University of California
Santa Barbara

Self-assembling DNA complexes: an *in vitro* platform to probe the relation between transcription and condensed DNA phases

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

Doctor of Philosophy
in
Biochemistry & Molecular Biology

by

Dan Trong Nguyen

Committee in charge:

Professor Omar A. Saleh, Chair
Professor Irene A. Chen
Professor Christopher S. Hayes
Professor Cyrus R. Safinya

September 2018

The Dissertation of Dan Trong Nguyen is approved.

Professor Irene A. Chen

Professor Christopher S. Hayes

Professor Cyrus R. Safinya

Professor Omar A. Saleh, Committee Chair

August 2018

Self-assembling DNA complexes: an *in vitro* platform to probe the relation between transcription and condensed DNA phases

by

Dan Trong Nguyen

I dedicate this dissertation to my parents - Viet & Sarabeth - and my brother - Andy - for their unconditional love & support.

Thank you for always being there. I would not get very far without you.

Acknowledgements

First and foremost, I would like to thank my advisor, Professor Omar A. Saleh. He took a chance on a 26-year-old Harvard kid with a weak quantitative background, and trusted that kid to develop over time. Thank you for your patience, for providing me with incredible opportunities (in research as well as life), and for your steady mentorship & guidance.

At its core, scientific research is a collaborative enterprise. I have had the good fortune to work with many people throughout my doctoral studies. Irene Chen, Chris Hayes, Cyrus Safinya, Deborah Fygenson, Bob McMeeking, Fyl Pincus, Luc Jaeger, Kevin Plaxco, Matt Helgeson, Craig Montell, Jennifer Ross, Rae Robertson-Anderson, Achim Wixforth, Tim Liedl, Chris Westerhausen, Megan Valentine, Bekele Gurmessa, Ben Lopez, Mary Raven, Lior Dassau, Nelly Traitcheva, Jack Reifert, Yali Zhang, Jamie Wilcox, Jen Mobberley, Ranajay Saha, Lukas Schnitzler, Manuel Brugger, Greg Ekberg, Grant Gucinski, Greg Campbell, Evan Janzen, Claire Tran, Sam Verbanic, Lourdes Velazquez, Tuan Nguyen, Elmer Guzman, Stephanie Khairallah, Sarah Grundeen, Billy Wonderly, Peter Aziz, & Jeremy Robins...thank you for helping me learn & do science.

To the graduate students and postdoctoral fellows with whom I worked on a daily basis in the Saleh lab, you deserve the lion's share of my gratitude: Byoung-jin Jeon, Chang-Young Park, Andrew Dittmore, Bob Lansdorp, John Berezney, David Jacobson, Sarah Innes-Gold, Ian Morgan, Nate Conrad, & Gabi Abraham. Thank you so much for the wonderful discussions, "Aha!" moments, and immeasurable help & support.

To all my colleagues, friends, and family (that means you too Tally Dog), thank you for being there with me. Grad school can be quite taxing, both physically and mentally. You helped me make it to the finish line. For that, I will be forever grateful. Thank you all so very much.

Curriculum Vitæ

Dan Trong Nguyen

Education

2018	Ph.D. in Biomolecular Science & Engineering, University of California, Santa Barbara.
2012	P.S.M. in Nanoscience, Arizona State University.
2008	A.B. in Biochemical Sciences, Harvard University.

Honors and Awards

2017	UC Santa Barbara Graduate Division Dissertation Fellowship
2017	American Physical Society Shirley Chan Student Travel Award
2016-2017	Bavaria California Technology Center (BaCaTeC) Supported Project
2014-2016	UC Santa Barbara Chang Prize for 'Best Graduate Talk' (annual)
2014	UC Santa Barbara Chang Merit Fellowship for Biochemistry & Molecular Biology
2012	Amgen Graduate Fellowship

Publications

- B.-j. Jeon, **Dan T. Nguyen**, G. Abraham, N. Conrad, Deborah K. Fygenson, and Omar A. Saleh. (2018) Salt-dependent properties of a coacervate-like, self-assembled DNA liquid. *Soft Matter* 14(34): 7009-7015.
- B.J. Gurmessa, N. Bitten, **Dan T. Nguyen**, Jennifer L. Ross, Omar A. Saleh, Moumita Das, and Rae M. Robertson-Anderson. (2017) Active disassembly and reassembly of actin networks induces distinct biphasic mechanics. *ArXiv* 1710.03781.
- **Dan T. Nguyen** and Omar A. Saleh. (2017) Tuning phase and aging of DNA hydrogels through molecular design. *Soft Matter* 13(32): 5421-5427.
- C.-Y. Park, D.R. Jacobson, **Dan T. Nguyen**, S. Willardson, Omar A. Saleh. (2016) A thin-permeable membrane device for single-molecule manipulation. *Review of Scientific Instruments* 87(1): 014301.
- H. Yamashita, **Dan T. Nguyen**, E. Chung. (2014) Blood-based assay with secreted Gaussia luciferase to monitor tumor metastasis. In: Badr C. (eds) *Bioluminescent Imaging. Methods in Molecular Biology (Methods and Protocols).* Humana Press 1098: 145-151.
- R. Maini, **Dan T. Nguyen**, S. Chen, L.M. Dedkova, S.R. Chowdhury, R. Alcala-Torano, Sidney M. Hecht. (2013) Incorporation of β-amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorganic & Medicinal Chemistry* 21(5): 1088-1096.

Abstract

Self-assembling DNA complexes: an *in vitro* platform to probe the relation between transcription and condensed DNA phases

by

Dan Trong Nguyen

DNA organization within a cell is multifaceted and dynamic. Not only does chromatin structure vary with position, as there is significant heterogeneity in the degree of local compaction within the global DNA complex, but it also changes with time, showing dependence on, e.g., the growth stage of a cell. The condensation state of DNA is intimately coupled to gene accessibility and expression: changes to DNA organization can modulate transcription and, in turn, transcribed RNA can reorganize chromatin structure. We sought to develop a more tractable, *in vitro* platform to probe this complex relationship and utilized a self-assembling DNA nanostar (NS) as the principal component of a chromatin mimic. In this dissertation, I will discuss our research with condensed phases of self-assembled NSs, describing structural properties of the system's two phases (i.e. NS-networks & NS-liquids), features of NS-liquids that enable application as a model of membraneless organelles, and initial work examining the transcription of genes integrated within NS-liquids.

Contents

Chapter 1

Introduction - Towards artificial chromatin

The 20^{th} century witnessed a revolution in biological research that illuminated many of the core principles governing cellular behavior, including the central role played by DNA. A mechanistic understanding of the constituent molecular processes has advanced to the extent of birthing entire disciplines predicated on controlling gene expression from DNA (e.g. bioengineering and synthetic biology). However, while many intricacies of molecular biology have been elucidated, one particular relationship, unequivocally deemed essential to biological understanding, has persisted in remaining unresolved: the interplay between DNA organization and gene expression.

Before presenting our research towards developing an *in vitro* platform to probe this relationship, I use this first chapter to establish the scientific framework with which to contextualize the work. I briefly review fundamental features of both gene expression & chromatin organization, and follow with a description of the importance of *in vitro* approaches toward biological research. I end with a coarse overview of the contents in the remaining chapters of this dissertation.

1.1 Gene expression

1.1.1 Hierarchical coding of genetic information

At its foundation, biological information is encoded in deoxyribonucleic acid, better known by its famous abbreviation: DNA. While proteins are predominantly responsible for constituting and implementing the molecular processes and biochemistry that sustain life, it is DNA that contains the instructions for producing these proteins (and other biomolecules) for use in cells. This genetic information – stored in DNA as sequence combinations of nucleotides (nt) exhibiting one of 4 nitrogenous bases: adenine (A), thymine (T), cytosine (C), & guanine (G) – is first converted to ribonucleic acid (RNA) before subsequent conversion to protein. This hierarchical flow of information, coded from one bio-polymer to the next, was elucidated near the mid-20^{th} century and is colloquially known as biology's "Central Dogma" [1].

While this directional flow of information has proven generally true, and remains the presiding principle in biology, extensive work in the latter half of the 20^{th} century (and continued to present day) has led to additions and clarifications, as well as revealed

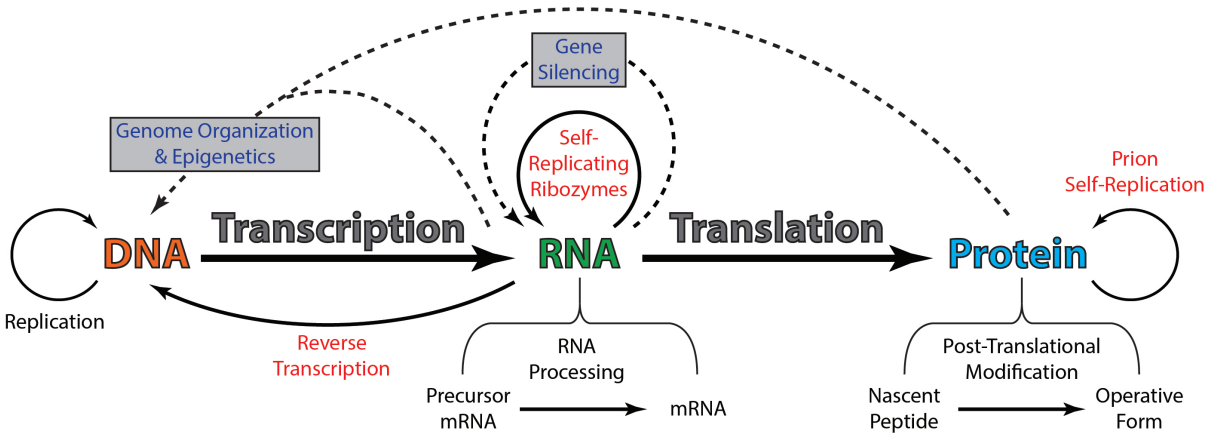


Figure 1.1: **Flow of biological information.** Solid arrows convey the flow of information. Dashed arrows indicate regulatory effects on gene expression.

exceptions, to the original simplistic model of $DNA \rightarrow RNA \rightarrow Protein$ (summarized in Fig. 1.1). It should be acknowledged that increased complexity was not unexpected by scientists at the time when the original model was proposed [2].

- **Additions/Clarifications:** From the original information encoded by DNA, more information (and thus biological diversity and complexity) can be added at the RNA level, through RNA processing (e.g. alternative splicing), as well as at the protein level, through post-translational modifications (e.g. phosphorylation, methylation, acetylation, ubiquitylation, glycosylation, and proteolysis) [3, 4].

- **Exceptions:** Concerning the flow of information, various cases have been identified that go against the "dogmatic" grain: RNA retroviruses utilize reverse transcriptases to convert their RNA into DNA before continuing the expected genetic progression [5, 6]; scientists studying the 'RNA world' hypothesis [7] have devised systems comprised of catalytic RNA capable of replication without the aid of a DNA template [8, 9, 10, 11] (however, self-replicating ribozymes have yet to be found in natural systems); and prion proteins, though originally derived from DNA, self-replicate (e.g. infect and amplify in other cells) without any nucleic acid mediation [12, 13].

1.1.2 Gene regulation

Increased sophistication was also revealed in the context of gene regulation, which encompasses processes that control when a gene (i.e. a sequence of DNA that encodes a product with biological function(s)) is expressed and to what degree. A basic understanding of regulation centers on the behavior of transcription factors (TFs). At the end of a protein signaling cascade, perhaps in response to an environmental stimulus or through homeostatic maintenance, a TF binds to a genetic operator and aids (or hinders) the

binding and function of RNA polymerase (RNAp) at a promoter. RNAp transcribes the gene into RNA, which is then translated into protein by ribosomes. It is now abundantly clear that once a gene is transcribed into RNA, it is by no means guaranteed to continue expression to the protein level.

Extensive work has revealed the sizable regulatory roles played by non-coding RNA (ncRNA), defined as RNA that does not encode biologically-active proteins. A recent accounting of genes in the human genome suggests that there are as many non-coding genes as there are protein-coding genes, both numbering around 21,000 [14, 15]. NcRNAs can regulate expression in both pre- and post-transcriptional manners. Short RNAs (including microRNAs and short interfering RNAs, known as miRNAs and siRNAs, respectively, that are roughly 20-30 nt long) play essential roles in enzymatic complexes that degrade messenger RNA (mRNA) transcripts in a sequence-specific manner, thereby preventing their translation into protein [16]. In addition to these RNA silencing capabilities, these short ncRNAs, along with a class of much longer ncRNAs (candidly, though uncreatively, named long non-coding RNAs, or lncRNAs), also modulate gene regulation through epigenetic mechanisms [17, 18]. That is, they control expression at the 'ground' level, dictating whether a gene is transcribed in the first place.

1.1.3 Epigenetic effects

The meaning of the term *epigenetics* has acquired ambiguity since its coinage by Waddington in 1942 [19], diverging over the years as it was embraced by disparate scientific fields [20]. For clarity, I use the term in this manuscript to describe processes that affect gene expression through control of the chemical and/or organizational state of DNA, thereby adding another layer of information control beyond DNA sequence. While the term typically describes expression patterns that are inherited between cellu-

lar generations, I do not use it in that way, preferring to adopt a broader definition that does not hinge on inheritance and can be used to describe a means of transcriptional regulation reliant on DNA accessibility.

For each cell in the human body, approximately 2 meters of DNA [21], collectively referred to as **chromatin**, must be stored in a volume with a radius of approximately 5 μm, depending on the cell type. By necessity, cells must package their DNA into dense structures with hierarchical organization. This tight-packing can make it difficult for cellular machinery to sterically access DNA and transcribe genes. For example, work with a hormone-activated TF demonstrated that, when examined soon after hormone induction, the TF was bound almost exclusively (approx. 95%) to pre-existing regions of accessible chromatin [22]. The TF had difficulty interacting with tightly-packed, less accessible (or inaccessible) DNA. This situation is analogous to reading a book with all its pages glued together; one must first unglue the pages before the encoded information can be read. Taking this analogy further, the different cell types in the body look and act in dramatically different ways because, though they have the exact same 'instruction manual', certain 'chapters' are open & available to some cell types but remain 'glued together' & unreadable to others. That is, cellular differences in DNA packaging and organization determine whether a gene can be expressed at the DNA level, the very start of biology's "Central Dogma".

1.2 DNA organization in the cell

While the history of chromatin research dates as far back as the early 19^{th} century [23], recent decades have seen a surge in interest (and scientific study) of chromatin due in large part to the emergence of powerful experimental and computational techniques, including fluorescence *in situ* hybridization methods; chromatin conformation capture

strategies & other sequencing-based approaches; and super-resolution/single-molecule fluorescent microscopy & force spectroscopy techniques; among many others [24]. They have worked to spur collaborative efforts between diverse disciplines (ranging from biology to physics to computer science) toward attaining a better understanding of the structure and organization of genomic DNA.

1.2.1 Eukaryotic chromatin

In eukaryotic cells, the first level of DNA organization is the **chromosome**, one of several extremely-long individual DNA strands that occupy a cell nucleus. During interphase, the stage of the eukaryotic cell cycle where a typical cell spends most of its time and during which the bulk of cellular gene expression occurs, each chromosome occupies an expansive (though distinct) region known as a **chromosome territory (CT)** (Fig. 1.2), in contrast to the famous X-structure that is briefly adopted during mitosis. Within each CT, there are distinct regions of DNA that appear to consistently share physical contact or proximity. The formation and maintenance of these macrodomains have been attributed to the activity of DNA loop-extruding enzymes (Fig. 1.2 bottom right) and they have been observed to persist through multiple rounds of cell division as well as be conserved in related species [25]. These macrodomains exist on the order of hundreds to thousands of kilobases long and are customarily referred to as **topologically associated domains (TADs)**.

Much of the work supporting (if not defining) the existence of TADs was done with chromatin conformation capture techniques [26, 27]. There have been many adaptations of the technique since its initial development by Dekker and Kleckner [28], but they are all based on the same general idea of: (1) Crosslinking physically-interacting chromatin (i.e. DNA strands in contact) using formulin; (2) Subsequent enzymatic DNA digestion

& re-ligation, so as to link interacting DNA segments into a single entity; and (3) DNA sequencing followed by identification of the genome location of each interacting segment using a sequenced genome. In this fashion, researchers are able to generate contact frequency maps between different regions of the genome. These experiments demonstrated that though two segments of a chromosome can be distantly separated along a linear DNA sequence (e.g. a million bases apart), they can be in immediate physical contact in three dimensions. Further examination has revealed that TAD domains appear to corre-

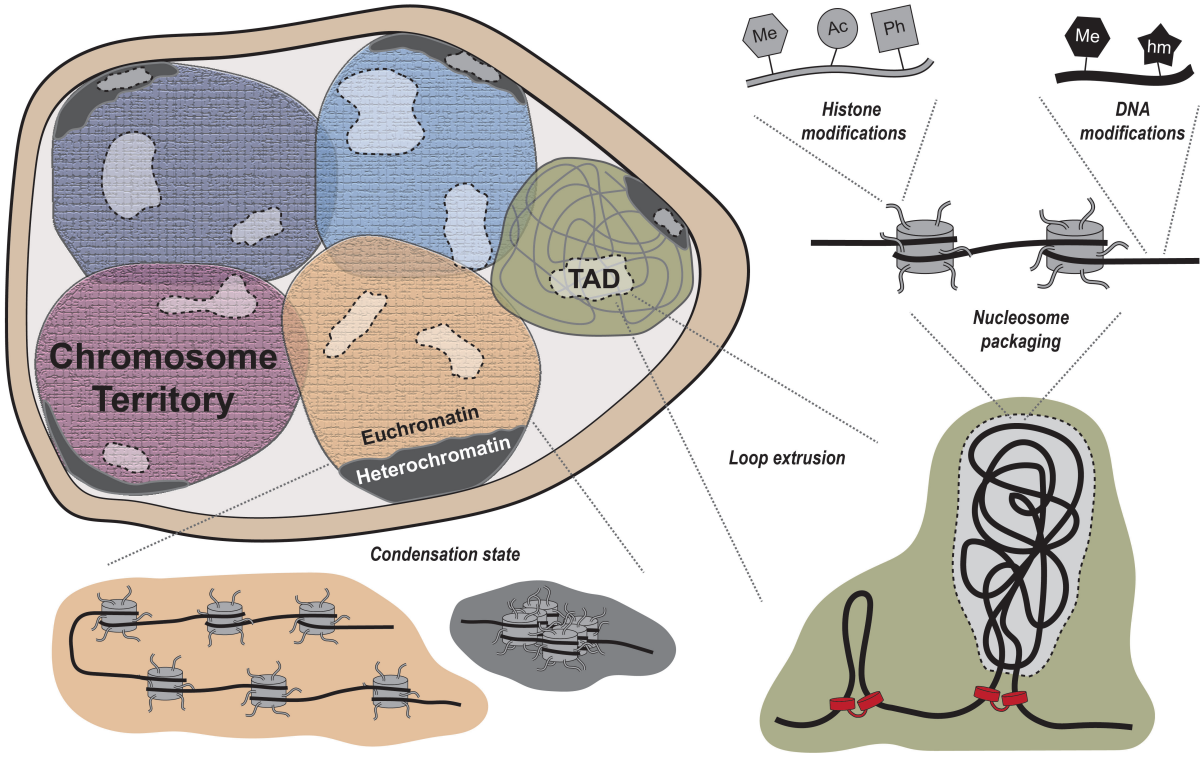


Figure 1.2: **Eukaryotic Chromatin Organization.** DNA structure in the eukaryotic cell (during interphase) exhibits multiple levels of organization. Individual chromosomes occupy chromosome territories that contain topologically-associated domains (TADs) that are hypothesized to be formed by enzymatic complexes utilizing loop extrusion mechanisms. Within each TAD, DNA is further packaged through non-covalent interactions with histones, which are mediated by epigenetic signatures on both the histone proteins and DNA. The degree of compaction (i.e. condensation state) is closely correlated with the level of gene expression, further classifying DNA as either euchromatin or heterochromatin.

late with markers of chromatin activity; that is, distinct TADs can be further classified as either containing *active* or *silent* chromatin [29].

