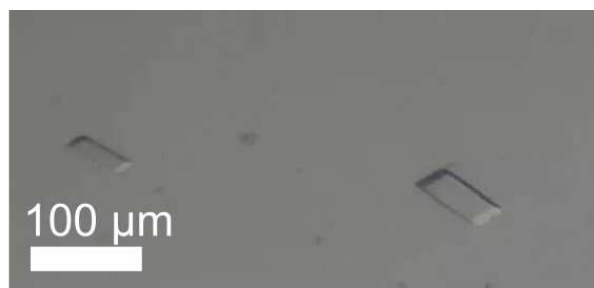


Supplementary Figure 3.18: Confocal loading of enzymes trace-labeled with NHS-fluorescein. Scale bar is 100 μm . Crystals were imaged with differential interference contrast (left panels) and with 488nm excitation laser and 515 nm emission filter after 48 hours of soaking with trace labelled enzymes at concentration 20 μM (with that consisting of 1:20 labelled:unlabelled ratio). **A.** CC1+21 in solution with HhaI. **B.** CC1+21 in solution with AmeLig. **C.** CC1+10 in solution with HhaI. **D.** CC1+10 in solution with AmeLig.

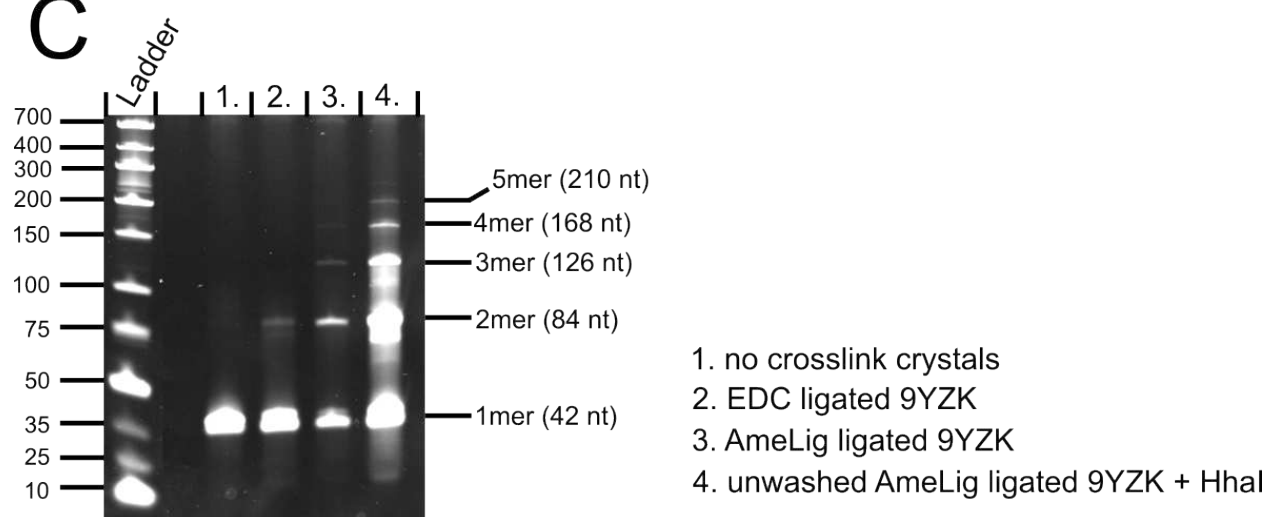


1. no crosslink oligos
2. HhaI + ligated oligos in mother liquor
3. oligos + AmeLig room temp
4. oligos + AmeLig 40 °C
5. oligos + AmeLig 53 °C

Supplementary Figure 3.19: Free solution building block reactions. DNA-only crystal building blocks (dsDNA without RepE54) were added to solution at a final concentration of 12.5 μ M. Lane 1 has oligos with no crosslinking in mother liquor. Lane 2 has DNA-only building blocks ligated by AmeLig that had been incubated with HhaI in mother liquor. Lanes 3-4 are AmeLig reactions on free solution building blocks at increasing temperatures in mother liquor. Temperatures are room temperature, 40 °C, 53 °C, from 3-5.

A**B**

5' -pAGTATCTTCCCAC**CTGTGACAAATTGCCCTCAG**ACCTCATCC-3'
 3' -ATAGAAGGGTG**GACACTGTTTAACGGGAGTCT**GGAGTAGGTCp-5'

C

Supplementary Figure 3.20: 9YZK crystal growth and ligation. **A.** 9YZK crystals in white light microscopy showing typical CC1+21 parallelepiped growth habit. **B.** Crystal DNA oligo sequence, RepE54 binding site is bolded. This sequence lacks the GCG/C cut site recognized by HhaI. **C.** Denaturing PAGE showing dissolved crystals with no crosslink, crystals ligated by EDC, crystals ligated by AmeLig, and unwashed AmeLig-ligated crystals incubated with HhaI, from 1-4. Lane 4 shows higher order species than Lane 3 due to lack of washing of free solution ligation products prior to incubation with HhaI.

Supplementary Table S3.1: CC1+10 and CC1+21 oligonucleotide sequences

DNA sequences used for crystallization. Oligo orientation is noted with either 5' or 3'. RepE54 binding site highlighted in grey.