In eukaryotic cells, the general dichotomy, initially proposed by Heitz in 1928 [30], between tightly-packaged, genetically-silenced DNA (called heterochromatin) and loosely-packaged, actively-expressed DNA (called euchromatin) is now well established [31, 32]. DNA packaging is mediated by direct interactions with specific structural proteins. Double-stranded DNA (dsDNA) is wrapped around a series of histone octamers, each comprising a single **nucleosome**, that self-assemble into higher order complexes, where assembly is mediated by chemical modifications on both DNA and histones that tune inter- and intra-nucleosome interactions [32] (Fig. 1.2). DNA modifications center predominantly around methylation (and variants thereof), while histone modifications are very diverse, including methylation, acetylation, ubiquitination, ADP-ribosylation, sumolation, and phosphorylation, among others. These various modifications dictate whether DNA is condensed into tightly-packed heterochromatin or into more accessible, loosely-packed euchromatin. There is extensive ongoing work, e.g. the Roadmap Epigenomics Project [33], that aims to elucidate the "histone code" [34] or, more broadly, the "epigenetic code" [35], so as to attribute each epigenetic modification (or set of modifications) to specific chromosomal condensation states and then, ultimately, to particular gene expression patterns. Work with the Myc TF, for example, showed that a particular "histone signature" was a strict prerequisite for Myc binding, superseding the DNA sequence specificity of the TF [36]. That is to say, in a general sense, a specific "epigenetic signature" enabled a chromatin structure that provided steric access of TFs and other biomolecules (e.g. transcriptional machinery) to interact with a gene of interest.

We can thus regard the transcription of DNA as being inexorably coupled to its state of compaction. There are many examples that support this relationship. For instance, recent work in mouse embryonic stem cells, looking at the expression of a

reporter integrated at thousands of random sites in the genome, showed that the local chromatin compaction state at the site of integration was predictive of reporter expression [37]. Perhaps the best demonstration of the effect of DNA condensation state on gene expression is a phenomenon known as position-effect variegation, which has been well-studied in *Drosophila melanogaster*, showing that the expression level of a gene of interest depends on its proximity to heterochromatic DNA. In the classic case regarding white eye color [38], Hermann Muller observed that, upon translocation of the *white* gene (without damage to the gene itself) to regions directly adjacent to heterochromatin, flies exhibited red patches in their eyes, demonstrating that, in certain cells, there was definitive silencing of the *white* gene [39, 40]. The gene repression was attributed to "spreading" of the heterochromatic packaging structure to encompass the newly positioned DNA, due to removal of a poorly-understood "structural barrier", upon translocation of the gene, that delimited euchromatin from heterochromatin. Further work identified the specific details of the patchy (or "mottled") eye color phenotype, but the clear result was that, when a gene was artificially positioned near (or within) regions of the genome that are normally densely-packaged as heterochromatin, expression of that gene was repressed.

1.2.2 Prokaryotic chromatin

There is arguably less known about the structural organization of DNA in prokaryotic cells, partially due to the early perception (since corrected) that bacteria lack higher order chromosome organization. Much of what is known comes from work done with *Escherichia coli*. As such, I focus my discussion of prokaryotic chromatin on that model organism. It should be noted that the growing body of research with other organisms (particularly *Bacillus subtilis*, *Caulobacter crescentus*, and *Salmonella enterica*), though revealing clear differences in certain cases, supports the notion that the structural features

observed in *E. coli* are, in a general sense, shared by other bacteria [41, 42, 43].

The *E. coli* genome, called the **nucleoid**, exhibits a contour length of approximately 5 million base pairs (bp); if uncondensed, the diameter of the circular genome would measure ∼0.5 mm. Three basic mechanisms have been attributed towards condensing the genome so that it may fit within a rod-shaped cell that is only about 2 μm in length and 1 μm in diameter: (1) entropic forces associated with DNA polymer dynamics and molecular crowding; (2) supercoiling of DNA; and (3) the interaction of DNA with structure-modifying biomolecules like nucleoid-associated proteins [44]. These mechanisms must compact the chromosome in a coordinated manner that still allows DNA to move dynamically during the various processes vital to the cell (e.g. replication, segregation, and transcription). Similar to the case with eukaryotes, it is only within the past couple decades when notable progress was made towards elucidating the details of chromatin compaction and organization.

The bacterial nucleoid exhibits multiple layers of organization (summarized in Fig. 1.3). It was determined, using a recombination-based method that probes the frequency

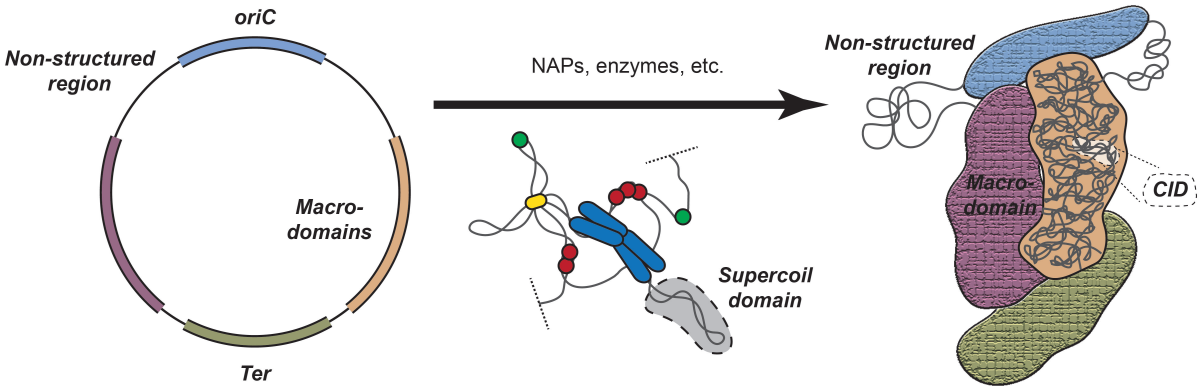


Figure 1.3: **Prokaryotic Chromatin Organization.** Through interaction with structural proteins, including nucleoid-associated proteins (NAPs), DNA is organized into macrodomains that contain many chromosomal interaction domains (CIDs). Within each CID exist many individual supercoil domains (SDs) that are topologically isolated.

with which different DNA regions on the chromosome physically interact, that the circular genome is initially organized into **macrodomains** of ∼1 million bp [45], where the DNA within each macrodomain sterically interacts with DNA in its own macrodomain much more preferentially than with DNA in any of the others. In 2018, chromosome conformation capture results were published that correlate nicely with these macrodomains and also provide evidence for smaller domains (regions of ∼40 to 300 kilobases, kb) within each macrodomain, referred to as **chromosomal interaction domains (CIDs)** [46]. At the lowest level of this nested, matryoshka doll-like scheme of prokaryotic organization, there are a multitude of topologically-independent looped microdomains, often called **supercoil domains (SDs)**, within each CID. Each SD is topologically isolated in the sense that relaxation of supercoiled DNA, due to an interruption of DNA (e.g. a nick), only affects that individual microdomain [47, 43, 46]. SDs are on average ∼10 kb long, but the full distribution of lengths can range between 5x more or less of that average.

To generate these various domains, *E. coli* utilize a wide variety of crosslinking DNA-binding proteins, collectively referred to as nucleoid-associated proteins (NAPs), to modulate chromosomal structure. Each NAP enforces a particular structural effect on DNA, for example: integration host factor (IHF) induces a distinct ∼160° U-turn in dsDNA; factor for inversion stimulation (Fis) bends dsDNA less severely, binding to AT-rich regions and inducing a ∼50-90° turn; heat-unstable nucleoid protein (HU), while also bending dsDNA, exhibits a wrapping behavior, forming multi-component complexes that can coil dsDNA around itself; and histone-like nucleoid structuring protein (H-NS) can oligomerize so as to bridge different DNA strands (thereby forming DNA loops), as well as coat sections of dsDNA, perhaps occluding other proteins from binding [48, 41, 43].

As with eukaryotic cells, structural modifications that alter nucleoid condensation states appear to affect gene transcription. Work with cells expressing a mutant NAP,

specifically HU, showed that the modified NAP not only induced a different genomic structure but that this altered nucleoid organization also led to expression of a different set of genes not seen with the wildtype [49]. These HU mutants also exhibited hindered formation of so called transcriptional factories/foci, which are dense clusters of actively transcribing RNAps that are characteristic of elevated gene expression [50, 51]. Further evidence supporting the transcriptional influence of nucleoid structure is that NAP expression levels are growth-stage dependent: Fis is abundant during log phase while absent in stationary phase; IHF is generally more highly expressed during stationary phase than log phase; and H-NS is expressed at about the same level throughout growth [52, 41]. The different NAP populations presumably induce distinct nucleoid condensation states that lead to different expression patterns during the various stages of growth. Perhaps the clearest (and simplest) piece of evidence that supports the interrelationship of chromatin structure and gene expression is that rapidly growing cells in exponential phase exhibit de-condensed nucleoids while those of stationary phase cells are highly condensed [53]. This dichotomy between exponential and stationary phase cells — the high vs. low transcription states, the de-condensed vs. condensed nucleoids, and the varied NAP populations — parallels the eukaryotic euchromatin/heterochromatin distinction and strongly suggests that NAPs may regulate transcription via their structural influence on the nucleoid. While this idea is now generally accepted, the exact mechanism(s) remain poorly elucidated.

1.2.3 Dynamic chromatin

The growth stage-dependent NAP distribution shows that the bacterial nucleoid is actively modulating its degree of condensation. It nicely illustrates a general property regarding chromatin organization and behavior (for both prokaryotes & eukaryotes): it's

dynamic! Along with the fluctuations stemming from the binding/unbinding of structural proteins (e.g. NAPs or histones), chromatin is acted upon by a vast array of motor proteins during active processes, which include replication, repair, recombination, transcription, translation, and even transertion (i.e. a process of simultaneous transcription, translation, and membrane insertion observed in bacteria). As such, DNA organization should not be thought of as being static and permanent, but rather as innately fluid and adaptive.

The physics associated with transcriptional activity attests to this dynamic description. Protein expression can vary considerably between cells in an isogenic population, as a result of the inherent stochasticity of gene expression [54, 55]. This observation has been attributed to the phenomenon of transcriptional bursting, where mRNA synthesis switches randomly between "bursts" of high transcriptional activity and long periods of inactivity [56]. It was determined that the mechanism behind bursting relies on intracellular gyrase concentration [57]. Chromosomal DNA is negatively supercoiled in the cell, which facilitates local unwinding for RNAp to bind and initiate transcription [58]. When active, RNAp actually adds more supercoiling to its substrate DNA, building up (+) supercoiling in front of the transcriptional machinery and (−) supercoiling behind it [59]. Researchers demonstrated that the accumulation of (+) supercoiling on the working front inhibits transcription, but this inhibition can be removed by the topological relaxation activity of gyrase; with additional analysis, it was concluded that general supercoiling dynamics is the primary origin of transcriptional bursting [57]. Now consider the role of structural proteins. In prokaryotes, for example, it has been demonstrated *in vitro* that certain NAPs, specifically HU, Fis, and H-NS, constrain (−) supercoiling while others, namely IHF and Dps, do not [60]. It seems likely that the particular NAP population at any given time can modulate SD formation within the nucleoid as well as the state of supercoiling within each SD. With these considerations regarding supercoiling, it is

easy to visualize the dynamic nature of the system as well as identify yet another means through which transcription can be dependent on DNA structure & organization.

However, the relationship between transcription and chromatin structure is not completely one-sided. There is indeed bi-directionality in the inter-dependence, where transcription's effect on DNA organization goes beyond simply expressing the structural proteins. In prokaryotes, for instance, transcription is thought to affect chromosome organization through its influence on boundary formation between different CIDs and/or SDs. Experiments have revealed notable variation in the total number of SDs within a nucleoid. Such variability has led many to believe that SDs are transient and may have different characteristics depending on the particular stage of growth or environmental condition [61]. This could be due to the binding/unbinding of topological limiters (e.g. NAPs) that disallow/allow supercoil diffusion. Recent work has suggested that transcriptional activity could be one of these topological limiters. Chromatin conformation capture studies showed that domain boundaries are enriched with highly expressed genes and that artificial positioning and insertion of a highly expressed gene is sufficient to produce a new domain boundary [62]. The hypothesis is that active transcription could open up local chromatin regions, preventing supercoil diffusion [63]. While much of this work was done with *C. crescentus*, recent work with *E. coli* also suggests a notable role played by transcription in domain interactions [46].

In eukaryotic cells, it is becoming increasingly clear that transcription may also affect chromatin organization, particularly through the action of ncRNAs, which were introduced earlier in this chapter. In the "nascent transcript" model, an actively-forming RNA transcript acts as a physical template to recruit heterochromatin assembly machinery, including DNA methyltransferase [18]. RNA can thus play a regulatory role in epigenetic modifications, affecting chromosome structure at the histone level. Certain histone/epigenetic signatures can act as epitopes to recruit chromatin-remodeling complexes

[64]; these chromatin remodelers utilize the hydrolysis of ATP to transiently unwrap DNA from histones and move (or even remove) nucleosomes from DNA, thereby modulating DNA accessibility [31, 65]. Besides considerations concerning RNA-dependent chromatin-remodeling machinery, transcriptional activity alone appears enough to induce distinct organizational changes. Transcription of a gene has been shown to trigger relocation of that gene from the periphery to the interior of a nucleus [66, 67, 68]. Additionally, in the context of imprinting (where a cell selectively expresses either the maternal or paternal copy of a gene, while silencing the other), it was shown that just the act of transcribing the *Airn* gene, and not any function attributable to the Airn RNA transcript itself, was sufficient to silence the paternal allele of the Igf2r gene, through a mechanism related to occlusion of the Igf2r promoter during *Airn* transcription [69].

These examples work to demonstrate the complex inter-dependence of transcription and chromosome organization/re-organization. The relationship remains actively studied today. In fact, a research consortia — called the 4D Nucleome Project/Network [70] — was recently started, bridging researchers together from diverse disciplines, ranging from microscopy to genomics to computer science to molecular biology to physics, with a shared goal:

> *To map the structures and dynamics of the genome in space & time...and to relate these features to its biological activities.*

1.3 Biology *In Vitro*

This far from exhaustive review illustrates the immense complexity of the biological cell, demonstrating the challenge of definitively identifying causal relationships between chromatin structure and gene expression (and vice versa). The plethora of interacting components confounds the attribution of particular effects to specific actors. *In vitro*

approaches, which enable direct control of reaction components and conditions, provide more tractable means to untangle biological intricacies and to explicitly identify & confirm distinct effects.

The gold standard in biological research is to prove your hypothesis *in vivo*, or within living systems. However, work that is performed *in vitro*, or "in the glass", has also proven worthwhile (if not integral) towards gaining a full understanding of biology. The term is quite inclusive, encompassing any work performed outside of a living cell or, more strictly, outside of a living organism. Though not fully representative of the complexity of living systems, the simplicity of *in vitro* systems enables a level of control and scientific clarity that is difficult (or perhaps impossible) to achieve *in vivo*.

The impact of *in vitro* research is perhaps best embodied by the history of the "Central Dogma" and elucidation of the genetic code. Several pivotal contributions are highlighted below:

- Nucleic acid synthesis: Arthur Kornberg and colleagues purified DNA polymerase from *E. coli*, and showed that the enzyme, under specific conditions, could accurately assemble long strands of DNA [71, 72, 73]. Work led by Severo Ochoa led to the purification of polynucleotide phosphorylase, an enzyme enabling the accurate construction of RNA-like polymers [74, 75].

- RNA to Protein: Marshall Nirenberg and colleagues showed that a poly-uracil sequence of RNA codes for the synthesis of a phenylalanine polypeptide chain [76]. They were the first to clearly demonstrate the role played by mRNA in coordinating protein synthesis and contributed greatly to deciphering the full genetic code.

- Reading direction: Building off his work with polynucleotide phosporylase, Ochoa and colleagues demonstrated that short poly-A fragments ending with a cytidine

at the 3' end led to synthesis of lysine polypeptides with a carboxy-terminal asparagine, demonstrating that RNA is read in the 5' to 3' direction [77, 78].

- Transcriptional regulation: Using yeast (i.e. *Saccharomyces cerevisiae*) as their primary model [79], Roger Kornberg and colleagues, in addition to elucidating the crystal structure of RNA polymerase II, determined the importance and roles of many transcription factors & regulators of eukaryotic transcription [80].

All of these scientists received Nobel prizes for their contributions to biological understanding. Notably, ALL of this work was performed *in vitro*, using custom cell-free gene expression systems (CFGE), where precise control of experimental conditions enabled the scientific discovery.

The interrelationship of DNA organizational state and genetic output can be studied through similar means. By modifying CFGE systems so that the gene-encoding DNA templates are integrated into higher order nucleic acid complexes (akin to chromatin), scientists can tractably address, as stated by the 4D Nucleome Project/Network, the...

> *...lack (of) mechanistic insights into the relationships between chromosome conformation and nuclear processes, including transcription, DNA replication, and chromosome segregation.*

Reconstituted model systems, though undoubtedly simplistic compared to actual chromatin, mimic chromosomal states and allow researchers to probe the organization-expression relationship in a more systematic and controlled manner, enabling research that can potentially separate structural effects from genetic ones or even distinct structural effects from each other.

1.4 Research Overview

In this dissertation, I present my work towards developing a simple *in vitro* reconstituted DNA platform to examine how the nucleic acid environment around genes can affect transcriptional activity, and vice versa. In **chapter 2**, I introduce the integral structural component of the system, namely a DNA nanostar, and describe interesting material features associated with condensed phases of self-assembled NSs, focusing primarily on the aging dynamics of NS-networks. In **chapter 3**, I spotlight the NS-liquid system, characterizing notable physical properties of the liquid and interpreting their dependence on DNA hybridization. In **chapter 4**, I briefly review the significance of liquid-liquid phase separation in biology and propose the use of the NS-liquid as an experimental model of "membraneless organelles", describing the factors that govern biomolecular partitioning into the system. In **chapter 5**, I describe our initial work examining how the integration of genes into a NS-liquid affects transcription and, inversely, how mRNA products restructure the DNA environment from which they are transcribed. In **chapter 6**, I conclude the dissertation by discussing how research with this reconstituted platform could proceed in the future and present notable preliminary results. The appendices contain more detailed descriptions of the materials & methods used in this work as well as a table of abbreviations used in this dissertation.

1.5 Permissions and Attributions

All the work described in the following chapters was done in collaboration with Professor Omar A. Saleh.

- Chapter 2 and Appendix A contain content that has previously appeared in the journal *Soft Matter* [81]. It is reproduced and adapted here with the permission of the Royal Society of Chemistry.

- Chapter 3 and Appendix A contain content that is the result of a collaboration with Dr. Byoung-jin Jeon, Gabrielle Abraham, Nate Conrad, and Professor Deborah K. Fygenson; and has previously appeared in the journal *Soft Matter* [82]. It is reproduced and adapted here with the permission of the Royal Society of Chemistry.

- Chapters 4 & 5 and Appendix A contain content that is the result of a collaboration with Dr. Byoung-jin Jeon, Greg Ekberg, and Professor Christopher S. Hayes.

- Appendix A also contains content that is the result of a collaboration with Dr. Chang-young Park, Dr. David R. Jacobson, and Sam Willardson; and has previously appeared in the journal *Review of Scientific Instruments* [83].

Chapter 2

Self-assembling DNA nanostars

2.1 DNA as a material

The ability of DNA to hybridize in a specific, sequence-dependent manner has led it to be embraced as a programmable structural material, particularly for self-assembling systems [84, 85]. Exploitation of DNA's sensitive tunability has enabled construction of complex nanostructures ranging from simple branched junctions to sophisticated 3D nanocages carrying therapeutic cargo [86, 87, 85, 88]. Through strategic placement of single-stranded DNA (ssDNA) overhangs, nanostructures can be further engineered to assemble into large-scale macrostructures. These materials are commonly called DNA hydrogels as they almost universally contain water, are comprised predominantly of crosslinked DNA, and exploit the biopolymer as a structural scaffold.

DNA hydrogels have been constructed for a variety of applications, including controlled drug delivery, gene therapy, tissue engineering, protein production, bio-sensing, and bio-capture [89, 90, 91, 92, 93]. Some systems have even been tailored to be environmentally responsive, enabling DNA hydrogels to, for example, degrade upon exposure to specific external stimuli [93, 94, 95, 96, 97, 98, 99]. In this dissertation, I describe our

work to expand the list of DNA hydrogel applications to include its use as a reconstituted model system of chromatin. Specifically, we construct DNA hydrogels through the self-assembly of multi-armed synthetic DNA nanostars (NSs) with palindromic overhangs (Fig. 2.1) [100]. The system's simplicity, versatility, and intrinsic biocompatibility enable it to be a practical core component of an *in vitro* chromatin mimic.

In our efforts to develop the system to probe biologically relevant questions, we discovered interesting features and properties that are of interest to materials research more generally. The hydrogel state is intrinsically a non-equilibrium, kinetically trapped state, yet discussions of non-equilibrium structural phenomena of DNA hydrogels often do not go beyond descriptions of the gel-to-solution transitions exhibited by the aforementioned environmentally-responsive systems. That is, there has been little careful consideration of how the path-dependent final structure is guided by molecular design. In this chapter, we examine the aging kinetics of DNA hydrogels comprised entirely of self-assembling DNA NSs. In this simple system, we first identify that placing unpaired bases at overhangs determines whether NSs form an arrested, solid-like percolated network or an equilibrium, phase-separated liquid. This result may have biological implications, as increasingly more evidence shows the existence, and the presumed functionality, of nucleic acid complexes within living cells undergoing liquid-liquid phase separations [101, 102]. We then show that the network phase exhibits aging in the form of global hydrogel contraction towards a denser state, and that the contraction rate increases with weaker overhang hybridization strength. Finally, we demonstrate that global contraction is ion-specific, where the phenomenon is retarded in the presence of divalent cations. Ultimately, we show that hydrogel structure can be controlled by tuning the hydrogel aging kinetics, specifically by modulating the hybridization lifetime through the inclusion of single-stranded gaps that disrupt base-stacking, the inclusion of unpaired bases that increase NS conformational entropy, and the variation of salt & temperature conditions of the system.

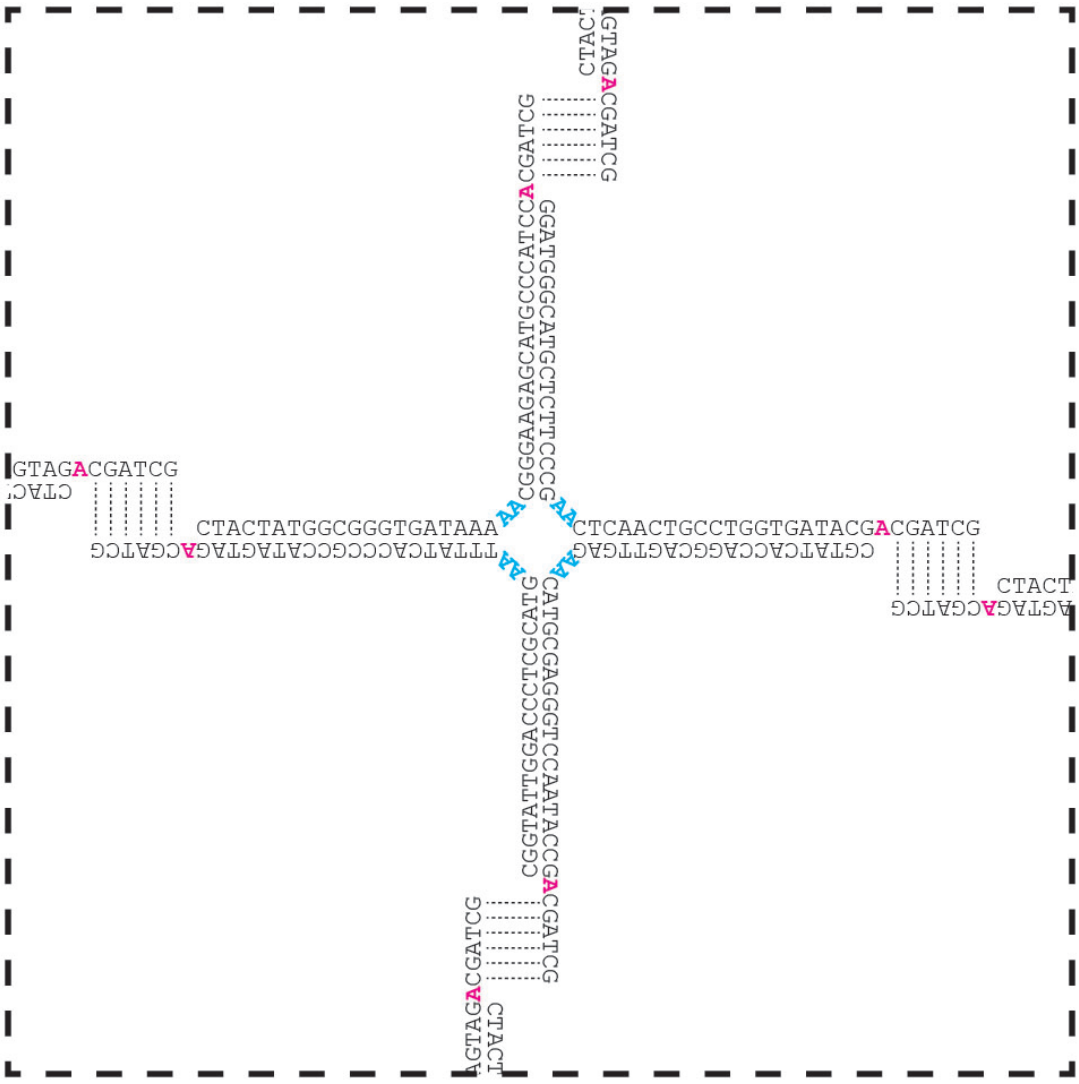


Figure 2.1: **DNA nanostar structure.** A DNA NS exhibits multiple dsDNA arms that each terminates with a palindromic overhang that enables hybridization with other NSs. Dotted lines indicate base-pairing. Shown is a NS with four arms, 20 bp in length, with 3' overhangs, where the palindromic region is 6 nt in length. NSs can be modified through the inclusion of additional bases, which are designed to remain unpaired, at the NS joint (cyan) and/or at each overhang (pink), preceding the palindromic regions.

2.2 Microscopic details determine phase

DNA hydrogels were constructed from tetravalent DNA NSs with four different molecular designs. As in previous methods [89, 100], individual NSs were self-assembled from four complementary ssDNA oligomers via a gradual temperature annealing scheme. Efficient NS formation was confirmed using polyacrylamide gel electrophoresis (PAGE) (Fig. A.1). Each NS exhibits dsDNA arms that are 20 bp in length and that terminate with a 3' overhang containing a 6 nt palindromic region. The NS designs differ with regards to the presence of additional unpaired bases at two distinct sites: one adenosine at the 5' end of each overhang and/or two adenosines between each NS arm at the joint (Fig. 2.1); these modifications increase the NS conformational entropy by allowing flexibility and rotation [100]. We thus examine four distinct NS architectures that we term fully-stiff (no unpaired bases at either site), joint-flexible (unpaired bases only at the joints), overhang-flexible (unpaired bases only at the overhangs), and fully-flexible (unpaired bases at both sites; Fig. 2.2a). NS samples were prepared at 10 μM concentrations in 10 mM Tris buffer with 10 μM of YOYO-1 fluorescent dye (i.e. 1:1 dye to NS ratio) and varying concentrations of NaCl. Samples were mixed via pipette and immediately wicked into cleaned glass capillary tubes for observation via fluorescence microscopy at 488 nm excitation. Experiments employing an internal fluorescein modification, in lieu of the YOYO-1 dye, exhibited qualitatively similar results, demonstrating that YOYO-1 has a minimal effect on the observed phenomena at these concentrations. Note that the methods described in this and the following chapters can be found in greater technical detail in Appendix A.

NSs self-assemble into macromolecular structures via overhang hybridization when the ionic strength of the solution is sufficient to screen the repulsive interaction between negatively-charged DNA strands. Under such conditions, the bond lifetime between NS

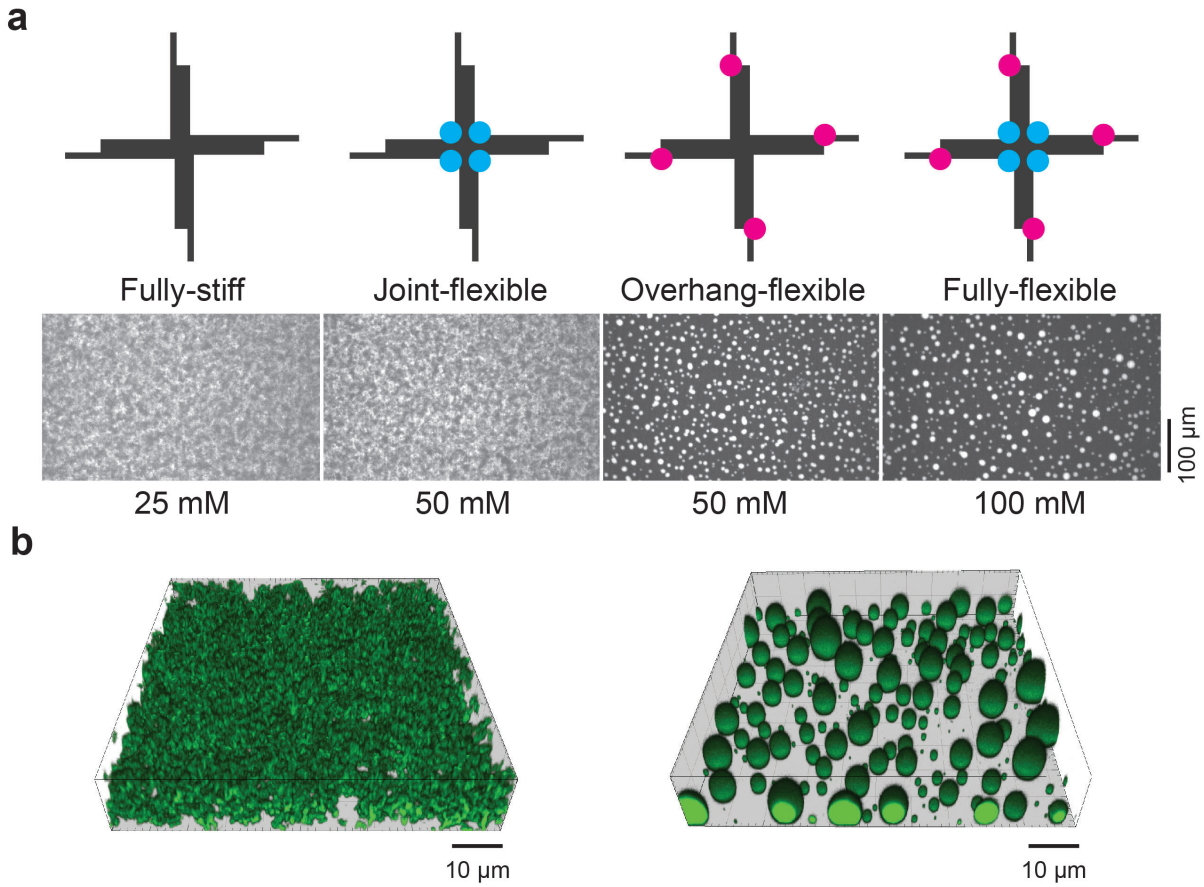


Figure 2.2: **Macromolecular phase.** All mixtures contain a 1:1 ratio of NS:YOYO-1 dye for fluorescence imaging with 488 nm excitation. **(a)** Each of the four different NSs was incubated with NaCl for 1 hour at room temperature. Grouped in columns are the NS design, the resulting self-assembled macromolecular structure, and the NaCl concentration used during incubation. **(b)** Confocal microscopy was performed on fully-stiff (left) and fully-flexible NS (right) samples

arms is sufficiently long to enable the formation of condensed NS phases. We confirm that NSs only interact when arms exhibit complementary overhangs: blunt-armed NSs remain in solution even at 1 M NaCl. From the four different NS architectures, all of which share the same sole means of self-assembly via hybridization of the palindromic overhang sequence, we observe two distinct macromolecular structures. The phase-determining feature is overhang flexibility. NSs lacking an unpaired overhang base — fully-stiff and joint-flexible NSs — form percolated networks that span the entire sample volume (Fig.

2.2a & b left), after 1 hour of incubation at room temperature with sufficient NaCl. In contrast, NSs exhibiting an unpaired overhang base — overhang-flexible and fully-flexible NSs — phase separate from water to form dense liquid droplets (Fig. 2.2a & b right). These DNA droplets sediment, become more spherical over time, and can coalesce, affirming their fluidic nature. Thus, the overhang flexibility, specifically the presence of single-stranded gaps after hybridization that remove additional stabilization due to base-stacking (Fig. 2.1), is a salient design parameter when constructing DNA hydrogels for specific applications, since its presence controls the final phase of the material.

2.3 Hybridization strength modulates network aging

DNA networks formed by fully-stiff and joint-flexible NSs exhibit time-dependent coarsening, similar to protein hydrogel systems [103], where the average pore size of a hydrogel increases with time. This coarsening is associated with the local contraction of globules within the network that straightens intervening connections, opening the pores (Fig. 2.3a). Indeed, when the non-ionic surfactant Tween® 20 is added, minimizing non-specific interactions with glass surfaces, local globule contractions drive global hydrogel contraction, where the tube-like network reduces its diameter over time. Contraction dynamics are dependent on both the ionic strength and the temperature: lower NaCl concentrations and higher temperatures (i.e. conditions that weaken base pairing interactions) induce faster contraction towards a denser hydrogel state (Fig. 2.3b & c). Coarsening and contraction only occur when NS interactions are non-covalent. If hybridized arms are covalently linked using enzymatic ligation, neither process is observed, suggesting that the reversible hybridization of palindromic overhangs is essential for these phenomena.

The dual-dependence on temperature and ionic strength suggests that hydrogel con-

traction is controlled by the hybridization strength between NS overhangs. We thus posit that contraction is a result of the dynamic rearrangement of overhang hybridization partners between interacting NSs. Upon formation of a kinetically trapped hydrogel, the majority of NSs will not have all arms hybridized. Additionally, the orientation of hybridized arms of an individual NS may constrain and stretch that NS in unfavorable

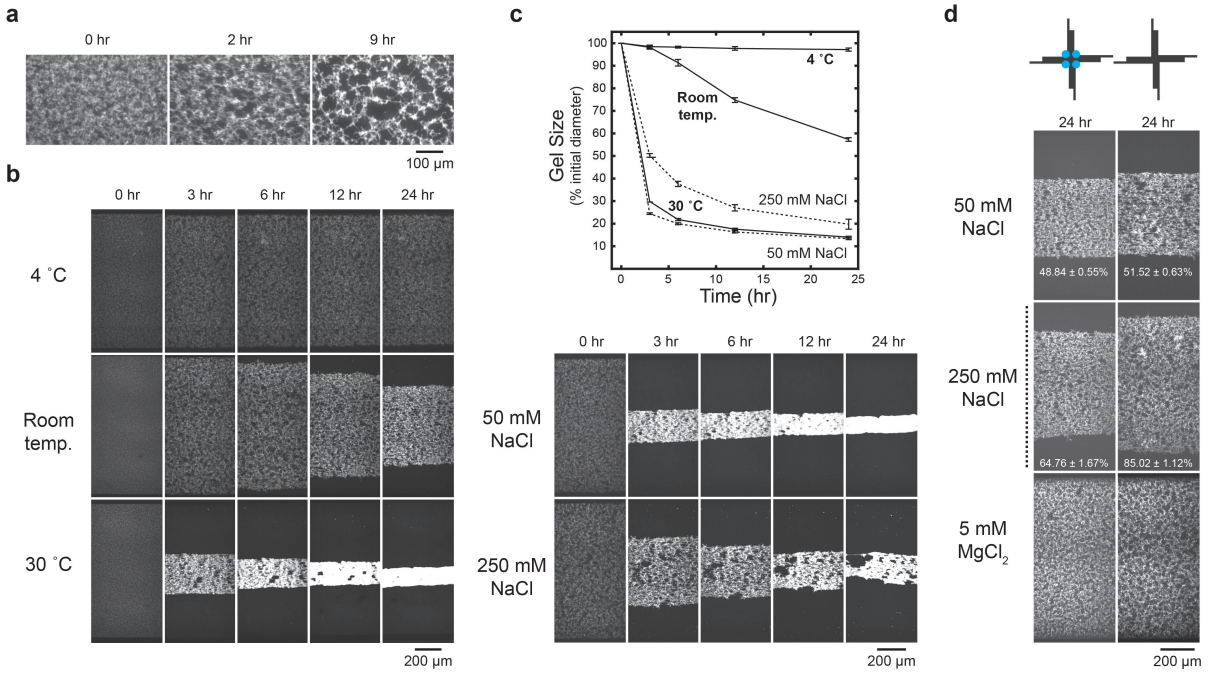


Figure 2.3: **Hydrogel aging is dependent on ionic strength and temperature.** All samples contain a 1:1 ratio of NS:YOYO-1 dye for fluorescence microscopy at 488 nm excitation. **(a)** Time series of a fully-stiff NS network in 50 mM NaCl at room temperature without Tween® 20. (b-d) Time series of fully-stiff NS networks with 0.1% v/v Tween® 20, where gels at 0 h span the entire width (top to bottom) of the field of view. **(b)** Temperature effect. Hydrogel aging in 50 mM NaCl at different temperatures (left). Salt effect. Hydrogel aging in different NaCl concentrations at 30 °C (right). **(c)** The change in gel diameter, as a percentage of initial size, of each image series shown in (b) as a function of time, in hours. Error bars refer to the standard deviation of gel diameter over the length of a gel. **(d)** Aging of joint-flexible NSs (left panels) and fully-stiff NSs (right panels) in different salt concentrations after 24 hours of incubation at room temperature. The dashed line indicates the initial diameter of all gel samples. Inset values describe the final gel diameters as a percentage of the initial diameter with standard deviation across the length of each gel.

conformations. By rearranging hybridization partners, a hydrogel can approach a more favorable state where a larger fraction of its NSs have all (or most) of their arms bound, and where those interactions minimize the conformational strain on each NS. As the rate of rearrangement depends on bond lifetimes between hybridized arms, conditions that weaken DNA hybridization would enable faster hydrogel contraction, as observed.

Joint-flexible NS gels contract faster than fully-stiff NS gels, under the same conditions. After 24 hours at room temperature, the diameters of contracted joint-flexible NS gels are smaller than those of fully-stiff NS gels, regardless of the NaCl concentration (Fig. 2.3d top & middle). Therefore, in addition to solvent conditions, the internal flexibility of NSs also regulates the contraction dynamics. This corroborates the importance of the dynamic hybridization rearrangement to contraction. On a microscopic level, it is possible that joint flexibility acts to shorten the NS bond lifetime, as the flexibility conferred by unpaired bases at the joint increases the conformational entropy of an unbound arm, which competes energetically with arm hybridization. Alternatively, or in addition, joint-flexible NSs could undergo faster rearrangement dynamics, since the arms are not as rigidly constrained, permitting faster exploration of their surroundings for new binding partners.

When $MgCl_2$ is used in lieu of NaCl for stabilizing DNA hybridization, contraction dynamics are dramatically slowed. 5 mM $MgCl_2$ exhibits a similar predicted binding free energy [104] at 25 °C, -6.2 kcal mol^{-1}, as 250 mM NaCl for the 6 nt NS overhang. However, neither fully-stiff NS gels nor joint-flexible NS gels contract after 24 hours of incubation with 5 mM $MgCl_2$ at room temperature (Fig. 2.3d bottom), though some coarsening is observed. This directly contrasts with the contraction seen with networks in 250 mM NaCl (Fig. 2.3d middle), suggesting that a specific interaction between $MgCl_2$ and NSs, not shared by NaCl, retards aging dynamics.