PDB Code / Purpose	DNA Sequences
1 SB 5'P <i>9NCR, Used in EDC trials</i>	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCG - 3' 3' - GGCCTGGACACTGTTTAACGGGAGTCTGGGCCp - 5'
0 SB 5'P <i>9YZJ Used in EDC trials</i>	5' - pCCTGTGACAAATTGCCCTCATAATTAGGCCG - 3' 3' - GGACACTGTTTAACGGGAGTATTAATCCGGCp - 5'
1 SB 3'P <i>Used in EDC trials</i>	5' - CCCGGACCTGTGACAAATTGCCCTCAGACGGp - 3' 3' - pGGCCTGGACACTGTTTAACGGGAGTCTGCCG - 5'
2 SB 3'P <i>Used in EDC trials</i>	5' - CCCGGACCTGTGACAAATTGCCCTCAGACGGp - 3' 3' - pGCCTGGACACTGTTTAACGGGAGTCTGCCGG - 5'
1 SB 5'P <i>HhaI cut sites</i>	5' - pGCGCGACCTGTGACAAATTGCCCTCAGGCGC - 3' 3' - GCGCTGGACACTGTTTAACGGGAGTCTCGCGCp - 5'
2 SB 5'P <i>HhaI cut sites</i>	5' -pCGCGCCTTCCCACCTGTGACAAATTGCCCTCAGACCTCATCG-3' 3' -GCGGAAGGGTGGACACTGTTTAACGGGAGTCTGGAGTAGCGCp-5'
2 SB 5'P <i>9YZK</i>	5' -pAGTATCTTCCCACCTGTGACAAATTGCCCTCAGACCTCATCC-3' 3' -ATAGAAGGGTGGACACTGTTTAACGGGAGTCTGGAGTAGGTCp-5'

Supplementary Table S3.2: X-ray diffraction statistics for CC1+10 with HhaI cut sites

Data collection	
Light source	Synchrotron
Wavelength (Å)	0.979
Resolution Range (Å)	32.95 - 3.38 (3.48 - 3.38)
Space group	C121
Unit cell dimensions	
a, b, c (Å)	156.404 122.607 74.623
α , β , γ (°)	90 117.97 90
Unique reflections	26058 (2111)
Multiplicity	7.0 (6.8)
Completeness (%)	97.42 (94.36)
Mean I/sigma(I)	6.51 (0.18)
Wilson B-factor	137.52
R-merge	0.1221 (4.791)
R-meas	0.1324 (5.184)
R-pim	0.05072 (1.963)
CC1/2	0.994 (0.131)
Refinement	
Reflections used in refinement	17025 (1355)
Reflections used for Rfree	1710 (140)
R-work	0.2091 (0.4003)
R-free	0.2379 (0.4277)
Number of non-hydrogen atoms	2900
Macromolecules	2899
Ligands	1
Solvent	0
Protein residues	221
RMS (bonds) (Å)	0.004
RMS (angles) (°)	0.75
Ramachandran favored (%)	97.21
Ramachandran allowed (%)	2.79
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0
Clashscore	8.39
Average B-factor	236.21

Supplementary Table S3.3: Crystallization conditions of CC1+10 and CC1+21
Sitting drop vapor diffusion conditions for crystallization.

PDB Code:	Crystallization Condition:
Abbreviations: (pending release)	MgAc ₂ = magnesium acetate, Mg(CH ₃ COO) ₂ ; Li ₂ SO ₄ = lithium sulfate; MES = 2-(N-morpholino)ethanesulfonic acid.
9NCR	30 mM MgAc ₂ , 1 M Li ₂ SO ₄ , 50 mM MES pH 6.5
9YZJ	30 mM MgAc ₂ , 1.7 M Li ₂ SO ₄ , 50 mM MES pH 6.5
1 SB 3'P	40 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5
2 SB 3'P	40 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5
CC1+10 HhaI cut sites	30 mM MgAc ₂ , 1.5 M Li ₂ SO ₄ , 50 mM MES pH 6.5
CC1+21 HhaI cut sites	30 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5
9YZK	60 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5

Supplementary Table S3.4: Protein sequences

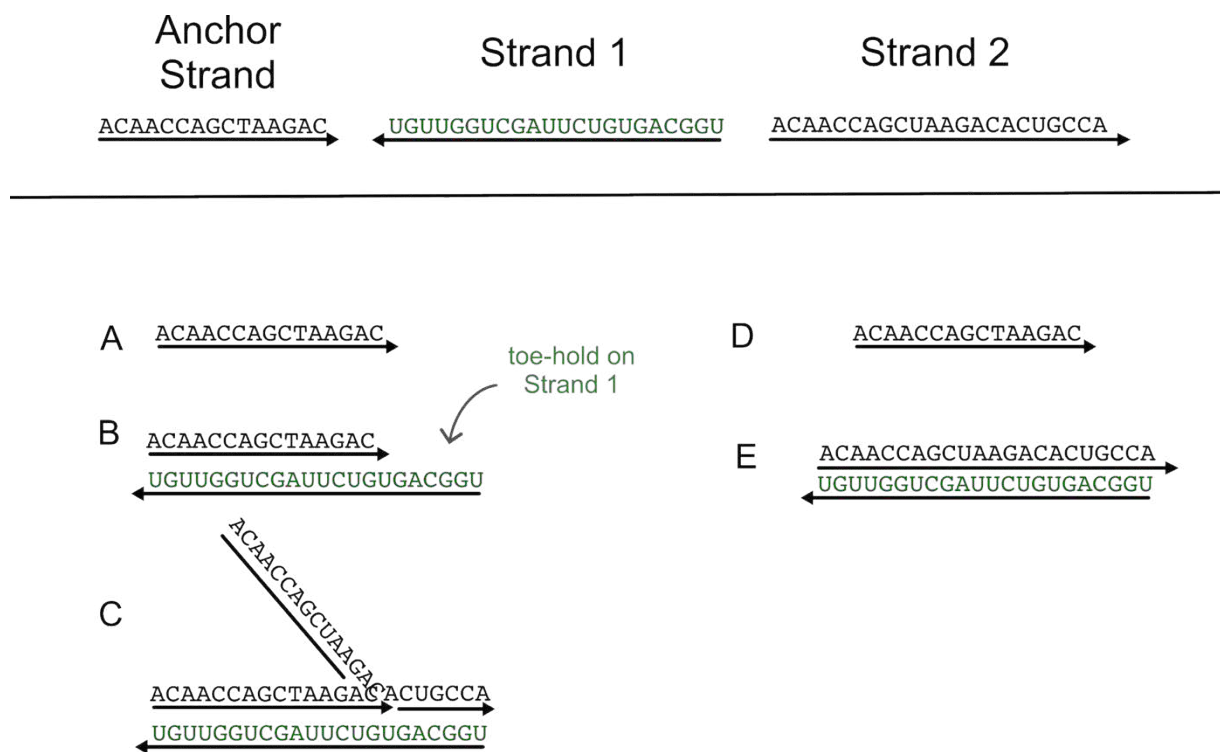
RepE54 Transcription Factor	MRGS HHHHHH GSMAETAVINHKKRKNSPRIVQSNDLTEAAYSLSRDQKRM LYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEASKDIRQALKSFAG KEVVFYRPEEDAGDEKGYESFPWFIKPAHSPSRGLYSVHINPYLIPFFIGLQN RFTQFRLSETKEITNPYAMRLYESLCQYRKPDGSGIVSLKIDWIIERYQLPQS YQRMPDFRRRFLQVCVNEINSRTPMRLSYIEKKKGRQTTHIVFSFRDITSMT TG
AmeLig	MRGS HHHHHH GSGEDKHDIVDGLQLAKQYGHSRQDINIAEYVWSEKLDGI RARWDGTELRTNRNNKIDAPAWFTANWPKATIDGELWIARGQFERTASIVL SKLTSVAPQSVAGSLPRTESTVDAMTATHSLPSKRWAKVRFMAFDMPVAG QSFDSRLNMLNMLKEATPNPTFAVVSQFTLSSVNALEEKLEQVTLSGGEG MLHHKKAFYHSGRSDKLIKVKQFEDAEAKVLAHFAGKGKFKGMMGSLLV ETPAGVQFKLGTGFSEKERRAPPAVGSWVTFKFYGVTKNGKPRFASFLRVR PPSDLPK
HhaI (pdb 6UKF not expressed in house)	MNWKEFEVFCVTYLNKTYGNKFAKKGESDSTTSIDILFTGNNPFYIEAKMPH SQCGQFVLIPNRAEYKFDYSPKNKSEINPYTQKIMQFMSENFSEYANLSTKG KIIPLESVFNWIKYKSKSVKFFITSNGDFIIFPIEHFEHYFNVSCYRIKK SGSRHLNSKSLPDFKQALDKKGISYTMRGLELHSDENIHDKRISGDDKDFLI KENNGAYHVKILSNTFNANVIFSISLKNNISLFILNEDRKAFEAAISL