2.3.1 Contraction relies on gel-embedded NSs

We performed fluorescence recovery after photobleaching (FRAP) experiments with hydrogels comprised of fully-stiff NSs, covalently labeled with fluorescein, to probe the molecular mechanisms underlying network dynamics (see Appendix A.2 for more technical details). As a control, we first bleached enzymatically-ligated gels and observed no fluorescence recovery, confirming the importance of transient non-covalent interactions to the aging phenomena. Then, we bleached fully-stiff NS gels in 250 mM NaCl. Though gel networks under these conditions display contractile aging on a time scale of hours (Fig. 2.3d middle right), fluorescence recovery is visually discernible within minutes (Fig. 2.4a), indicating the presence of a rapid NS exchange process that is unrelated to

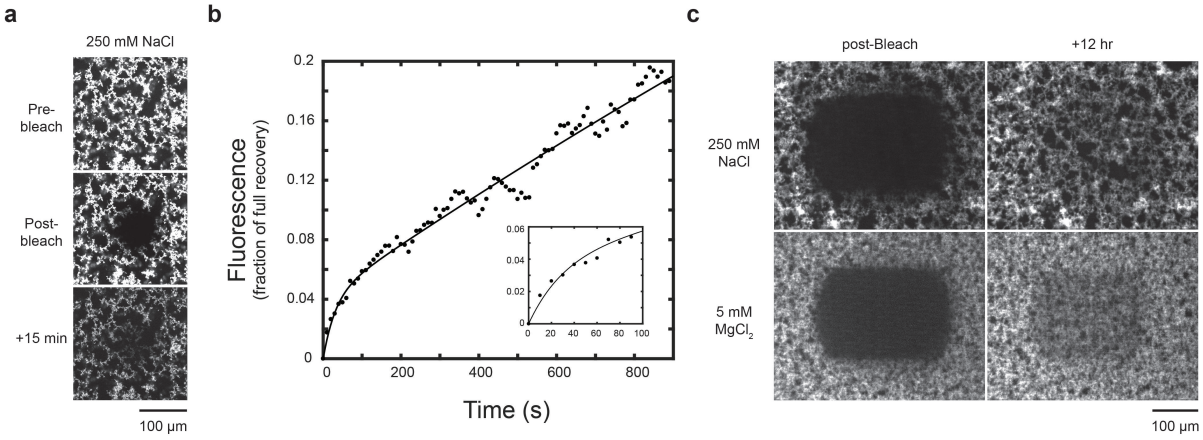


Figure 2.4: **Fluorescent Recovery After Photobleaching (FRAP). (a)** Representative time series of FRAP for fully-stiff NS networks. Fluorescein-labeled NSs in 250 mM NaCl were bleached with 488 nm light and imaged over time at room temperature. **(b)** Representative recovery curve. The data are well-fit by a double exponential equation $f(t) = a(1-e^{-t/b})+(1-a)(1-e^{-t/c})$, where a indicates the fraction of gel recovering with time constant b, and $(1-a)$ is the fraction recovering with time constant c. Across three data sets, the best-fit parameters are $a = 0.040 \pm 0.002$, $b = 31 \pm 11$ s, and $c = 4700 \pm 600$ s, each reported as mean $\pm$ standard deviation. The inset highlights the short time regime. **(c)** Even after 12 hours, the bleached spot is still visible in both 250 mM NaCl and 5 mM $MgCl_2$; this supports the slow dynamics of the gel/gel exchange process.

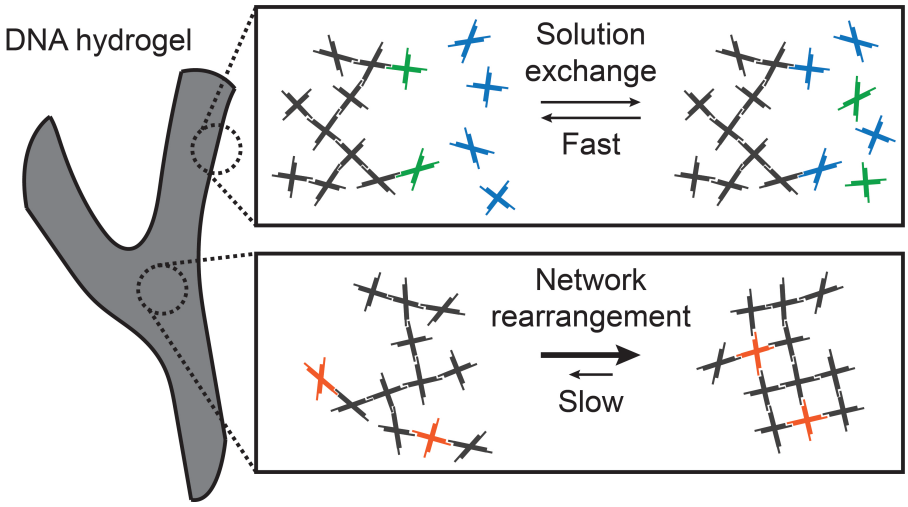


Figure 2.5: **Model of NS exchange.** One exchange process involves fast exchange with solution-borne NSs (top) while the other involves NSs rearranging hybridization partners within the network without ever fully releasing from the gel (bottom).

the slow aging. This process is also seen with $MgCl_2$-stabilized gels. The fluorescence recovery within bleached regions is well fit with a double exponential equation, indicating two separate recovery processes that exhibit time scales on the order of seconds and hours, respectively (Fig. 2.4b). To rationalize this FRAP data with the contractile aging results, we posit the existence of two dynamic exchange processes that underlie aging. In the first, a solution-borne NS binds to the gel, or vice versa (Fig. 2.5c top); this sol/gel exchange is rapid and occurs for all gels, but does not lead to contraction. Our analysis suggests that a small fraction of the NS population (approximately 4%) is involved in this fast process, consistent with the idea of a surface-dependent exchange. In the second exchange process, a NS moves within the gel, swapping binding partners but never fully releasing from the network (Fig. 2.5c bottom). This gel/gel exchange is relatively slow, but highly sensitive to salts and temperature, and, critically, is the process that leads to structural rearrangement. We see that, even after 12 hours of recovery, the bleached regions are still identifiable (Fig. 2.4c), indicating that bleached NSs buried within the gel have not fully exchanged. This picture clarifies that it is the mobility and exchange

of NSs within the solid phase that drives contraction.

2.3.2 Salt-specific contraction

Using flow cells with permeable membranes, we can observe the response of a single gel system to controlled changes under salt conditions with time. Fully-stiff NSs were prepared in flow cells containing permeable membranes [83], comprised of crosslinked polyethylene glycol-diacrylate (PEG-diacrylate), that separate a central sample channel from two buffer exchange channels (Fig. 2.6) (see Appendix A.2 for more technical details). The membranes enable the exchange of the salt while sequestering NSs to the sample channel. After transferring high salt solutions (e.g. 50 mM NaCl) into the exchange channels via a syringe pump, salt equilibration into the sample channel occurs in ∼5 min at room temperature, inducing complete network percolation within ∼30 min. The resulting hydrogel exhibits dense edges since it is those regions near the permeable

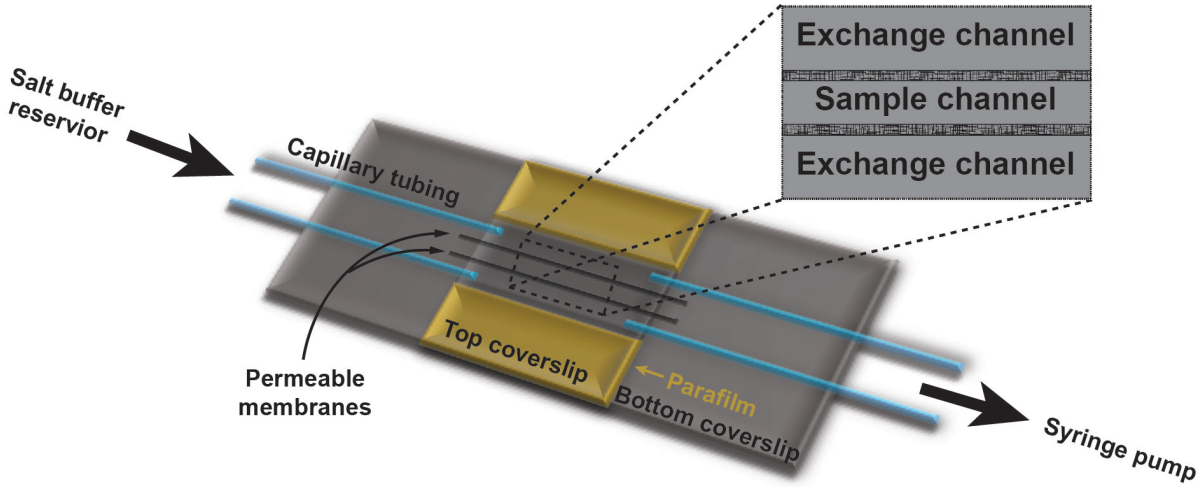


Figure 2.6: **Permeable membrane flow cell.** Within a flow cell constructed by melting parafilm between two glass coverslips, a central channel is separated from two buffer exchange channels by two PEG-diacrylate permeable membranes. Thin capillary tubes are inserted into the exchange channels and connected to a buffer reservoir and a syringe pump, enabling salt exchange to the system. NS samples within the central sample channel are unperturbed by flow in the exchange channels.

membranes that first encounter the diffusing salt (Fig. 2.7a left). In a direct reflection of the capillary tube results, exchange towards lower salt conditions (i.e. 15 mM NaCl) induces a rapid global contraction (Fig. 2.7a middle left), where a change typically observed after 12 hours at 50 mM NaCl is seen after only ∼1 h. This contraction is irreversible: when salt concentrations are returned to 50 mM NaCl, new networks form over the already contracted system (Fig. 2.7a middle right). Subsequent exposure to 15 mM NaCl induces further contraction, even to the point of tearing the hydrogel (Fig. 2.7a right).

Hydrogel contraction in response to a decreased salt concentration, as observed in the permeable-membrane experiments, does not follow the expectation from simple electrostatic considerations. When a charged hydrogel is moved from higher to lower salt conditions, the reduced screening of the electrostatic repulsion between the charged network strands should induce the gel to reversibly expand. Covalently-crosslinked DNA hydrogels behave in this manner [105]. Our non-covalent NS systems do exhibit transient expansion, just before dissolution, when salt concentrations are reduced to concentrations below that which sustains network percolation, e.g. 0 mM NaCl. However, when the salt reduction is to concentrations that remain network-stabilizing, e.g. 15 mM NaCl, we observe irreversible contraction (Fig. 2.7a). The non-covalent nature of the system thus enables a hydrogel response that is not shown by covalently-crosslinked DNA systems.

When $MgCl_2$ is used in lieu of NaCl for ionic strength modulation in permeable membrane experiments with fully-stiff NS, the gels exhibit reversible expansion when moved from high to low $MgCl_2$ concentrations (Fig. 2.7b), contrasting the irreversible contraction seen with NaCl but paralleling the behavior of covalently-crosslinked DNA hydrogels. This improved network stability, i.e. decreased overhang hybridization rearrangement dynamics, agrees with the minimal contraction observed in capillary tube experiments (Fig. 2.3d bottom). It is well established that Holliday junctions, i.e. bio-

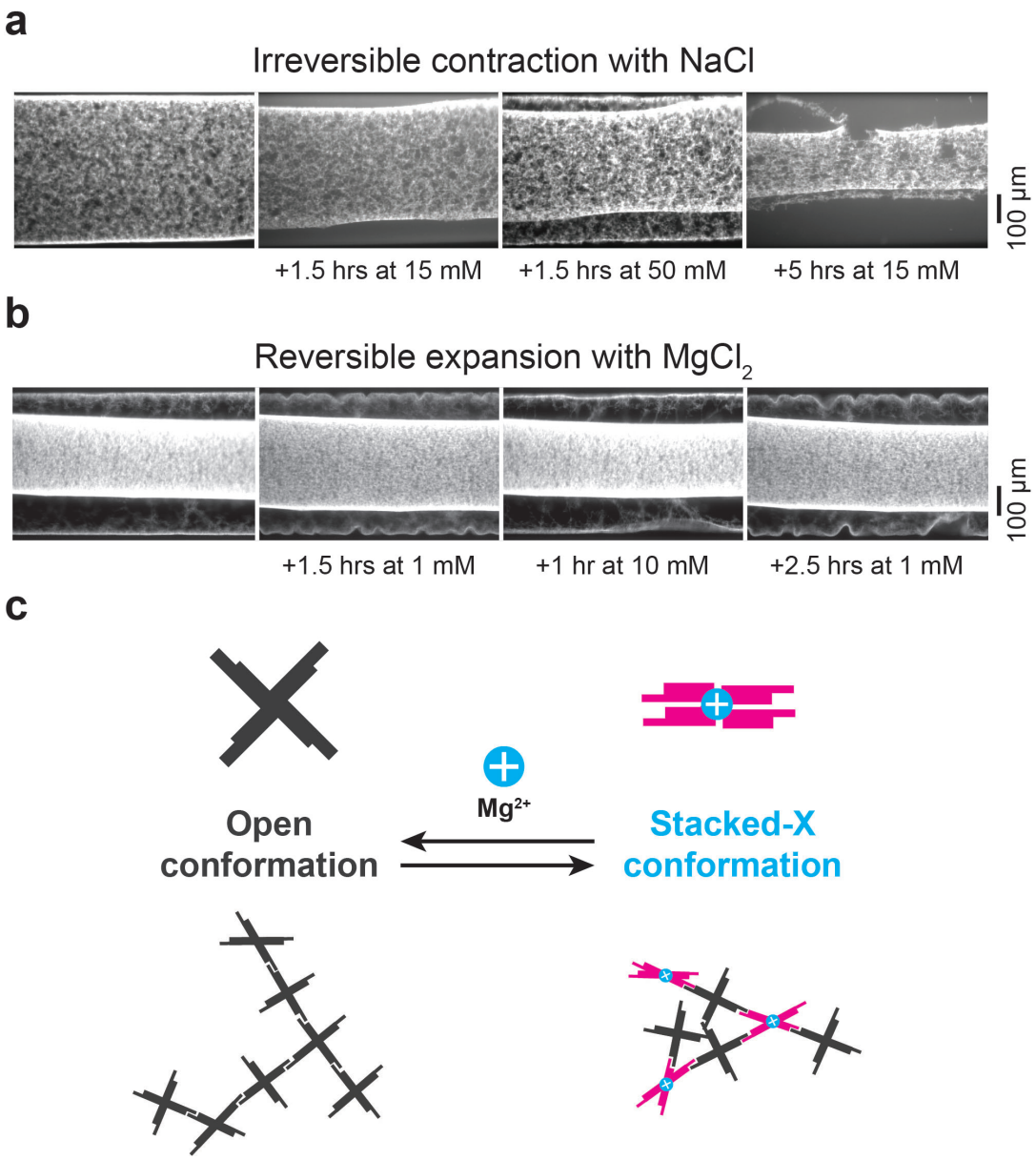


Figure 2.7: **Salt-specific contraction.** Hydrogels exhibit irreversible contraction and reversible expansion upon decreasing concentrations of NaCl and $MgCl_2$, respectively. (a and b) Permeable membrane flow cells. **(a)** A network was first formed through exchange of 50 mM NaCl (left). The gel was then exposed to different NaCl conditions; the concentrations and incubation periods are noted below each panel. **(b)** A contracted network was formed through initial exposure to 50 mM NaCl and subsequent equilibration with 10 mM $MgCl_2$ (left). The gel was then exposed to different $MgCl_2$ conditions; the concentrations and incubation periods are noted below each panel. **(c)** Synthetic Holliday junctions (i.e. tetravalent NSs) are known to adopt open, planar conformations in the absence of Mg^{2+} ions and then switch to stacked-X conformations in their presence [106].

logically relevant tetravalent DNA NSs, switch from flexible open conformations to rigid stacked-X conformations upon Mg^{2+} binding (Fig. 2.7c top) [106]. We posit that such a conformational constraint, i.e. internal rigidification of NSs, would reduce the mobility of NS arms, hinder NS rearrangement dynamics, and thereby retard gel contraction. In permeable membrane experiments, this conformational switching could also induce reversible expansion (Fig. 2.7c bottom). We probed the conformational switch using PAGE, comparing the degree of migration in the presence and absence of added $MgCl_2$. Relative to the ladder, both stiff and flexible NSs migrate further with added $MgCl_2$ than without (Fig. A.1), supporting the notion that NSs can adopt the more compact stacked-X conformation. Although more work is needed to isolate and probe the effect of stacked-X conformational switching on global gel behavior, our results ultimately demonstrate that DNA NSs exhibit an ion-specificity that enables them to be suitable components in 'smart', chemically specific, responsive hydrogels.

This work demonstrates that the non-equilibrium kinetic behavior of DNA hydrogels has a significant effect on the hydrogel structure and, as a result, is a programmable parameter when constructing hydrogels for various applications. These aging kinetics can be rationally tuned through careful design of both the self-assembling components and the solution conditions, all of which modulate the dynamics of NS binding partner rearrangement. NS are also able to confer ion-specificity to hydrogel behavior, giving them potential as functional elements in environmentally-responsive DNA hydrogels. As NSs and Holliday junctions are increasingly common structural components of DNA hydrogels and of DNA origami nanoassemblies [107], an improved understanding of their interaction kinetics will enable the field to better tailor DNA systems for specific applications and may also inspire novel applications for DNA nanotechnology.

Chapter 3

Characterization of DNA liquids

NSs self-assemble into percolated networks when the hybridization strength between interacting overhangs is strong enough to support network growth. However, as mentioned in the previous chapter, if that interaction is tempered through inclusion of an unpaired base at each overhang (Fig. 2.1), thereby perturbing base-stacking interactions and weakening the binding free energy, then NSs form dense liquid droplets that phase separate from aqueous solution (Fig. 2.2). While this π-stacking perturbation nicely illustrates the NS phase dichotomy, any means of weakening NS interactions (e.g. shorter overhang sequence), while still allowing sufficiently strong binding to support NS self-assembly, can lead to adoption of the dense liquid phase. This chapter describes our work towards characterizing the physical properties of these DNA NS-liquids.

3.1 Liquid-liquid phase separation

This NS-liquid system nicely illustrates the phenomenon of liquid-liquid phase separation (LLPS), where macromolecular components (often polymers) form a macromolecule-rich phase in equilibrium with a macromolecule-dilute aqueous phase. Phase separation

is frequently driven by electrostatic attraction between oppositely charged biopolymers, a process termed complex coacervation. This process has been the subject of intense study in part because of its variety of potential (and often profitable) applications, which range from drug delivery to food processing to cosmetics [108, 109, 110, 111]. More recently, electrostatically-driven biomolecular LLPS has been identified in biological contexts, where the resulting droplets form within cells, creating biochemically-distinct spaces that are typically termed membraneless organelles [112, 113], which will be discussed further in chapter 4.

Typical coacervate systems involve either multiple components, a wide range of molecular configurations, or a corresponding variety of types of intermolecular interactions, and combinations thereof. This complexity has hindered both understanding of microscopic coacervate structure, and examination of coacervate properties. In contrast, the DNA NS system involves only a single type of self-attractive particle and has interactions dominated by well-characterized DNA basepairing interactions; thus, we suggest the NS system can serve as a useful model for LLPS properties and behavior. To that end, we carry out a rigorous investigation of the salt-dependent density, viscosity, self-diffusivity, and surface tension of a NS-liquid. These parameters exhibit a contrary response to salt compared to typical coacervates, due to the dominant role played by salt-induced electrostatic screening in DNA hybridization. Indeed, quantitative modeling shows that hybridization thermodynamics can explain many of the measured properties; yet, effects beyond hybridization are needed to explain the measured variation in density, the low surface tension, and the observed breakdown of the Stokes-Einstein relation. Altogether, our data implies that the NS-liquid has a heterogeneous, clustered structure. More broadly, NS-liquids exhibit low densities, high viscosities, and anomalous structural properties; these distinct qualities have implications for other macromolecular LLPS systems, as well as indicating the potential of the NS system as a substrate for biomolecular engineering.

3.2 Salt-dependent physical properties

Fully-flexible NSs exhibiting unpaired adenosines at both the NS joint and overhangs were used to form NS-liquids. In contrast to the previous chapter, the NSs used here exhibit 5' ssDNA overhangs with the sequence 5'–*CGATCGA*–3', where the palindromic *CGATCG* portion constitutes a sticky-end that permits NS-NS binding, while the 3' *A* is left unpaired (Fig. 3.1 left).

The NS-liquid phase is expected to be salt-sensitive, due to charge screening effects on the repulsion between negatively-charged phosphate groups. Screening affects sticky-end hybridization as well as the electrostatic repulsion between NSs and between arms within a single NS. To investigate this, we first examined how changes in salt concentration affect the density of DNA within the NS-liquid. With sufficient salt, and at low temperatures, NSs interact to form DNA-rich, spherical NS-liquid droplets within a

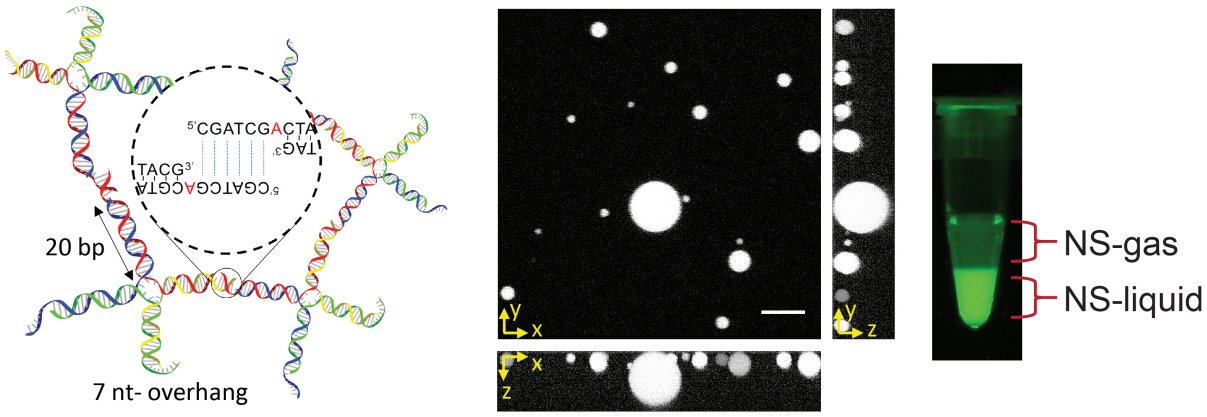


Figure 3.1: **NS-liquid droplets**. Schematic of individual NS constructs comprised of four ssDNA oligos (each strand is a distinct color) and of the formation of a NS-liquid via DNA hybridization between NS overhangs. Center: Confocal image of phase separated NS droplets in 0.5 M NaCl at 20 °C. A mixture of untagged and Cy3-tagged NSs (99:1) was used for fluorescent visualization at 561 nm excitation. XZ(right) and YZ(bottom) projections show the spherical shape of the droplets (scale bar, 20 μm). Right: Bulk phase separation of ~100 μL of NS solution at 20 °C. NSs are visualized by adding YOYO-1 at 1:100 dye:NS molar ratio.