Hexahistag in blue

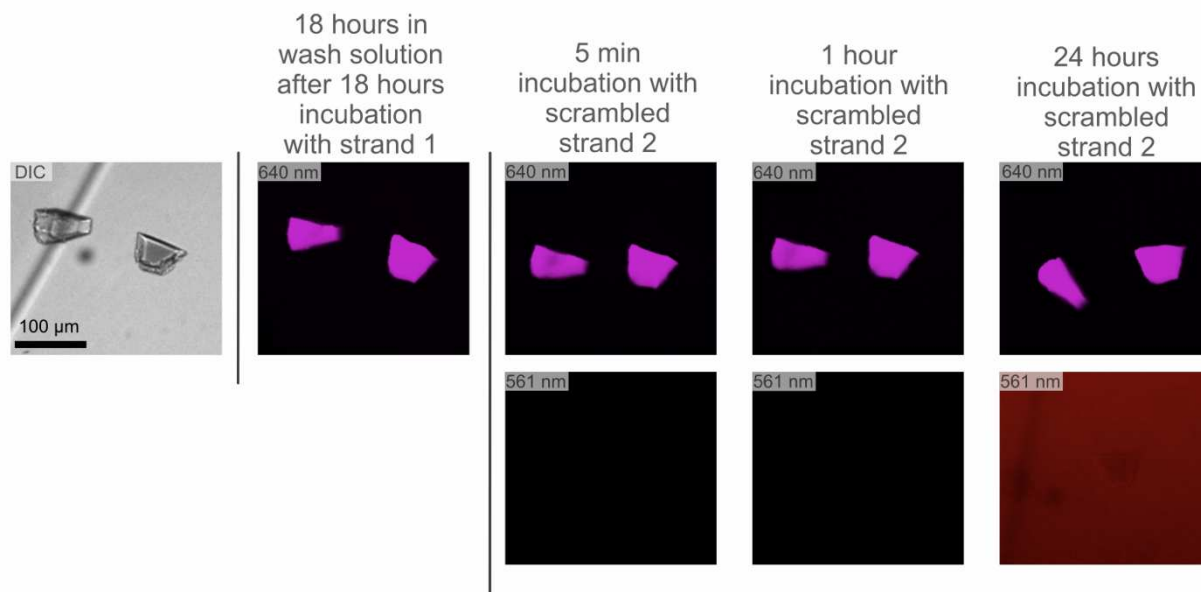
APPENDIX III. Supplemental Information for Chapter 4: Programmable Capture and Toe-hold

Mediated Release of Nucleic Acid Cargo in Porous Protein-DNA Co-Crystals

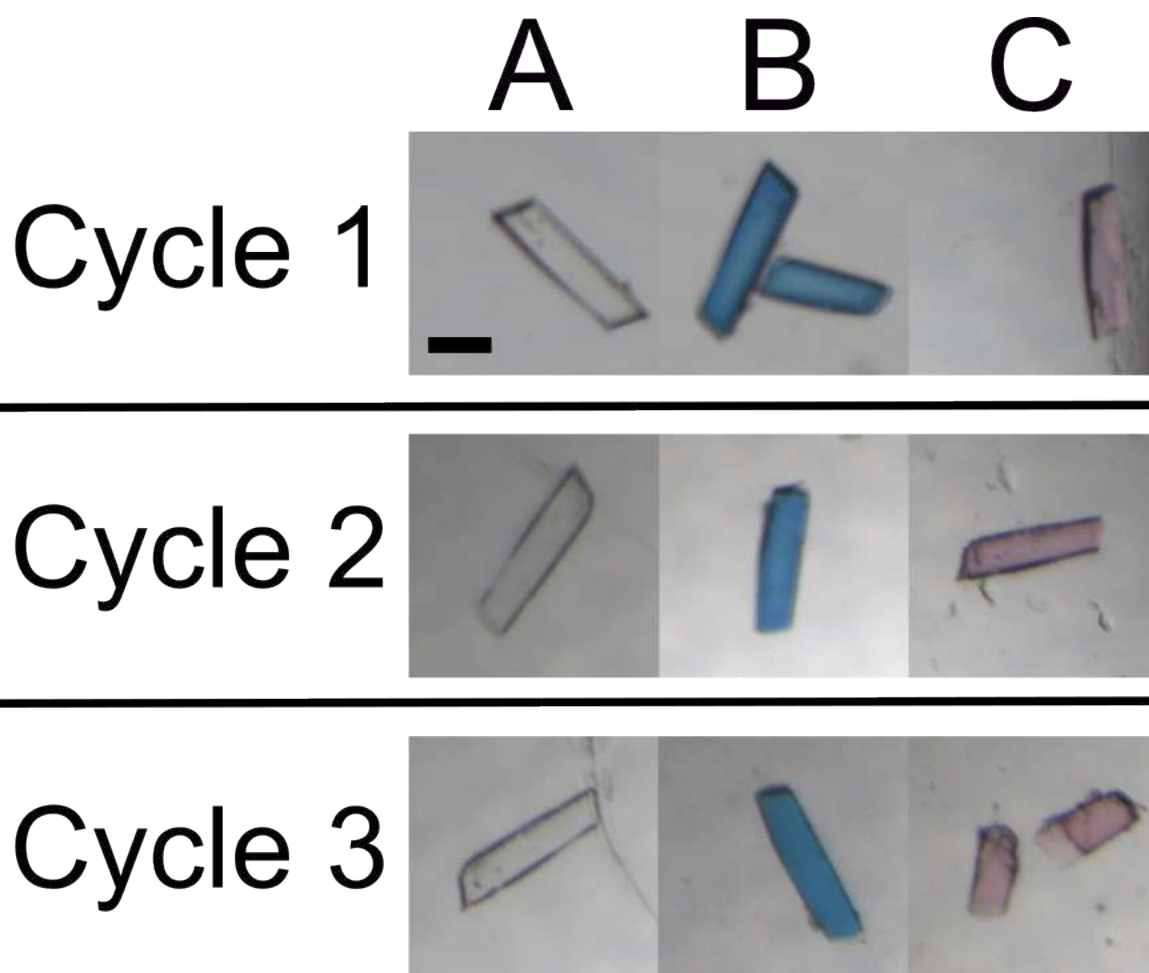
- Supplementary Figure 4.1:** Strand displacement mechanism in the RNA capture device
- Supplementary Figure 4.2:** DNA device scramble control confocal microscopy
- Supplementary Figure 4.3:** Multiple DNA capture cycles
- Supplementary Figure 4.4:** RNA capture in tailed and no-tail crystals confocal microscopy
- Supplementary Figure 4.5:** RNA capture scramble control confocal microscopy
- Supplementary Figure 4.6:** Time lapse loading of miR-34a strand 1
- Supplementary Figure 4.7:** Time lapse unloading of dsRNA miR-34a
- Supplementary Figure 4.8:** XRD electron density map comparison between mir-34a capture crystals with and without guest strand 1
- Supplementary Figure 4.9:** Custom plate reader adaptor schematics.
- Supplementary Table 4.1:** XRD statistics for CC1+10 with 15-nt tail for mir-34a capture
- Supplementary Table 4.2:** XRD statistics for CC1+10 with 15-nt tail for mir-34a capture and guest
- Supplementary Table 4.3:** XRD statistics for CC1+10 with 15-nt tail for DNA capture
- Supplementary Table 4.4:** Protein sequences
- Supplementary Table 4.5:** CC1+10 oligonucleotide sequences
- Supplementary Table 4.6:** Crystallization conditions of CC1+10



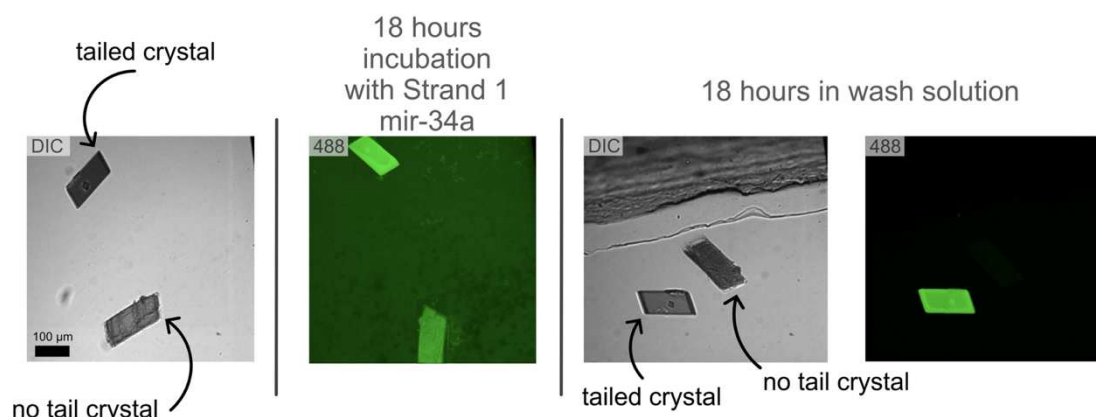
Supplementary Figure 4.1: Strand displacement mechanism in the RNA capture device. **A.** The ssDNA 15-nt anchor strand extending from the DNA component of a CC1 crystal. **B.** Strand 1 ssRNA (containing a fluorescein fluor) hybridizes with the anchor strand, leaving 7-nt of Strand 1 unbound. **C.** Strand 2 ssRNA invades and displaces the anchor strand using the 7-nt toe-hold, releasing dsRNA into solution, resulting in **D.** ssDNA anchor strand and **E.** Hybridized Strand 1 and Strand 2.