DNA-dilute aqueous 'NS-gas' (Fig. 3.1 right). The two phases can be separated using centrifugation and extracted for analysis with absorbance spectroscopy. As expected for a phase equilibrium, the DNA concentration in the NS-liquid phase is independent of the overall NS concentration (Fig. 3.2). The DNA concentration of the NS-liquid phase increases from 19.2 (±1.8) to 35.1 (±1.1) mg/mL at 20 °C, corresponding to DNA volume fractions $\phi = 0.011$ to 0.021, as [NaCl] increases from 0.25 to 1 M (Table 3.1). Alternate density estimates, based on sedimentation velocities of isolated droplets (Fig. A.6), are consistent with these concentrations, indicating that neither the centrifugation process nor the transformation of NS-liquid droplets into a bulk phase affects NS-liquid density.

We measured the viscosity of NS-liquids at various [NaCl] via microrheology (Fig. 3.3) and found results that match estimates from bulk rheology (Fig. A.7). Microrheology was performed by embedding fluorescent probe particles (200 nm diameter) into large NS droplets (diameter >50 μm) and tracking their mean squared displacement (MSD) over time at 20 °C (see Appendix A.3 for more technical details). NS-liquids exhibit viscous behavior at long time scales, with a diffusive exponent α, from $\langle MSD \rangle \sim \tau^{\alpha}$

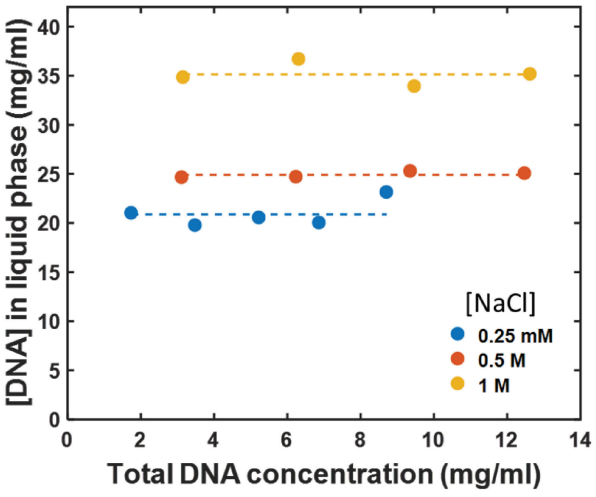


Figure 3.2: **DNA concentration.** The concentration of DNA in the NS liquid phase is independent of the total DNA concentration of the NS solutions at 0.25, 0.5, and 1 M NaCl at 20 °C.

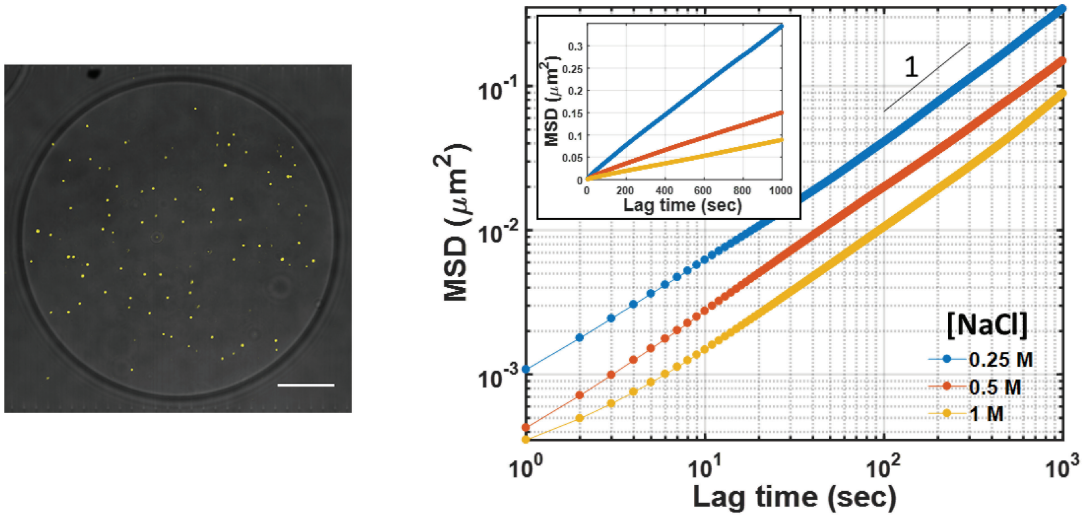


Figure 3.3: **NS-liquid viscosity**. Left: Confocal image of fluorescent probe particles (yellow) embedded in a large NS droplet at 20 °C (scale bar, 20 μm). Right: Log-log plots of averaged mean squared displacement (MSD) of the probe particles at three different concentrations of NaCl. (inset) Linear-linear plot.

with lag time τ, that is near-unity at all measured [NaCl] (Fig. 3.3). We note that, at short time scales, α deviates slightly from unity. This might be due to the onset of an elastic response, as observed in bulk rheology measurements (Fig. A.7), or it might be a consequence of displacements at short time scales being of similar magnitude to particle-tracking noise [114, 115]. Regardless, our focus here is on the well-resolved long-time dynamics. Diffusivity of the probe particles in the NS-liquid, D_{probe}, was obtained by fitting the long-time data ($\tau > 100$ sec) to $\langle MSD \rangle = 4D_{probe}\tau^{\alpha}$, where the factor of four is used because the particles are tracked in two dimensions. Viscosity η of the NS-liquid, computed from the Stokes-Einstein relation, increases with ionic strength, changing more than 3 fold, from 24 to 90 Pa·sec, as [NaCl] increases from 0.25 to 1 M (Table 3.1).

The internal dynamics of NS-liquids (i.e. the mobility of individual NSs) can be probed using FRAP. To achieve this, one of the component NS oligos was modified to include an internal Cy3 dye. Mixtures of 99:1 untagged:Cy3-tagged NSs were used to form fluorescent NS-liquid droplets. Using confocal microscopy, a point in the center of

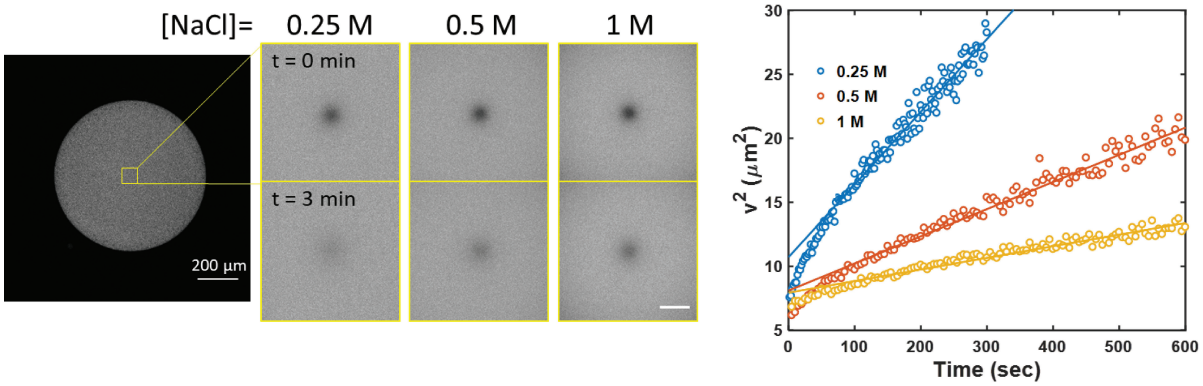


Figure 3.4: **NS self-diffusivity**. Left: Representative images of FRAP within NS droplets taken right after (i.e. postbleach) and 3 min after bleaching at [NaCl]=0.25, 0.5, and 1 M (scale bar, 10 μm). Experiments were carried out at 20 °C. Right: v^2 vs. time, where v^2 is the variance of the Gaussian function of the radial intensity profile, fits to a straight line at long time scales with slope $2D$ at each [NaCl], where D is a diffusion coefficient.

each droplet was bleached and the fluorescence recovery was subsequently tracked. Time scales of recovery were extracted from the time-dependent variance v^2 of the Gaussian function describing the intensity profile of a bleached spot (Fig. A.5) (see Appendix A.2 for more technical details). For each [NaCl], we observe a fast response within the initial seconds of recovery followed by a slow relaxation (Fig. 3.4). The origin of the fast recovery is unclear; it might be due to diffusion of free (unbound) DNA NS, or a consequence of photodamage that occurs during the bleaching step. We focus on the slow recovery (t >60 sec) where v^2 grows linearly with time, consistent with expectations for diffusion [116]. Since the slope of that linear dependence is $2D$, we can use it to calculate diffusion coefficients D; we find that D significantly decreases as [NaCl] increases (Table 3.1).

The properties of any liquid/gas interface are defined by the surface tension, which here is again expected to be salt-dependent for NS-liquid droplets. To quantify this, we exploit the frequent observation of coalescence events, in which two or more NS droplets collide and merge (Fig. 3.5). Experiments were done with droplets collected

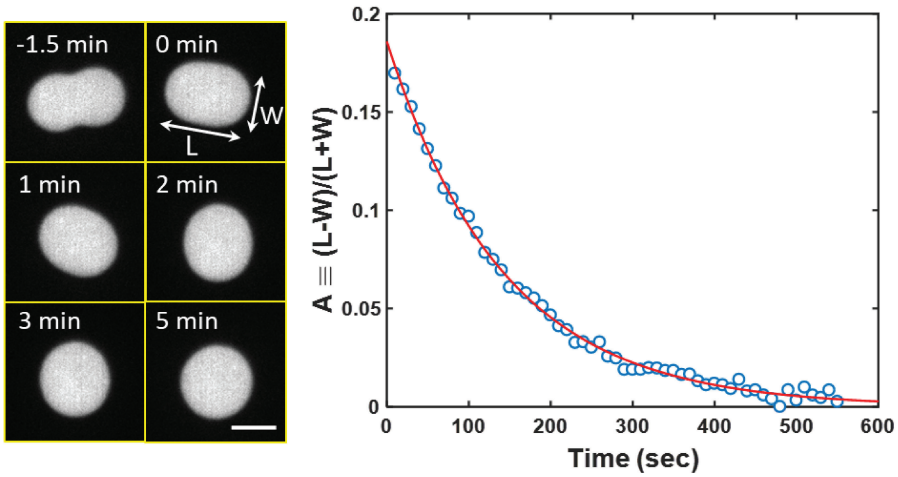


Figure 3.5: **NS-liquid surface tension**. Left: Coalescence event of two NS-liquid droplets imaged over time at 20 °C (scale bar, 5 μm). Right: Aspect ratio of a NS-liquid is defined as $A \equiv (L-W)/(L+W)$ where L and W are the length and width of the deformed droplet under relaxation.

on a surfactant-coated oil-water interface so as to minimize the effects on coalescence dynamics from adhesion to, or excess drag from, external surfaces (see Appendix A.3 for more technical details). At late stages, coalescing droplets form ellipsoids that relax to spheres over a characteristic time τ that varies with final droplet radius R and surface tension σ [117]. In the present case, where the internal (i.e. NS-liquid) viscosity η_{int} is much larger than the external (i.e. NS-gas) viscosity, the relation is

$$\tau \cong \frac{19}{20}\frac{\eta_{int} R}{\sigma} \tag{3.1}$$

Using τ from exponential fitting to the trajectory of the droplet aspect ratio and η_{int} from microrheology, we find that as [NaCl] increases, surface tension also increases (Table 3.1).

Table 3.1: Physical properties of NS-liquids: DNA concentration (c_{DNA}), viscosity (η), self-diffusion coefficient (D), and surface tension (σ).

[NaCl] (M)	c_{DNA} (mg/mL)	η (Pa·sec)	D (10^{-3} μm^2/sec)	σ (μN/m)	$D\eta$ (Pa·μm^2)
0.25	19.2 ± 1.8	24 ± 4	27 ± 5	1.23 ± 0.06	0.65
0.5	24.9 ± 0.3	45 ± 3	9 ± 3	2.28 ± 0.13	0.41
1	35.1 ± 1.1	88 ± 17	3.6 ± 0.9	3.7 ± 0.4	0.32

3.3 Role of DNA basepairing thermodynamics

Since the NS-liquid condenses through DNA hybridization of the overhangs, we expect a major mechanism of salt modulation of liquid properties will be through its effects on DNA hybridization. Indeed, this is qualitatively observed in most of the properties measured (see Table 3.1): higher [NaCl] leads to more stable bonds, which increases the dissipation upon flow (increasing the viscosity) and keeps each NS bound longer to its neighbors (decreasing the diffusivity). Further, stronger binding increases the preference for NS-NS interactions over interactions with the solvent, thereby increasing the surface tension. An estimate for such effects can be made by calculating the salt-dependent hybridization probability of the overhangs using existing DNA thermodynamics databases, and by ignoring connectivity of overhangs within a NS (see Appendix A.3.4 for more technical details); this gives upper bounds of the hybridization probability of 98.7, 99.2, and 99.5%, per overhang, for 0.25, 0.5, and 1 M salt, respectively. This rough estimate does clarify the relatively strong nature of the bonding in the system; however, as it shows very high and practically unchanging values with [NaCl], it does not help explain the relatively strong dependence with salt of our results.

Quantitative analysis of the free energy of overhang hybridization, rather than the binding probability, better illuminates the role of salt-dependent hybridization in determining D, η, and σ. SantaLucia has shown that the free energy of DNA duplex formation can be estimated, in units of kcal/mol, as $\Delta G = -(3.68 \times 10^{-4})NT \ln[\text{salt}] + C$, where

N is half the number of phosphates in the duplex, T is temperature in K, [salt] is the monovalent salt concentration in molar, and the constant $C \approx -9.5$ kcal/mol is the hybridization free energy at 1 M salt and $T = 293$ K for the particular sequence design used here [118]. We expect the surface tension to be sensitive to unbound NS arms at the liquid surface; thus, we expect $\sigma \propto |\Delta G|/{v_{mol}}^{2/3}$, where v_{mol} is the volume occupied by one molecule of NS in the NS-liquid phase. The denominator tracks the area per molecule on the surface and assumes the available space per molecule is the same on the surface as in the bulk. We thus expect $\sigma {v_{mol}}^{2/3} \propto \ln$ [NaCl], which is confirmed by the data (Fig. 3.6c). v_{mol} was calculated via $v_{mol} = MW_{NS} = (c_{DNA} \cdot N_A)$, where MW_{NS} is the molecular weight of a NS and N_A is Avogadro's number. The values of v_{mol} used here are 5190, 4000, and 2840 nm^3 for [NaCl] = 0.25, 0.5, and 1 M, respectively.

Interestingly, the absolute values of $\sigma {v_{mol}}^{2/3}$ are 100-fold below the energy of a single bond, e.g. at 1 M salt, $\sigma {v_{mol}}^{2/3} \approx 4 \times 10^{-22}$ J, while $|\Delta G| \approx 9$ kcal/mol $\approx 6 \times 10^{-20}$ J. We posit this can be explained by a structure in which the majority of the NSs on the surface retain the same number of bonds as NSs in the bulk, with internal flexibility permitting these surface NSs to orient all arms inwards. The surface tension could then be attributed to bonds lost (relative to bulk) by a minority of surface NSs. A more

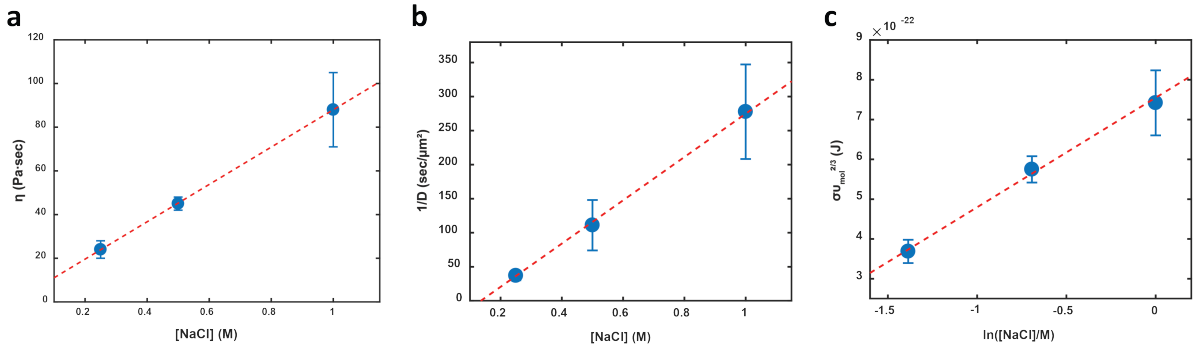


Figure 3.6: **Salt-dependence**. The relationships of (a) viscosity, (b) diffusivity, and (c) surface tension with [NaCl] indicate each parameter is sensitive to the energetics of DNA hybridization.

complex picture is needed to account for the effect of reorientation of NSs at the surface, which will create a surface-excess entropy that would also contribute to σ. Regardless, while the scaling with salt demonstrates that σ is controlled by overhang bonding, the 100-fold factor between $\sigma {v_{mol}}^{2/3}$ and $|\Delta G|$ indicates that there are relatively few unbound arms at the surface.

We relate the transport properties, η and D, to ΔG by positing that transport within the liquid is an activated process, in which a NS breaks bonds with its neighbors, moves, then rebonds in a new position. The activation energy of this process, E_a, will thus be given by the free energy, ΔG, of the bonds, and the number, m, of bonds broken: $E_a = m|\Delta G|$. The time scale, τ, of transport should then follow Arrhenius behavior, $\tau \sim \exp(E_a/RT)$, where R is the gas constant. Using the experimental conditions ($N = 5$, $T = 293$ K), this gives a power-law dependence, $\tau \sim [\text{salt}]^{0.93m}$. We expect this time scale controls both viscosity and diffusivity: $\eta \propto \tau$, $1/D \propto \tau$. The experimental data both show near-linear dependencies: $\eta \propto [\text{salt}], 1/D \propto [\text{salt}]$ (Fig. 3.6a & b). This generally indicates that hybridization thermodynamics account for the transport properties, and specifically indicates that $m \approx 1$, i.e. that the basic process limiting transport is the breaking of a single NS-NS bond. Although η and D are expected to have a prefactor that depends on density [119], that effect is minor compared to the salt dependence. Finally, we note these results are fully-consistent with the results of Bomboi *et al.* [120], who observed ln [salt] dependency of the free energy barrier associated with NS dynamics.

Overall, the results in Fig. 3.6 imply that all three NS-liquid physical properties explored in this study are dominated by NS-NS overhang hybridization. This is confirmed using NSs with a stronger overhang sequence (i.e. larger $|\Delta G|$), which indeed exhibits a higher viscosity (Fig. 3.7), confirming the potential of sequence-dependent control of NS-liquid properties.

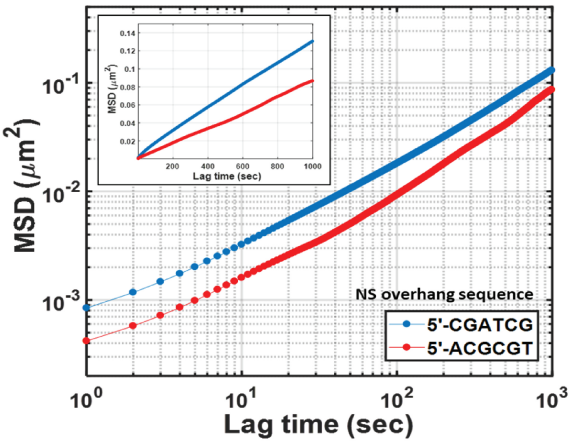


Figure 3.7: **Sequence-dependent viscosity.** Averaged MSDs of 200 nm probe particles in liquid droplets of two different NSs, NS-CGATCG and NS-ACGCGT, as a function of lag time at 0.5 M NaCl. The lower MSD values for NS-ACGCGT lead to a higher viscosity, when calculated from $MSD = 4D_{probe}\tau^{\alpha}$, reflecting the stronger hybridization strength of the sequence [118].