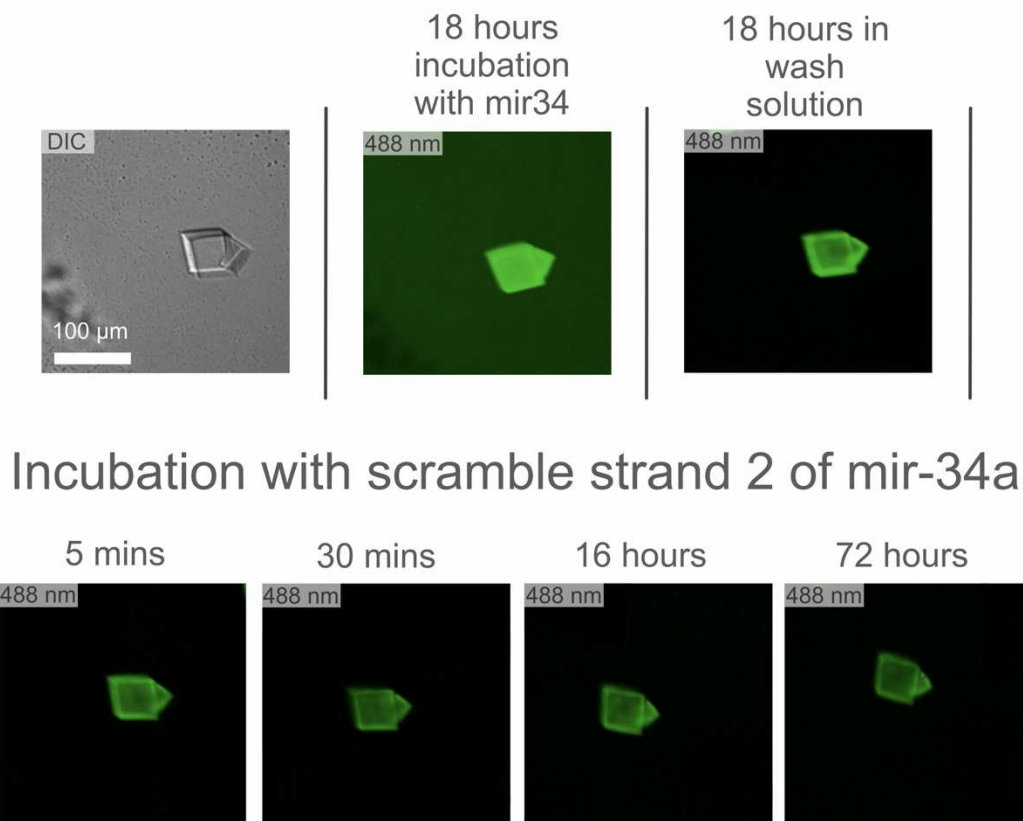
Supplementary Figure 4.2: DNA device scramble control confocal microscopy. A broken crystal with 15-nt anchor strand for capture of ssDNA species is shown in differential interference contrast. The crystal halves were soaked with Cy5-tagged Strand 1 of the DNA device cycle for 18 hours, then moved to a wash buffer for 18 hours and imaged with 640 nm laser excitation, detecting with a 660 nm emission filter. A scrambled version of Cy3-tagged Strand 2 is added to solution at 100 μM . For subsequent time points (5 minutes, 1 hour, and 24 hours) crystal halves were imaged with 640 nm laser excitation, detecting with a 660 nm emission filter as well as imaged with 561 nm laser excitation and 570 nm emission filter. As opposed to when Strand 2 (unscrambled) is added to solution, the crystal halves maintain their Cy5-Strand 1 emission profile and show Cy3-Strand 2 emission is equivalent to background levels.



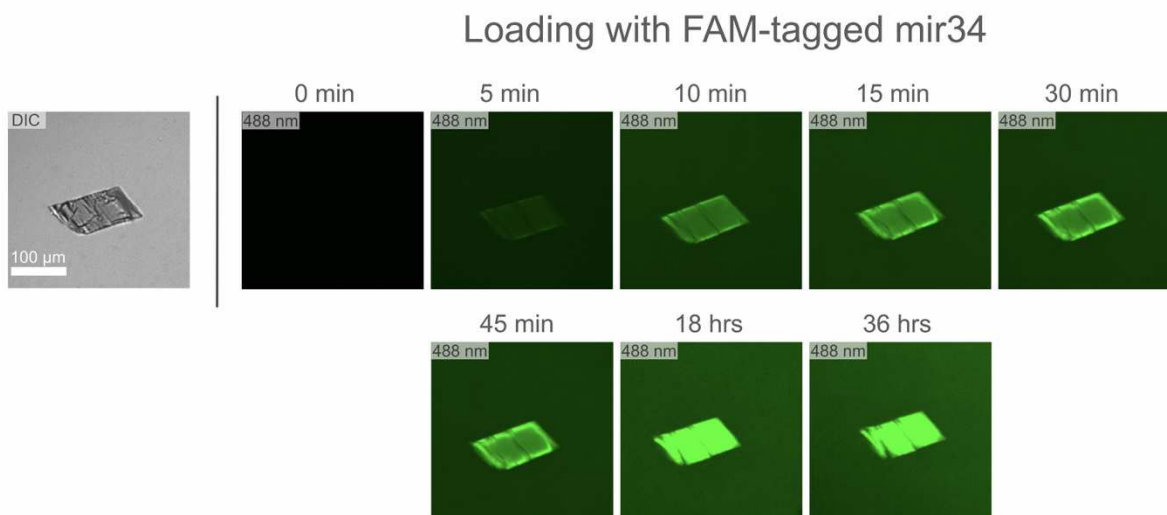
Supplementary Figure 4.3: Multiple DNA capture cycles. Scale bar is 100 μm . One DNA capture crystal is cycled three times through **A.** unbound, empty anchor strand. **B.** Cy5 Strand 1 addition. **C.** Cy3 Strand 2 addition then Strand 3 and Strand 4 addition to cycle back to A. Crystal was looped into loading solution, the image was captured using white light microscopy, then looped into a wash solution before the next step of the cycle. After 16 solution changes the crystal broke on the loop before the final loading of Strand 2 shown in Cycle 3 C.



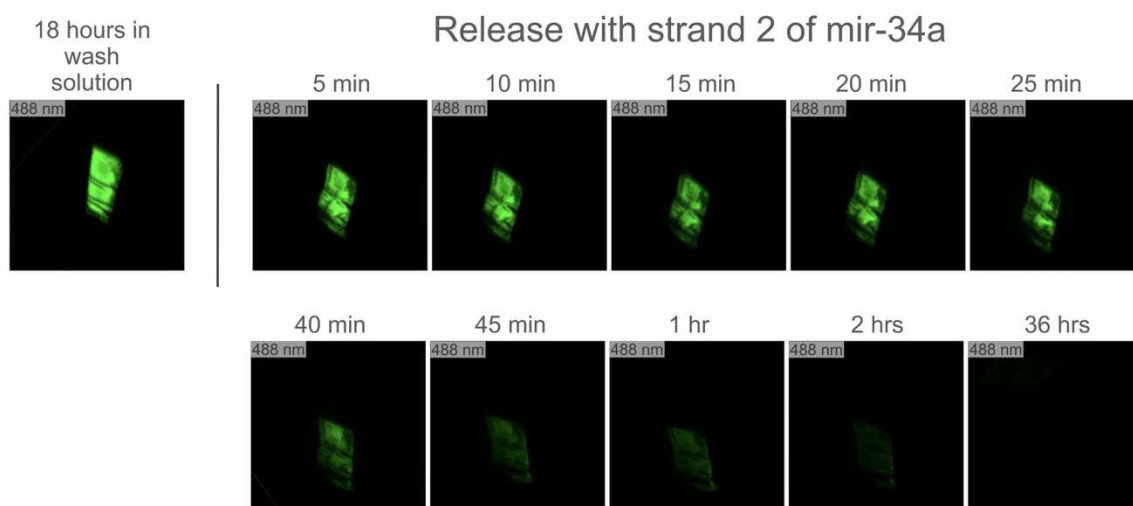
Supplementary Figure 4.4: RNA capture in tailed and no-tail crystals confocal microscopy. Roughly equivalently sized crystals, one with a mir-34a 15-nt capture tail and one without, were imaged in a wash solution using differential interference contrast. 100 μ M of FAM-tagged Strand 1 mir-34a is added to solution and crystals soaked for 18 hours. Crystals were imaged using 488 nm excitation and 515 nm emission filter. Both crystals were moved to a fresh wash solution and imaged using 488 nm excitation and 515 nm emission filter, as well as differential interference contrast. Only the tailed crystal maintained the fluorescent signal, indicating that base complementarity is necessary for co-crystals to stabilize ssRNA cargo through wash steps.



Supplementary Figure 4.5: RNA capture scramble control confocal microscopy. A crystal a second crystal grown off the side, with 15-nt anchor strand for capture of mir-34a is shown in differential interference contrast. The crystals were soaked with FAM-tagged Strand 1 of mir-34a for 18 hours then imaged using 488 nm excitation and 515 nm emission filter. The crystals were then moved to a wash buffer for 18 hours and imaged using 488 nm excitation and 515 nm emission filter. A scrambled version of mir-34a Strand 2 is added to solution at 100 μ M. For subsequent time points (5 minutes, 30 minutes, 16 hours, and 72 hours) crystals were imaged using 488 nm excitation and 515 nm emission filter. As opposed to when Strand 2 (unscrambled) is added to solution, the crystals maintain their FAM-Strand 1 emission profile, indicating that sequence specific triggers are needed to release the dsRNA mir-34a.

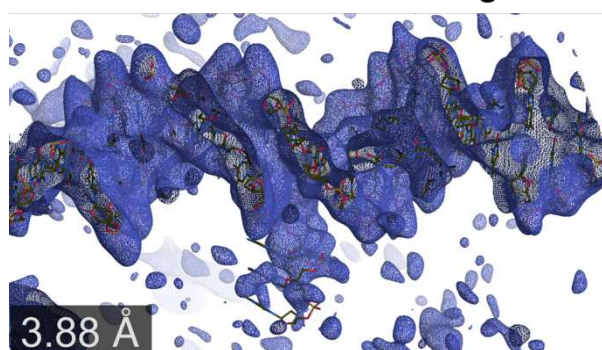


Supplementary Figure 4.6: Time lapse loading of miR-34a strand 1. A crystal with 15-nt tail for miR-34a capture is imaged using differential interference contrast. 100 μ M of Strand 1 miR-34a is then added to solution and the crystal is imaged using 488 nm excitation and 515 nm emission filter over the next 45 minutes, then at 18 hours and 36 hours. Time point one (0 min) shows no background fluorescence prior to Strand 1 being added to solution. Time point two (5 min) shows an increase in background fluorescence, as well as the beginnings of fluorescence increase at the crystal edges. The subsequent time points (10 min, 15 min, 30 min, and 45 min) show an increase in fluorescence at the crystal edges, with the fluorescence edge proceeding further into the crystal interior. This edge loading indicates specific binding within the crystal, as opposed to nonspecific binding that would likely appear as increase fluorescence evenly throughout the crystal interior. After 18 hours the crystal appears saturated, with little change in fluorescence after 36 hours.