3.3.1 Beyond basepairing: Density, Stokes-Einstein behavior, and clustering

Interestingly, certain aspects of our data cannot be attributed to effects of salt-dependent hybridization. This includes the effect of salt on NS-liquid density, c_{DNA}. Qualitatively, salt-induced bond stabilization should lead to more neighbors per NS, and an increase of density with salt, as observed. Quantitatively, one might expect this trend to saturate when the number of neighbors approaches four (i.e. the NS valence) at a value near the density of NSs in a diamond lattice, here corresponding to $c_{DNA} \approx 13.3$ mg/mL [100]. However, we measure droplet densities that far surpass this estimate. We speculate that high salt concentrations alter NS-liquid structure and enhance NS-liquid density not just by saturating NS bonds, but by screening electrostatic repulsion between NSs, or even within a NS. The latter would permit closer approach of NS arms, or

base-stacking across the NS junction; such effects are seen in Holliday junctions, 4-armed DNA structures of biological significance [121, 122].

Simple hybridization arguments also fail to explain the relation of the measured transport properties, D and η. In the NS-liquid phase, both diffusive transport and shear flow are hindered by NS-NS binding; thus, it is reasonable to expect that D and η will vary in a correlated fashion as binding interactions are perturbed by changes in salt. A standard prediction for this correlation is the Stokes-Einstein relation, which predicts that the product $D\eta$ is a constant, so long as the temperature and the particle size do not vary [123, 124]. Here, however, we find that $D\eta$ is sensitive to salt (Table 3.1), showing a strong decrease as [NaCl] increases. This might be attributed to a salt-dependent change in particle size, but the trend is opposite of what is expected. Our density measurements indicate the available volume per particle, v_{mol}, decreases with salt; further, intra-NS screening effects would lead to a smaller size of NSs at higher salts. If such salt-dependent compaction were the only effect, $D\eta$ would increase with salt, in contrast to what is observed. The strong decrease in $D\eta$ is likely due to the increased importance of collective effects at high salt concentrations: as bond strength increases at higher salt, NSs form larger and more stable effective clusters within the liquid, which diffuse progressively more slowly. Such collective dynamic behaviors are a characteristic feature of liquids with network structures similar to the present NS system [125, 126], lending support to this interpretation.

The existence of stably-bonded clusters of NSs is somewhat at odds with the observation that transport is controlled by a single-bond activation barrier (i.e. $m \approx 1$): The former implies a relatively integrated structure, with each NS participating in multiple bonds, yet the latter suggests that breaking only a single bond enables NS mobility. We speculate that this can be reconciled by the existence of a heterogenous structure in the liquid, in which NSs are indeed well-bonded to neighbors within the same cluster,

but with relatively few bonds connecting neighboring clusters, such that breaking of a single cluster-cluster bond induces mobility. This would indicate that most NSs at a cluster-cluster interface orient their bonds inwards (into their cluster), rather than across the interface; this is consistent with our interpretation of the magnitude of the surface tension, as discussed above.

3.4 Comparison to other coacervate systems

The unique features of NS-liquids become clear when compared to those of other reported biomolecular liquids, such as synthetic coacervates and liquids derived from components of intracellular droplets (see Table 3.2). The most salient difference is that NS-liquid properties respond to salt in the opposite manner as the other systems. This occurs because the NS-liquids form due to salt-stabilized DNA hybridization, while the other systems form due to electrostatic attractions, which are weakened by added salt.

Table 3.2 also points out that the various biomolecular liquids vary strongly regarding their density; note that the variation in the remaining properties (i.e. viscosity and surface tension) can roughly be understood as being sensitive to density. Synthetic coac-

Table 3.2: Physical properties of various coacervate and biomolecular liquid systems.

System	**Fraction macromolecule**	**Viscosity (Pa·sec)**	**Surface tension (N/m)**
DNA NS droplets	2–3.5 wt% $\phi \approx 0.01$–0.02 (increase)	10^{1} (increase)	10^{-6} (increase)
Synthetic coacervates (Polyelectrolytes/polypeptides/ polysaccharides) [126, 127, 128, 129, 130, 131, 132]	15–35 wt% $\phi \approx 0.1$–0.4 (decrease)	10^{1}–10^{5} (decrease)	10^{-5}–10^{-4} (decrease)
Intracellular droplets [115, 133, 101, 134, 135]	0.3–0.8 wt% $\phi \approx 0.002$–0.006 (decrease)	10^{0}–10^{1} (decrease)	10^{-6}–10^{-4} (decrease)

* (*response to increasing [salt]*)

ervate systems are generally at least 10-fold more dense than the other systems, likely because the constituent polymers tend to be highly-charged and flexible, permitting each chain to contact many oppositely-charged chains, thus driving higher liquid densities [127]. In contrast, the present NS-liquid system has a limited valence: strong interparticle contact is permitted only at the four overhangs, which decreases particle packing and liquid density. Further, we expect that the stiff nature of the DNA arms contributes to the low density by pushing bound particles away from each other. Interestingly, single-component liquids formed from purified LAF1, a protein derived from an intracellular droplet system, achieve an extraordinarily low density ($\phi \approx 0.3\%$). The researchers attributed the expansive occupied volume in their system to large configurational fluctuations of the protein [135]. However, it is interesting to note that LAF1 also consists of a rigid, non-interacting core decorated with an external domain capable of limited-valence interactions.

It has been suggested that the low density of intracellular droplets may have evolved to accommodate the free diffusion of small molecules within the meshwork of the macromolecular liquid, thereby allowing biochemical reactions to occur [135]. We suggest that NS-liquids could similarly permit chemical activation through the action of permeable species, perhaps permitting the DNA system to act as an *in vitro* model for testing the biological implications of sparse liquid structure. In this context, the ability to tune NS structure through sequence is a potent feature that enables control of liquid physical characteristics in varied solution conditions.

Chapter 4

Partitioning into DNA liquids

The low density of NS-liquids (Table 3.2) implies a capacity to accommodate other components. In this chapter, we explore the determinants of NS-liquid partitioning and examine the potential functionalization of the liquid droplets with permeated solutes as a model for membraneless organelles.

4.1 Membraneless organelles

A core principle in biology is the concept of a cellular compartment, or organelle, that provides a *specific* environment for *specific* biomolecules to interact and for *specific* biochemistry to occur. Organelle environments are delimited from the cellular cytoplasm by lipid membranes, relying on sensitive transport mechanisms to control biomolecular entry & exit. In recent decades, researchers have reported many instances of liquid-liquid phase separation (LLPS) within cells, where the polymer-rich liquid phase exhibits incredibly low volume fractions [82], as described in the previous chapter (section 3.4). As the interacting biomolecules (e.g. proteins, RNA, and/or DNA) that constitute the phase-separated liquid occupy very little space, it has been proposed that such sparsity

enables accommodation of other biomolecular components [135]. The complete hypothesis is that a cell could employ LLPS to generate a membraneless liquid compartment, or "membraneless organelle", to perform specific biochemistry on demand. Thus, by controlling the formation and dissolution of these biomolecular liquids, the cell not only has spatial control of biomolecular interactions but also gains temporal control as well. Researchers are actively probing the extent to which this hypothesis is true, developing tools to precisely induce (or reverse) phase separation and examining the cellular/physiological consequences [136].

Within chromatin environments, specifically the eukaryotic nucleus, there are many instances of LLPS. The nucleolus, paraspeckles, nuclear speckles, Cajal bodies, and PML nuclear bodies are all examples of nuclear liquids, and each has been attributed with a particular cellular role [137]. Researchers are working to better understand how the physical properties of liquid environments might contribute to and/or affect these biological functions. The nucleolus, for instance, is well-known to be the site of ribosome biogenesis. Closer study has revealed that the nucleolus exhibits a nested, multi-layered, droplets-within-droplets organization: The innermost droplet (called the fibrillar center, FC) is the site of RNAp I activity and rRNA transcription; the intermediate droplet layer (known as the dense fibrillar component, DFC) engulfs the FC droplet and contains proteins associated with rRNA processing; and the outermost droplet layer (or the granular component, GC) is the site of ribosome assembly [138]. Recent biophysical characterization has demonstrated that the multi-layered nesting is driven by differences in surface tension between the different nucleolar liquids, where the enveloped liquid has a higher surface tension than the enveloping liquid (FC > DFC > GC) [134]. We thus see that the cell has utilized LLPS, exploiting differences in surface tension that are attributable to differing biomolecular compositions, to establish an assembly line for ribosome biogenesis, assigning different tasks to specific nucleolar liquid compartments.

Recent work has highlighted liquid domains of heterochromatic DNA [139, 140], specifically identifying *constitutive* heterochromatin (i.e. silenced regions of eukaryotic DNA that are common to most, if not all, cell types) as liquid domains arising from LLPS. As described in chapter 1 (section 1.2), the gene-silencing property of heterochromatin is commonly attributed to its condensed organization, which sterically limits access to DNA. However, enzymatic processes have been reported to act on heterochromatic DNA: DNA repair proteins/complexes can correct damaged heterochromatic DNA [141, 142]; and RNA polymerases can indeed transcribe heterochromatic DNA, though at a low level, to recruit structural proteins as well as to produce RNA (e.g. repeat-associated siRNA, rasiRNA) to reinforce gene silencing via RNA interference [143, 144, 145]. Similar to the nucleolar case, heterochromatin could be utilizing the liquid environment as a membraneless organelle, selectively allowing certain components to enter (and act) when necessary.

NS-liquids, being comprised of DNA and exhibiting its sequence-specific tunability, can act as a model system to probe the biological implications and mechanistic features of phase-separated membraneless organelles that contain DNA (e.g. liquid heterochromatin domains). In this chapter, we examine the factors that govern the partitioning of biomolecules into NS-liquid droplets, describing features related to charge, DNA-binding affinity, and sequence-specificity that are perhaps expected of a negatively-charged DNA liquid. Also, we demonstrate the engineering capacity of NS-liquids by using DNA-modifications to induce biomolecular partitioning and, in so doing, identify a length-dependence for integrating dsDNA. Altogether, we demonstrate that NS-liquids can concurrently accommodate many types of biomolecules (including proteins and nucleic acids), supporting their application as an *in vitro* model of membraneless organelles.

4.2 General factors governing partitioning

We utilized fluorescent microscopy to examine a variety of factors that could affect biomolecular partitioning into NS-liquids. Using the same NS system described in the previous chapter (i.e. fully-flexible NSs with 4 arms and 5' overhangs) (Fig. 3.1), NS-liquid droplets were pre-formed by mixing NSs with salt in 10 mM Tris buffer (pH 7.5) and incubating on a rotator at 20 °C for >1 hr. Fluorescent biomolecules were subsequently added without pipette mixing and transferred to a flow cell using an end-cut pipette tip. After >3 hr of further incubation at 20 °C, samples were examined using confocal microscopy (see Appendix A.2 for more technical details).

4.2.1 Size

To probe the size-dependence of partitioning, we examined the success with which varying molecular weights of FITC-labeled dextran, frequently used to characterize liquid- and mesh-like systems [146, 135], partition into NS-liquid droplets in 1 M NaCl. We can qualitatively see that dextran of higher MW is less able to enter NS-liquids, as evidenced by dark spherical voids corresponding to the droplets (Fig. 4.1a). This is confirmed through quantitative examination of the fluorescence intensity difference between the inside & outside of NS-liquids. Images were normalized to the maximum intensity of FITC-dextran in bulk solution and the intensity profile (radially averaged) was calculated from the center to the outside of each droplet (Fig. 4.1b). The measured reduction in fluorescent intensity within a NS-droplet is extremely consistent among samples at each MW. The intensity of FITC-dextran fluorescence within NS-liquids (as a fraction of bulk intensity) for 4, 10, 20, 40, and 70 kDa FITC-dextran are 0.838 ± 0.002, 0.729 ± 0.003, 0.6388 ± 0.0008, 0.505 ± 0.002, and 0.282 ± 0.005, respectively.

To approximate a mesh size for the NS system, we can make the assumption that

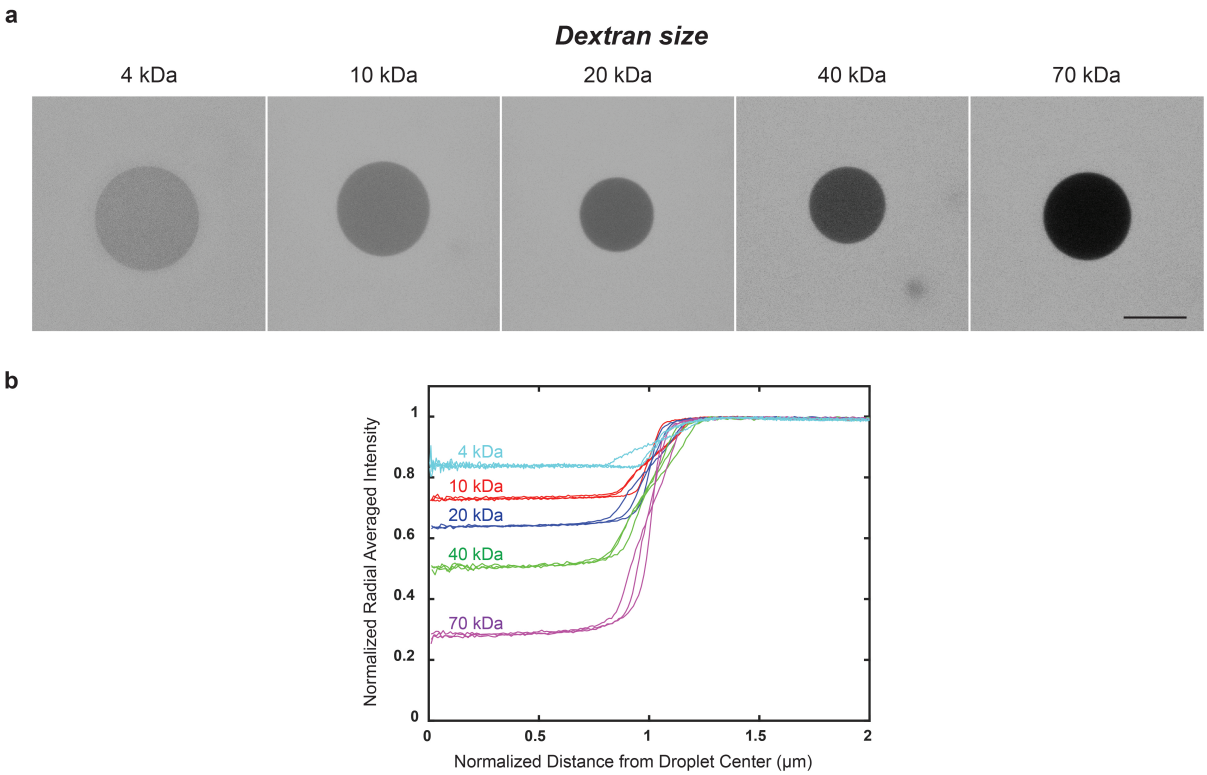


Figure 4.1: **Biomolecular size. (a)** Confocal images (488 nm excitation) of FITC-labelled dextran of varying molecular weights with unlabeled NS-liquid droplets. Scale bar = 10 μm. **(b)** Intensity profiles crossing from the center of a NS-liquid droplet into bulk solution for triplicate samples at each dextran MW.

each FITC-dextran sample (purchased from/provided by Sigma-Aldrich) exhibits a size distribution that centers around the designated MW. When the designated MW exhibits a hydrodynamic diameter that is equal to the mesh size, and assuming that larger molecules are excluded from the mesh while smaller ones can freely enter, then the fluorescence intensity within the mesh is expected to be 1/3 of that outside the mesh. In our image analysis, this outcome is observed with the 70 kDa FITC-dextran sample. Using the hydrodynamic diameter of 70 kDa FITC-dextran [147], we can approximate the mesh size of NS-liquids to be ∼12 nm. It is interesting to note that the distance between NS junctions with bound arms (assuming no arm bending; i.e. a straight line) is ∼15 nm.

4.2.2 Attractive interactions: charge & sequence-specificity

We used large, charged fluorescent particles to examine the effect of electrostatic interactions on partitioning. With DNA being an anionic polyelectrolyte, we confirm that, with 250 mM NaCl at 20 °C, negatively-charged components, specifically carboxylate-modified fluorescent particles, are excluded from NS-liquids, while positively-charged components, specifically amine-modified fluorescent particles, partition within NS-liquids (Fig. 4.2a). These particles are 2 μm in diameter. We thus see that, regardless of size, particles (and, by extension, biomolecules) will readily partition into NS-liquids if they exhibit a sufficiently strong attractive interaction with DNA NSs.

As the system is comprised of DNA, we examined how attractive interactions due to DNA hybridization affect biomolecular partitioning. Indeed, we confirm that sequence-complementarity to NS overhangs favors integration into NS-liquids. Streptavidin can be modified with DNA through its strong non-covalent interaction with biotin. When functionalized with biotinylated-ssDNA presenting the same palindromic region exhib-

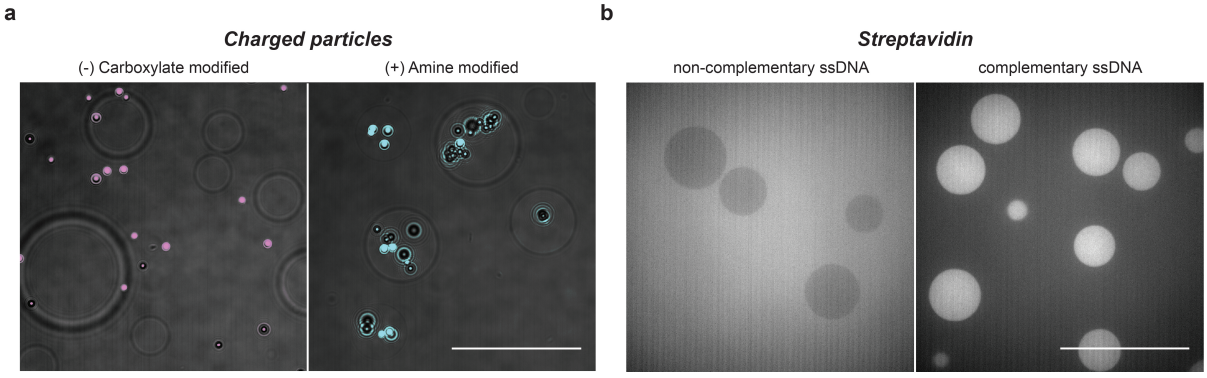


Figure 4.2: **Charge & DNA sequence specificity. (a)** Overlay of brightfield and confocal images (561 nm excitation) of fluorescent particles (2 μm diameter) with unlabeled NS-liquid droplets. Particles are false colored magenta and cyan, indicating negatively- and positively-charged particles, respectively. **(b)** Partitioning of streptavidin, labeled with FITC and imaged at 488 nm excitation, when non-covalently modified with biotinylated-ssDNA exhibiting sequences that are non-complementary (left) or complementary (right) to NS overhangs. All scale bars = 50 μm.

ited by NS overhangs, streptavidin (tagged with FITC for fluorescent visualization by BioLegend®) readily partitions into NS-liquids at 250 mM NaCl in 20 °C (Fig. 4.2b right). Without ssDNA modifications or with biotinylated ssDNA exhibiting a non-complementary sequence, streptavidin remains excluded from NS-liquid droplets (Fig. 4.2b left). DsDNA can be similarly modified with ssDNA overhangs, either through digestion with restriction enzymes or, as we do, through PCR (polymerase chain reaction) amplification using primers with internal abasic sites [148]. In 'auto-sticky PCR', as the technique is called (see A.1.2 for more details), each ssDNA primer contains an internal abasic nucleotide that stalls DNA polymerase activity. As a result, any DNA that is 5' of the abasic site remains single-stranded, leaving each PCR product with free 5' overhangs. As with streptavidin, only when dsDNA exhibits complementary overhangs to the NSs do they partition into NS-liquids at 250 mM NaCl (Fig. 4.3a).