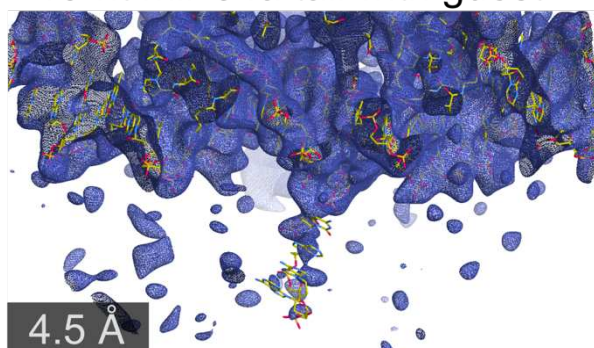


Supplementary Figure 4.7: Time lapse release of dsRNA miR-34a strand 1+strand 2. The same crystal from Fig. S4.6 is imaged after 18 hours in a wash solution post-loading. 100 μ M of Strand 2 mir-34a is then added to solution and the crystal is imaged using 488 nm excitation and 515 nm emission filter over the next 2 hours, then at 36 hours. After 2 hours the fluorescence has almost completely dissipated, implying that strand 2 has displaced strand 1 from the crystal anchor, hybridizing into dsRNA. Once unbound from the crystal the fluorescently tagged strand 1 diffuses out of the crystal and fluorescent signal decays to background levels. After 18 hours the crystal appears completely indistinguishable from the background fluorescence.

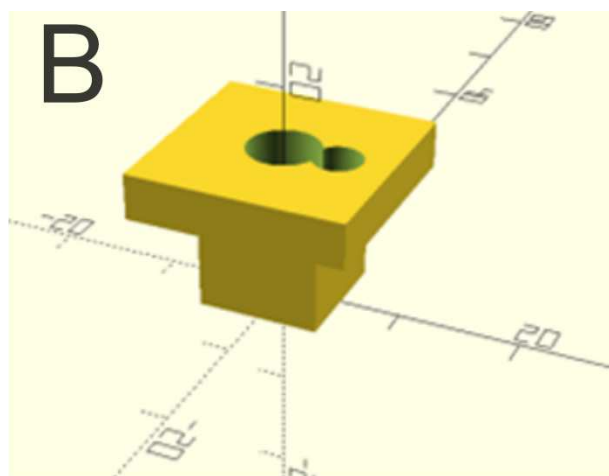
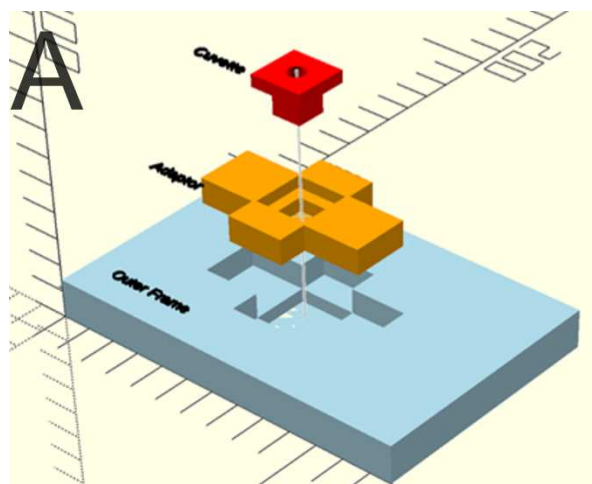
15-nt mir-34a tail without guest



15-nt mir-34a tail with guest



Supplementary Figure 4.8: XRD electron density map comparison between mir-34a capture crystals with and without guest strand 1. Maps are composite 2Fo–Fc map contoured at 1σ . The maps center on the anchor strand tails, showing the agreement between the electron density map and modelled in nucleotides of the anchor strands. In the absence of guest, 3 nucleotides of the 15 nucleotide tail can be reliably modelled. After addition of 100 μM Strand 1 of mir-34a an additional nucleotide for a total of 4 out 15 of the tail can be modelled. No bases from the duplex between anchor and Strand 1 ssRNA could be modelled.



Supplementary Figure 4.9: Custom plate reader adaptor schematics. **A.** Plate reader 3D printed prototype. In light blue the main outer frame, in orange the adaptor and in red the cuvette. **B.** The yellow prototype shows the final cuvette design with the side pocket well for the crystals. This 3D printed chassis maps on to a well in a standard fluorescence plate reader, allowing the crystal to remain in solution contact without falling into the light path.

Supplementary Table S4.1: X-ray diffraction statistics for CC1+10 with 15-nt tail for miR-34a capture

PDB code	
Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	65.61 - 3.88 (4.06 - 3.88)
Space group	C121
Unit cell dimensions	
a, b, c (Å)	143.004 131.214 72.473
α, β, γ (°)	90 118.6 90
Unique reflections	19830 (2355)
Multiplicity	6.9 (6.6)
Completeness (%)	97.72 (92.83)
Mean I/sigma(I)	1.83 (0.27)
Wilson B-factor	109.14
R-merge	0.8562 (6.304)
R-meas	0.9272 (6.837)
R-pim	0.3518 (2.61)
CC1/2	0.971 (0.139)
Refinement	
Reflections used in refinement	10689 (1255)
Reflections used for Rfree	1066 (122)
R-work	0.2807 (0.5069)
R-free	0.3086 (0.4891)
Number of non-hydrogen atoms	2927
Macromolecules	2924
Ligands	3
Solvent	0
Protein residues	225
RMS (bonds) (Å)	0.005
RMS (angles) (°)	0.89
Ramachandran favored (%)	95.02
Ramachandran allowed (%)	3.62
Ramachandran outliers (%)	1.36
Rotamer outliers (%)	0
Clashscore	9.83
Average B-factor	181.52

Supplementary Table S4.2: X-ray diffraction statistics for CC1+10 with 15-nt tail for miR-34a capture and guest ssRNA strand 1

Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	67.74 - 4.5 (4.85 - 4.5)
Space group	C121
Unit cell dimensions	
a, b, c (Å)	153.054 128.268 72.787
α , β , γ (°)	90 117.73 90
Unique reflections	37691 (5125)
Multiplicity	6.5 (4.9)
Completeness (%)	99.18 (99.66)
Mean I/sigma(I)	1.03 (0.27)
Wilson B-factor	50.30
R-merge	1.938 (1.522)
R-meas	2.104 (1.673)
R-pim	0.8084 (0.6783)
CC1/2	0.313 (0.120)
Refinement	
Reflections used in refinement	7401 (1472)
Reflections used for Rfree	745 (140)
R-work	0.2922 (0.2542)
R-free	0.3149 (0.3118)
Number of non-hydrogen atoms	3114
Macromolecules	3112
Ligands	2
Solvent	0
Protein residues	226
RMS (bonds) (Å)	0.014
RMS (angles) (°)	1.76
Ramachandran favored (%)	95.05
Ramachandran allowed (%)	4.05
Ramachandran outliers (%)	0.90
Rotamer outliers (%)	0
Clashscore	6.01
Average B-factor	161.33

Supplementary Table S4.3: X-ray diffraction statistics for CC1+10 with 15-nt tail for DNA capture

Data collection	
Light source	Synchrotron
Wavelength (Å)	1
Resolution Range (Å)	64.91 - 3.81 (3.96 - 3.81)
Space group	I121
Unit cell dimensions	
a, b, c (Å)	74.097 126.617 133.879
α , β , γ (°)	90 90.16 90
Unique reflections	23686 (2645)
Multiplicity	12.1 (7.3)
Completeness (%)	99.01 (97.25)
Mean I/sigma(I)	9.1 (0.5)
Wilson B-factor	232.24
R-merge	0.188 (2.928)
R-meas	0.208 (3.147)
R-pim	0.062 (1.292)
CC1/2	0.993 (0.419)
Refinement	
Reflections used in refinement	12040 (1308)
Reflections used for Rfree	1201 (136)
R-work	0.2741 (0.4109)
R-free	0.3082 (0.4860)
Number of non-hydrogen atoms	3112
Macromolecules	3111
Ligands	1
Solvent	0
Protein residues	221
RMS (bonds) (Å)	0.005
RMS (angles) (°)	0.77
Ramachandran favored (%)	94.88
Ramachandran allowed (%)	5.12
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	9.63
Average B-factor	264.6

Supplementary Table 4.4: Protein sequences

Hexahistag in blue, mutation underlined.

Protein Name	Amino Acid Sequence
RepE54 Transcription Factor	MRGS HHHHHH GSMAETAVINHKKRKNSPRIVQSNDLTEAAYSLSRDQKRM LYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEASKDIRQALKSFAG KEVVFYRPEEDAGDEKGYESFPWFIKPAHSPSRGLYSVHINPYLIPFFIGLQN RFTQFRLSETKEITNPYAMRLYESLCQYRKPDGSGIVSLKIDWIIERYQLPQS YQRMPDFRRRFLQVCVNEINSRTPMRLSYIEKKKGRQTTHIVFSFRDITSMT TG
RepE54 variant (Y184E) <i>Mutation underlined</i>	MRGS HHHHHH GSMAETAVINHKKRKNSPRIVQSNDLTEAAYSLSRDQKRM LYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEASKDIRQALKSFAG KEVVFYRPEEDAGDEKGYESFPWFIKPAHSPSRGLYSVHINPYLIPFFIGLQN RFTQFRLSETKEITNPYAMRLYESLC <u>Q</u> ERKPDGSGIVSLKIDWIIERYQLPQS YQRMPDFRRRFLQVCVNEINSRTPMRLSYIEKKKGRQTTHIVFSFRDITSMT TG

Supplementary Table 4.5: Oligonucleotide sequences

PDB Code / Purpose	CC1⁺¹⁰ (or 12) DNA Sequences
RNA capture crystal <u>ssDNA tail underlined</u> , <i>third strand italicized</i>	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCG - 3' 3' - GCCTGGACACTGTTTAACGGGAGTCGGGCCGp - 5' <u>CAGAATCGACCAACA</u>
DNA capture crystal <u>ssDNA tail underlined</u> , <i>third strand italicized</i>	5' - pCCCGGACCTGTGACAAATTGCCCTCAGCCCG - 3' 3' - GCCTGGACACTGTTTAACGGGAGTCGGGCCGp - 5' <u>TCTCATTCGTATCTG</u>
9NCR (1 SB 5'P control crystal from Fig. S4.X)	5' - pGCCGGACCTGTGACAAATTGCCCTCAGCCCG - 3' 3' - GGCCTGGACACTGTTTAACGGGAGTCGGGCCp - 5'
Strand 1 miR-34a ssRNA	5' - UGGCAGUGUCUUAGCUGGUUGU - 3'
Strand 2 miR-34a ssRNA	5' - ACAACCAGCUAAGACACUGCCA - 3'
Scramble Strand 2 mir-34a ssRNA	5' - CCAGCAGACCACAUAAAGCUCA - 3'
Strand 1 DNA device ssDNA	5' - TAAGCATAGAC - 3'
Strand 2 DNA device ssDNA	5' - AGAGTAAGCATAGACGCTA - 3'
Scramble Strand 2 DNA device ssDNA	5' - ACAATAGAAGGCTGCTGAA - 3'
Strand 3 DNA device ssDNA	5' - GTCTATGCTTA - 3'
Strand 4 DNA device ssDNA	5' - TAGCGTCTATGCTTACTCT - 3'

Supplementary Table 4.6: Crystallization conditions of CC1+10, sitting drop vapor diffusion conditions for crystallization.

PDB Code:	Crystallization Condition:
Abbreviations: (pending release)	MgAc ₂ = magnesium acetate, Mg(CH ₃ COO) ₂ ; Li ₂ SO ₄ = lithium sulfate; MES = 2-(N-morpholino)ethanesulfonic acid.
9NCR	30 mM MgAc ₂ , 1 M Li ₂ SO ₄ , 50 mM MES pH 6.5
DNA capture crystal	30 mM MgAc ₂ , 1.7 M Li ₂ SO ₄ , 50 mM MES pH 6.5
RNA capture crystal	40 mM MgAc ₂ , 1.6 M Li ₂ SO ₄ , 50 mM MES pH 6.5