Length-dependent dsDNA integration

For a specific overhang sequence in a particular environment (i.e. at a defined hybridization strength), we observe a length-dependence for the integration of dsDNA into NS-liquids. Using the 5'-*CGATCGA*-3' overhang sequence in a 250 mM NaCl environment, we examined the partitioning behavior of dsDNA of various lengths, finding that the shorter DNA fragments (131 and 202 bp) partition homogeneously into NS-liquids, while the longer fragments (278, 350, and 426 bp) remain excluded (Fig. 4.3b top). When [NaCl] is raised to 1 M, 278 bp fragments gain the ability to partition into NS-liquids (Fig. 4.3b bottom), albeit not as favorably as the shorter fragments. This length-dependence can be understood as balancing the binding energy of overhang hybridization (which drives partitioning) with the entropic cost of confining DNA into a NS mesh and the electrostatic cost arising from repulsion between the dsDNA fragment and the NSs (which, together, resist partitioning). While the binding energy is constant for

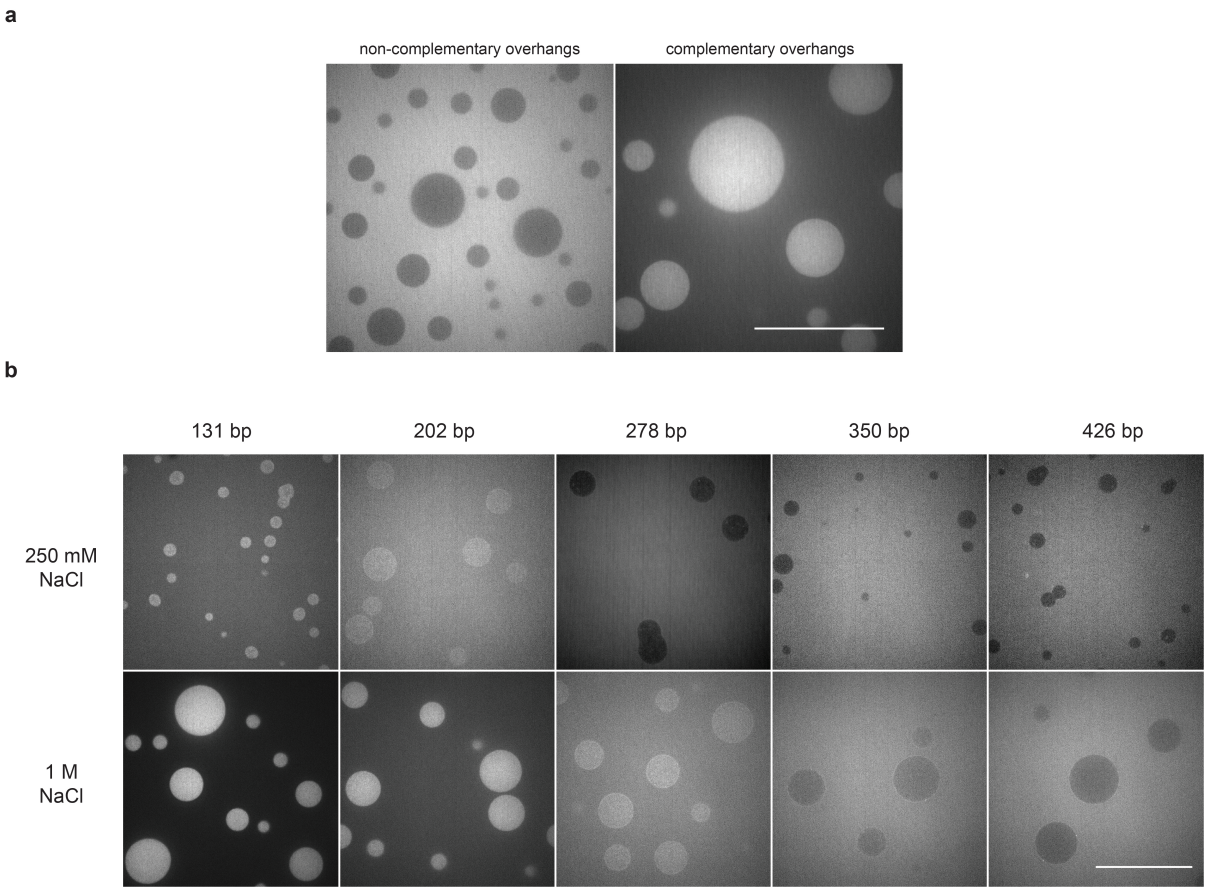


Figure 4.3: **Sequence-specific dsDNA partitioning.** Confocal images of fluorescently-labeled dsDNA with NS-liquid droplets after >3 hr incubation at 20 °C. **(a)** Partitioning of dsDNA, labeled internally with Cy3 and imaged at 561 nm excitation, when exhibiting ssDNA overhangs that are non-complementary (left) or complementary (right) to NS overhangs. **(b)** Partitioning of dsDNA of increasing lengths, exhibiting complementary overhangs to NSs, at 250 and 1000 mM NaCl. All dsDNA is labeled internally with Cy3 and imaged at 561 nm excitation. All scale bars = 50 μm.

a particular overhang, the costs of partitioning increase with increasing polymer length. The 278 bp dsDNA appears to sit near the boundary, where the energies driving and resisting partitioning are nearly balanced. Increasing [NaCl] to 1 M, and thereby increasing the binding energy, was sufficient to overcome the energetic costs, thereby enabling partitioning. However, as illustrated in Fig. 4.3, that same increase was insufficient for the longer 350 bp and 426 bp fragments.

4.2.3 DNA binding affinity

NS-liquid partitioning also shows a dependence on intrinsic DNA binding affinity. To examine proteins, salt conditions were adjusted to 50 mM NaCl and 8 mM $MgCl_2$, which are conditions that better reflect cellular environments where proteins are functional. We examined the partitioning behaviors of streptavidin, GFP, and T7 RNAp; for fluorescent visualization, streptavidin was conjugated to FITC by the manufacturer (BioLegend®) and T7 RNAp was fused with GFP at its C-terminus via a short linker. Both GFP and GFP-fused T7 RNAp were purified via Ni^{2+}-column chromatography with N-terminal His tags (see Appendix A.1.3 for more technical details).

Both streptavidin and GFP, proteins with minimal (if any) affinity to DNA, are ex-

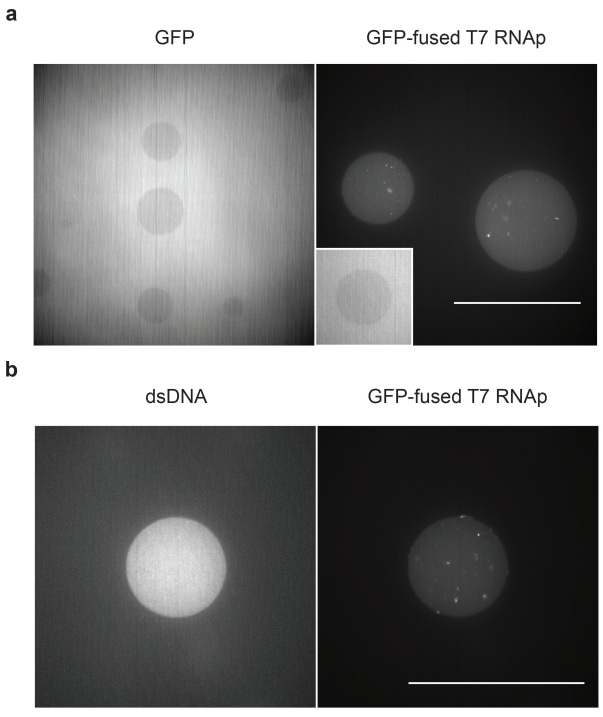


Figure 4.4: **Concurrent partitioning in physiological conditions. (a)** Partitioning of GFP and GFP-fused T7 RNAp with NS-liquid droplets in 50 mM NaCl & 8 $MgCl_2$. Inset: GFP-fused T7 RNAp partitioning at 250 mM NaCl. **(b)** Concurrent partitioning of dsDNA and GFP-fused T7 RNAp into the same NS-liquid droplet. Scale bars = 50 μm.

cluded from NS-liquids at 20 °C (Fig. 4.2b left and 4.4a left, respectively). However, GFP-fused T7 RNAp, which exhibits intrinsic DNA-binding affinity, readily partitions into NS-liquids (Fig. 4.4a right). Note that T7 RNAp has a molecular weight of ∼100 kDa, which is well above the partitioning size threshold suggested by the dextran results (Fig. 4.1), even without the added mass from the GFP fusion. At elevated salt (e.g. 250 mM NaCl), GFP-fused T7 RNAp no longer partitions (Fig. 4.4a right inset), recapitulating known DNA binding behaviors of T7 RNAp [149]. It thus appears that biomolecular partitioning will occur as long as the biomolecule has a sufficiently strong attractive interaction with the DNA NSs.

While there are some limitations (e.g. length-dependence of dsDNA integration), these results demonstrate that NS-liquids can act as membraneless organelles: they partition biomolecules in a specific manner that relies on affinity to the liquid's components. Notably, multiple components can partition into the same NS-liquid concurrently. For example, we see that dsDNA can co-localize with GFP-fused T7 RNAp in a single NS-liquid droplet (Fig. 4.4b). The work presented in this chapter also illustrates the engineering capacity of DNA NS-liquids: Biomolecules can be modified either covalently (e.g. GFP fusion with T7 RNAp) or non-covalently (e.g. streptavidin modified with biotinylated-ssDNA) to favor (or force) partitioning. For a model system, this tractability is advantageous, enabling greater control and versatility in devising the appropriate *in vitro* conditions to test & probe behaviors in membraneless organelles.

Chapter 5

Transcription from DNA liquids

After establishing that NS-liquids can act as viable *in vitro* models of membraneless organelles, through their ability to selectively partition biomolecules, we next considered how the NS-liquid environment might affect enzyme activity. Specifically, we sought to determine how transcriptional activity is affected when the gene of interest is integrated into a NS-liquid.

5.1 Cell-free gene expression

As briefly reviewed in chapter 1 (section 1.3), *in vitro* cell-free gene expression (CFGE) systems were instrumental in elucidating the genetic code and key biological processes. These systems can vary with regard to their degree of sophistication, where some perform transcription alone while others enable coupled transcription-translation. However, they all share one defining feature: the gene of interest is expressed outside of (and without the aid of) a living cell.

While gene expression itself occurs in the absence of cells, the transcriptional and/or translational machinery are indeed derived and purified from cells, where the type of

cellular extract (or lysate) that is used depends on the complexity of the protein of interest [150]. Bacterial lysate (generally from *E. coli*) is sufficient to produce simple proteins. For more complex proteins that require e.g. post-translational modifications for functionality, lysates derived from higher order organisms are required. These cell-free systems have garnered significant interest as a means of high-yield production of valuable biological products (usually proteins), since they overcome cell viability limitations encountered through traditional over-expression/purification methods. From malarial vaccines [151] to cytotoxic cancer treatments [152], CFGE systems are being applied to produce a diverse array of genes.

In efforts to increase efficiency and output, scientists/engineers/bio-pharmaceutical companies/etc. have been adapting CFGE systems to better mimic cellular environments. For example, research examining the influence of macromolecular crowding saw that transcription rates increased with higher concentrations of crowding agents (up to a certain point) [153]. More recent work, highlighting the increased interest in biomolecular liquids, examined gene expression in a cell lysate coacervate, showing that compartmentalization of expression machinery in a liquid-liquid phase separated droplet dramatically enhances transcription [154]. However, in these (and the vast majority of) CFGE systems, the DNA that encodes the gene of interest is allowed to freely diffuse in the reaction environment. Besides deciding whether the DNA should be in either a linear or a circular form, there is little other consideration concerning the structural state of the DNA to be transcribed. That is, there has been little effort to mimic chromatin compaction states in CFGE systems. It is perhaps not so surprising since, from a commercial perspective, the driving force has always been to maximize expression levels, and maximum DNA accessibility is presumably achieved when the DNA template is freely diffusing.

However, in academic-oriented efforts, a few CFGE systems have been developed that integrate the gene-encoding DNA template into nucleic acid complexes with higher

order structure, revealing some insight to how DNA structure can affect gene expression. For example, linearized plasmid was covalently embedded into hydrogels comprised of branched DNA nanostructures (of similar design to NSs). In wheat germ lysate, these *P-gels*, as they were named, exhibited a 94-fold increase in protein expression over an equal amount of freely diffusing templates [155]. The concentration of template into a localized region, and the presumed concentration of expression machinery, appears to account for the increased output. Genes have also been integrated into DNA brushes, where dsDNA strands are immobilized on photolithographic biochips that enable control of dsDNA spacing [156]. The researchers showed that high brush density can inhibit transcription by sterically hindering the accessibility of promoter sites. These systems, though undoubtedly simplistic compared to actual chromatin, illustrate how artificial chromosomal states can begin to probe the interplay between the structural organization of DNA and transcription from that DNA.

There have been an increasing number of reports describing the liquid-like behavior of chromatin, most of which highlight the lack of long range order and the highly dynamic mobility of chromatin domains [157]. Additionally, as described in chapter 3, there is increased interest regarding the role of liquid-liquid phase separation in biology. With these considerations, and after demonstrating the partitioning capacity of DNA liquids, we chose to implement the NS-liquid phase into a CFGE system using T7 RNAp. We find that T7 RNAp activity exhibits a dependence on NS concentration, where the enzyme is inhibited from promoter-dependent activity by interactions with non-template DNA. Notably, the inhibitory effect is diminished when non-template DNA NSs are concentrated into a liquid phase. Further, preliminary evidence suggests that transcription is relatively insensitive to gene integration into DNA liquids. Regarding the reverse direction of the relationship, we observe evidence that it is not just DNA structure that affects transcription but that transcriptional products can also reorganize DNA.

5.2 Spinach transcriptional reporters

With the aim of simplicity, we implement a "bare-bones" CFGE system that contains only the components and machinery needed for transcription. Specifically, we utilize T7 RNAp, as it is a single protein that can initiate and terminate transcription without additional protein factors [158]. For expedient transcriptional quantification, we utilize the *Spinach* family of fluorescent reporters as our gene(s) of interest, as they have been previously implemented as real-time monitors of RNA synthesis [159, 160]. Two gene templates were utilized in this work, both put under the control of the T7 promoter consensus sequence [161]: Spinach2 [162] with the T7 tϕ terminator [163, 164]; and iSpinach followed by the tZ terminator [165] (see Appendix B for DNA sequences). The latter pair, which are genetic variants improved for their respective functions (i.e. fluorescence and termination), were adopted later upon their invention. Traditional PCR was used

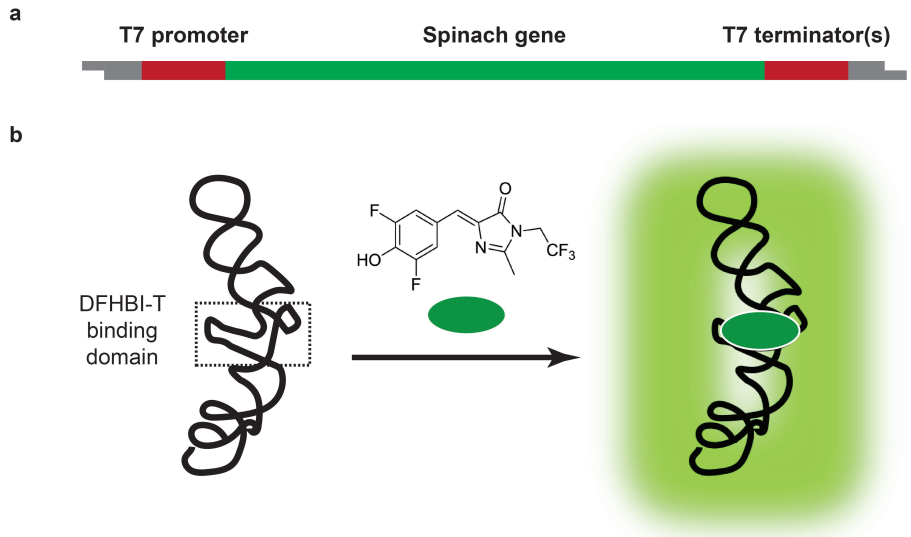


Figure 5.1: **Spinach transcriptional reporters.** **(a)** To quantify transcription, members of the *Spinach* family of RNA reporters were integrated into dsDNA templates under the control of the T7 consensus promoter and T7-specific terminator(s). Specific DNA sequences of gene templates are reported in Appendix B. **(b)** *Spinach* RNA adopts a folded structure that, upon interaction with the small molecule DFHBI-1T, leads to a dramatic increase in fluorescence from the complex.

to produce blunt templates while auto-sticky PCR was used to generate templates with NS-specific overhangs. Both Spinach2 and iSpinach RNA, upon transcription and proper folding, interact with the small molecule DFHBI-1T to fluoresce thousands of fold higher than DFHBI-1T alone (Fig. 5.1). Fluorescence was measured with a Tecan microplate reader using 480 nm excitation and 513 nm emission.

Transcription reactions were performed at 25 °C in standard transcription conditions: 40 mM Tris buffer (pH 8), 8 mM $MgCl_2$, 500 μM rATP/rUTP/rCTP, 1 mM rGTP, 1 U/μL SUPERase-In^{TM} RNAse inhibitor, and 20 μM DFHBI-1T. 50-100 mM NaCl was also added, depending on the experiment. Though excessive [NaCl] is known to hinder T7 RNAp activity [166], the added salt aids *Spinach* mRNA folding as well as sustains NS-liquid phase separation; notably, we still observe significant transcriptional output in these salt conditions. We confirm that when an essential transcriptional component (e.g. rNTPs, DNA template, or T7 RNAp) is removed from the reaction or when a non-*Spinach* gene is transcribed (e.g. GalE), the fluorescent signal does not increase with time (Fig. 5.2a). Only when all components are present and a *Spinach* family gene is expressed does the fluorescent signal increase with time. PAGE results confirm that

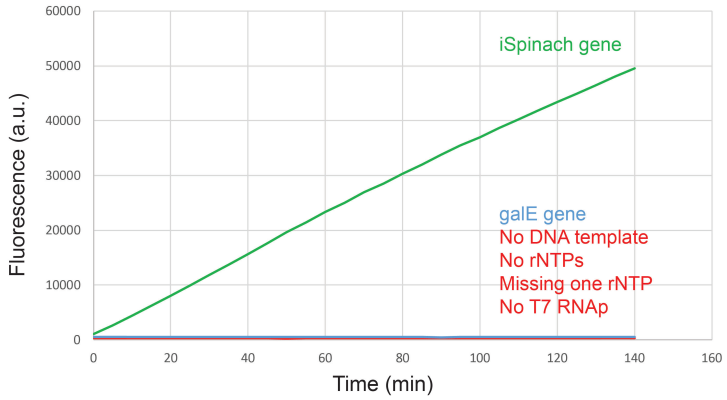


Figure 5.2: **Quantifying transcription.** Fluorescence signal increases over time only when a *Spinach* gene is transcribed and when all reaction components are present. If any essential elements are removed, then the fluorescent signal remains unchanged over time.

fluorescence intensity reflects the amount of *Spinach* RNA transcribed (Fig. A.8).

We examined how the presence of NSs perturbs T7 RNAp-dependent transcription, assessing whether transcription is affected by elevated concentrations of non-template DNA (i.e. DNA exhibiting minimal sequence homology with the T7 consensus promoter). To preclude any self-assembly and droplet formation, NSs and iSpinach dsDNA templates lacking overhangs (i.e. blunt) were used. With 100 nM T7 RNAp and 100 ng iSpinach template (∼18 nM), the transcriptional output, as indicated by the fluorescence intensity of the iSpinach RNA-DFHBI-1T complex, is significantly reduced with increasing concentrations of NSs (Fig. 5.3a). NSs exhibit an inhibitory effect, where T7 RNAp's affinity to non-template DNA, though weak compared to promoter affinity, is sufficient to distract/occupy RNAp from promoter-dependent activity. When NSs with complementary palindromic overhangs were used, forming NS-liquid droplets in this salt condition, we observe higher iSpinach expression compared to an equal concentration

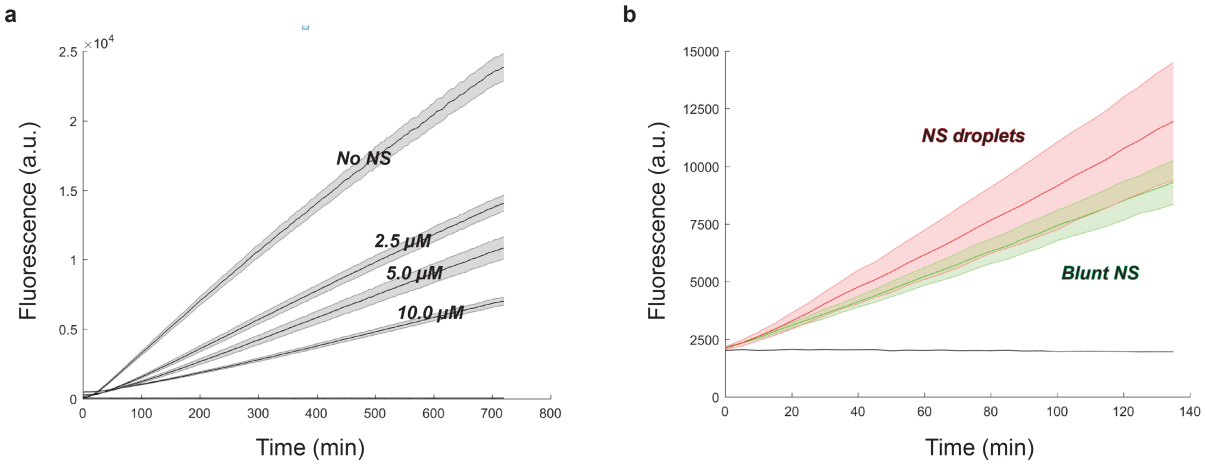


Figure 5.3: **Dependence on non-template DNA concentration. (a)** Transcriptional output from blunt iSpinach templates is sensitive to the amount of blunt NSs (i.e. non-template DNA) in solution. **(b)** When the phase of the inhibiting non-template DNA (i.e. NSs) is changed to that of a NS-liquid, inhibition of transcription appears reduced, as evidenced by an increase in fluorescence. Reactions were performed with 100 nM T7 RNAp and 100 ng iSpinach template (∼18 nM). Each curve shows a mean bounded with shaded regions denoting standard deviation.

(10 μM) of blunt NS (Fig. 5.3b). The phase of the inhibiting DNA appears to influence the degree of transcriptional suppression. Formation of NS-liquids appears akin to sequestering inhibitor, reducing the effective concentration of inhibitor in bulk solution.

We then compared transcription from *Spinach* templates integrated into NS-liquid droplets via complementary overhangs to that from freely diffusing blunt templates with droplets present, probing whether template integration into NS-liquid droplets affects gene expression. The Spinach2 template was utilized since its shorter 202 bp length, compared to the 426 bp of the iSpinach template, enabled homogeneous partitioning into NS-liquid droplets in these salt conditions. No significant difference in transcription was observed with: equimolar concentrations (50 nM) of both T7 RNAp and Spinach2 dsDNA; or T7 RNAp in excess (20 nM) over Spinach2 dsDNA (5 nM) (Fig. 5.4a and b, respectively).

A naive hypothesis might have expected the integration of dsDNA templates into NS-liquids to recapitulate the transcriptional silencing behavior attributed to condensed

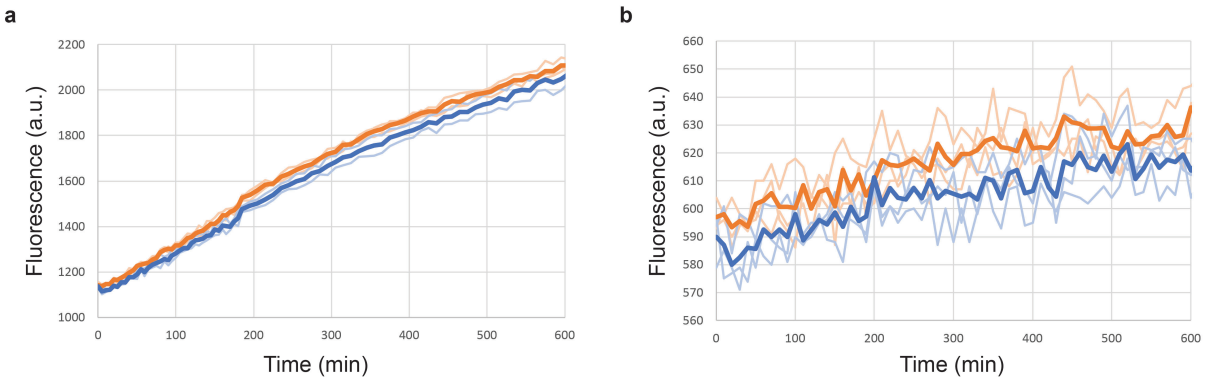


Figure 5.4: **Transcription from integrated gene templates.** Spinach2 fluorescence over time when transcribing from Sp2 templates either integrated into NS-liquid droplets (i.e. dsDNA with complementary overhangs to NSs; blue lines.) or freely diffusing (i.e. dsDNA without overhangs; orange lines) using (a) 50 nM T7 RNAp and 50 nM Spinach2 dsDNA and (b) 20 nM T7 RNAp and 5 nM Spinach2 dsDNA. Each curve represents one transcription reaction; bold curves correspond to mean values.

heterochromatin. While we do observe a slight decrease in the mean fluorescence intensity for samples with templates integrated into NS-liquids (Fig. 5.4a & b; blue lines), the variance among the triplicate samples is too large to confidently conclude that the difference is real. It is possible that different biochemical conditions (e.g. very low *Spinach* template concentration) might better illustrate or amplify a reductive effect associated with template partitioning, should one exist. However, the observed indifference to NS partitioning spawns interesting questions that can be pursued in future work, which will be discussed in the following chapter.

5.2.1 Transcription-dependent structural changes

In chapter 1 (section 1.2), I described how the relationship between chromatin organization and transcription is bi-directional. Not only can DNA structure affect gene expression but RNA products can organize/reorganize DNA structure. Indeed, we observe

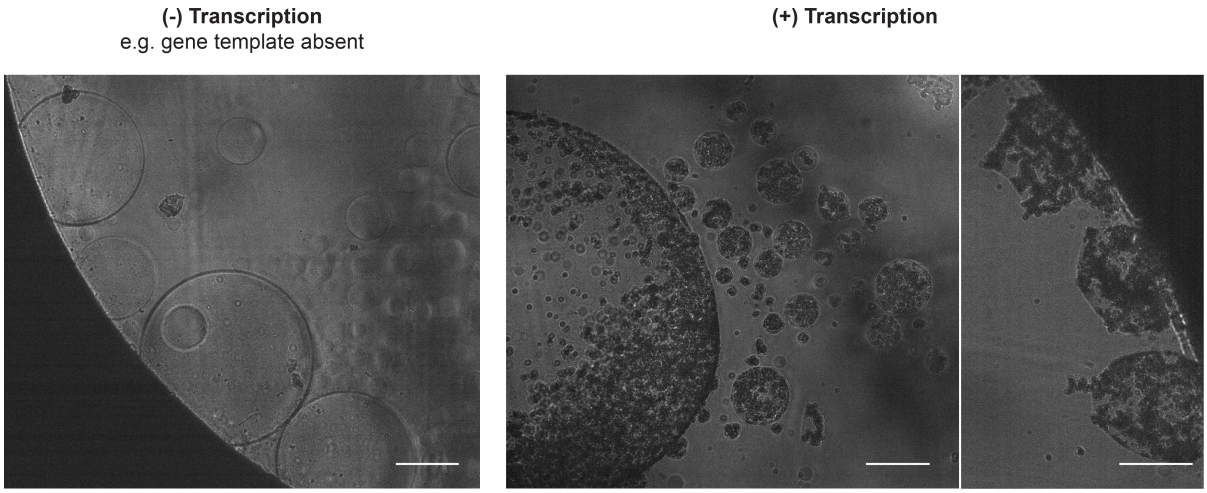


Figure 5.5: **Transcription-dependent structural changes.** *Spinach* RNA transcripts interact with DNA NS-liquid droplets, influencing structure through the formation of clusters and the deformation of spherical droplets (right). Images were taken using brightfield microscopy from individual samples in a microwell plate. Scale bars = 100 μm.

this behavior with the NS-liquid system. Many hours after the beginning of transcription, NS-liquid droplets exhibit internal speckles/clusters as well as shape deformations, where the previously spherical droplets exhibit non-spherical features (Fig. 5.5 right). Over the same time period, NS-liquid droplets in conditions that disallow transcription (e.g. missing a key transcriptional component) do not display those morphological changes (Fig. 5.5 left). Since these structural modifications only present when transcription is allowed to occur, we attribute these features to interactions of the transcribed RNA with NS-liquids. Future work needs to examine how this DNA liquid/RNA interaction affects *Spinach* folding and to clarify whether it confounds the *Spinach* fluorescence signal that is used for transcriptional quantification.

In this chapter, we demonstrated that self-assembling NSs can be utilized as an *in vitro* platform to probe the complex interplay of genome organization and transcription. While we could not clearly determine if integration into a DNA liquid environment affects transcriptional output, we did observe that LLPS could alter RNAp activity by modulating the inhibitory effect of non-template DNA. Additionally, we also identified that transcribed RNA can notably alter NS-liquid structure. While this initial work has perhaps spawned more questions than answers, it attests to the need for (and value of) an *in vitro* model, where conditions can be more feasibly adapted & engineered (compared to *in vivo* systems) so as to address each new question (or opportunity, rather) that arises.

Chapter 6

Future outlook

6.1 DNA organization & transcription

The results described in chapter 5 demonstrate that even with this simple system, implemented so as to have greater control and to reduce complexity, the identification/clarification of specific effects remains far from trivial. In trying to rationalize our finding that the partitioning of genes into a NS-liquid environment does not significantly alter transcriptional output, several interesting questions arise that inspire future work.

Where does the majority of transcription occur? We confirmed that both T7 RNAp and Spinach2 dsDNA partition into NS-liquids (Fig. 4.4c). As partitioning is an equilibrium, both components also exist in bulk solution. Transcription can thus occur within the NS-liquid as well as in the bulk 'NS-gas' solution. As described for dextran (Fig. 4.1), a similar partitioning analysis should be done to quantify the relative concentrations of enzyme and template in the two distinct phases. It would also be interesting to determine whether the presence of promoter-containing dsDNA templates within NS-liquids alters the degree of T7 RNAp partitioning. Such quantitative details would aid interpretation, providing a better idea of the biochemical conditions of transcription.

How is transcriptional activity different, if at all, within the NS-liquid? While the current setup examines transcription from a system containing both phases (i.e. NS-liquid and 'NS-gas'), enzyme activity needs to be characterized in the NS-liquid phase alone. Indeed, we have yet to explicitly demonstrate that enzymatic activity can occur in NS-liquids. In that regard, it may be prudent to first test a simpler process, perhaps e.g. the enzymatic cleavage of X-gal by β-galactosidase (which generates a distinct color change), before continuing with transcription. These experiments would help determine whether the NS-liquid environment mirrors either: the increase in transcription seen with *P-gels*, crowded environments, and coacervates [155, 153, 154]; or the decrease that is observed within a dense DNA brush [156] and often attributed to densely packed heterochromatin (as described in section 1.2).

Once these questions (and others like it) are resolved, and we gain a better understanding of how the NS-liquid environment affects transcription, additional features and conditions of greater complexity can be added to the platform so as to better mimic chromatin environments. For instance, fully-stiff NSs and fully-flexible NSs (Fig. 2.2) could be combined to construct a system that mixes NS-networks and NS-liquids. Indeed, chromatin-containing membraneless organelles are comprised of networks of chromatin enveloped by coacervates/liquids composed of proteins and RNA. Structural proteins, like histones and NAPs, can eventually be added and engineered to form particular organized domains of DNA, which can again be probed for their effect on transcriptional activity. And, of course, future work could look at how transcription can be utilized to reorganize DNA, probing how different genes (or lncRNAs) affect or reinforce the structure of the chromatin mimic. For example, it needs to be determined whether the clusters and deformations exhibited by *Spinach* transcription is general to any transcribed RNA or if it is specific to sequence properties of the current product. Recent theoretical work has suggested that, in a mixture of polymers, differences in activity can induce phase

separation of active polymers from passive ones [167]. That is, transcriptional activity itself (regardless of the transcribed sequence) could induce DNA clustering; indeed, the presence of transcriptional factories (or foci) in the cell lends some credence to the theory [168, 169]. Ultimately, there are many possible scientific questions that can be pursued using the NS platform, which I attribute to the immense engineering capacity of the NS and of DNA as a whole. In the remainder of this dissertation, I present some of our early work with NS engineering.

6.2 Distinct NS species

Throughout this dissertation, I focused my discussion on only one type of DNA NS. However, the innate sequence-tunability of DNA confers significant engineering potential, perhaps the most obvious of which being the construction of multiple species, or "flavors", of NS-liquids simply through overhang design. Two different NSs can be constructed such that their respective palindromic overhangs exhibit minimal (if any) complementarity to one another: 5'-$CGATCG$-3' and 5'-$GAGCTC$-3'. In conditions that enable self-assembly (e.g. sufficient salt and low temperatures), a mixed population of the two species of NSs will self-segregate into a binary system of two immiscible NS-liquids (Fig. 6.1a left). While droplets of the same NS species will coalesce upon contact, droplets of different species "bounce off" each other upon contact.

Different NS-liquid droplets can be engineered to interact through addition of a 'DNA surfactant', which is simply a NS with heterogeneous overhangs, e.g. two overhangs complementary to NS species "A" and two overhangs complementary to NS species "B" (Fig. 6.1b). The degree of interaction is dependent on the ratio of the three NS components. With a low amount of surfactant (e.g. 1/4 of the concentration of the other NSs), the distinct NS-liquid species are able to interact, forming flat interfaces with each other

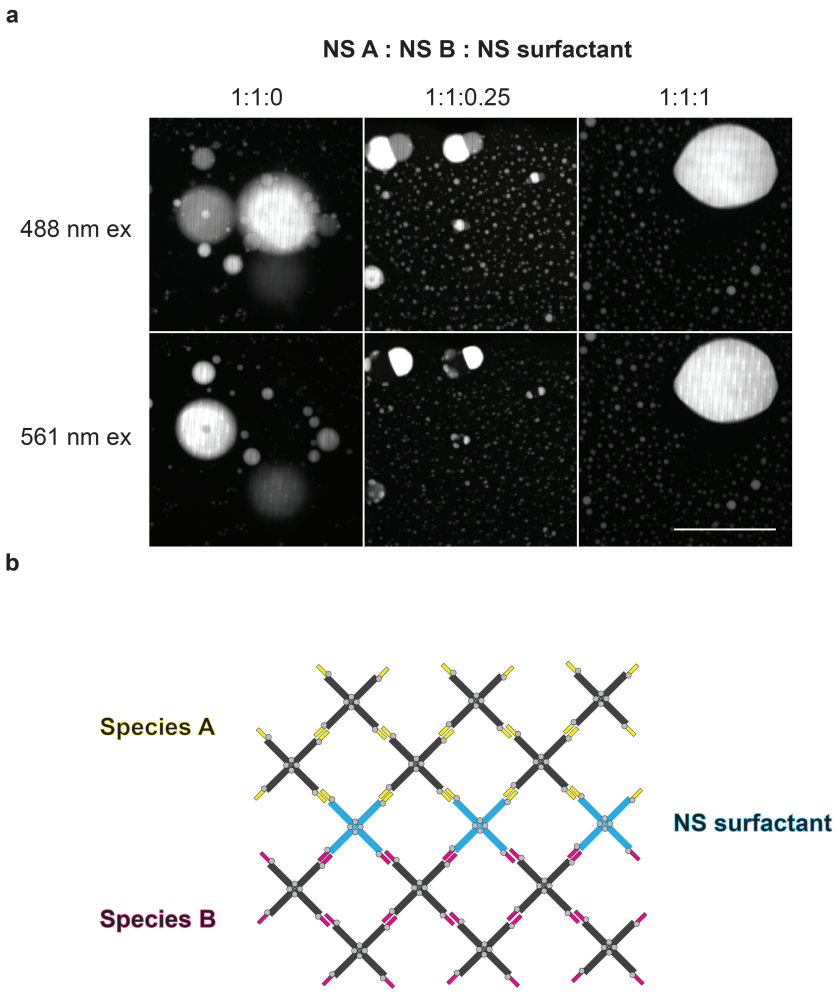


Figure 6.1: **Multiple NS species and NS surfactants. (a)** The degree of inter-species NS interaction is dependent on the amount of NS surfactant added, ranging from 0 to equimolar (left to right). NS A is labeled with Cy3; NS B is labeled with FAM. Cy3 is visible at both excitation wavelengths, as no emission filters were used. Scale bar = 100 μm. **(b)** Cartoon model of NS surfactant interface.

(Fig. 6.1a middle), where DNA surfactants presumably constitute the interface (Fig. 6.1b). When all three components are present in an equimolar ratio, we actually see no separation of distinct NS-liquid droplets (Fig. 6.1a right); there is sufficient DNA surfactant to support a well mixed system.

6.2.1 Partitioning into different NS-liquids

These distinct, immiscible compartments are an *in vitro* realization of membraneless organelles. Biomolecules can be engineered, through sequence complementarity, to specifically partition into the different NS-liquid species. For example, dsDNA with overhangs complementary to NS B and streptavidin modified with biotinylated-ssDNA also complementary to NS B, will only partition into B droplets, completely avoiding A droplets (Fig. 6.2).

Future work could strive to demonstrate enzymatic activity only in specific NS droplets, further supporting the application of NS-liquids as model membraneless organelles. Additionally, in conjunction with NS surfactants, NS-liquid droplets containing different components could be induced to fuse, bringing components together for binding (or even enzymatic activity) only after the addition of NS surfactant.

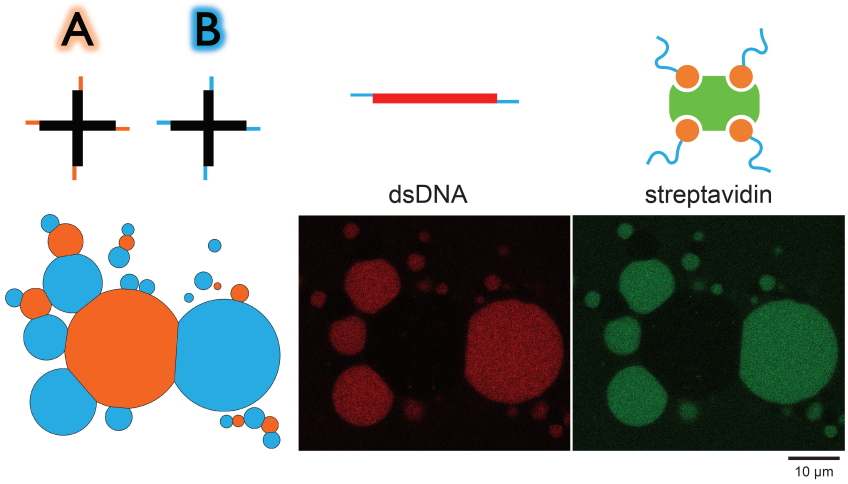


Figure 6.2: **Specific partitioning.** In a mixture containing multiple species of NS-liquids (i.e. A and B), streptavidin and dsDNA exhibiting ssDNA oligos and overhangs, respectively, that are complementary to NS B will specifically partition only into NS-liquid B species.

6.3 Transcriptional control of self-assembly

6.3.1 Single-stranded nucleic acid bridge

One advantageous feature that has been attributed to membraneless organelles is temporal control of biomolecular interactions, which is achieved by dictating when the liquid compartments form. Researchers have demonstrated, for example, that transcription can trigger liquid-liquid phase separation, where the newly generated negatively-charged RNA can form a complex coacervate with positively-charged polypeptides [101]. It is possible to mirror this using NS-liquids but, instead of any negatively-charged sequence of RNA, NS-liquids will only form when a specific sequence is introduced.

NSs can be designed to have non-palindromic overhangs, forcing them to remain dissolved in solution regardless of the environmental conditions (i.e. even at high salt concentration and low temperatures). In a mixture containing two species of NSs, one with non-palindromic 5' overhangs and the other with non-palindromic 3' overhangs, introduction of a single-stranded nucleic acid with sequence complementarity to both NSs will enable self-assembly into a condensed NS phase (Fig. 6.3). Following similar concepts as described in chapter 2, if the hybridization interaction is very strong, a percolated network will form; however, should hybridization leave ssDNA gaps that break base-stacking or, more generally, if the hybridization interaction is weak, then NS-liquids will form (Fig. 6.3b). Instead of using ssDNA, one can imagine how ssRNA can be used and how the phase of self-assembly can then be transcriptionally controlled through expression of a particular length of RNA (i.e. a gene). Another variation of transcriptional control would be the ability to dictate which NS species in a mixed population can interact. By transcribing different ssRNA linkers (i.e. genes), one could control whether NSs A and B self-assemble or whether NSs A and C do. Conversely, degradation could be triggered through the use of common nucleic acid engineering strategies, like addition

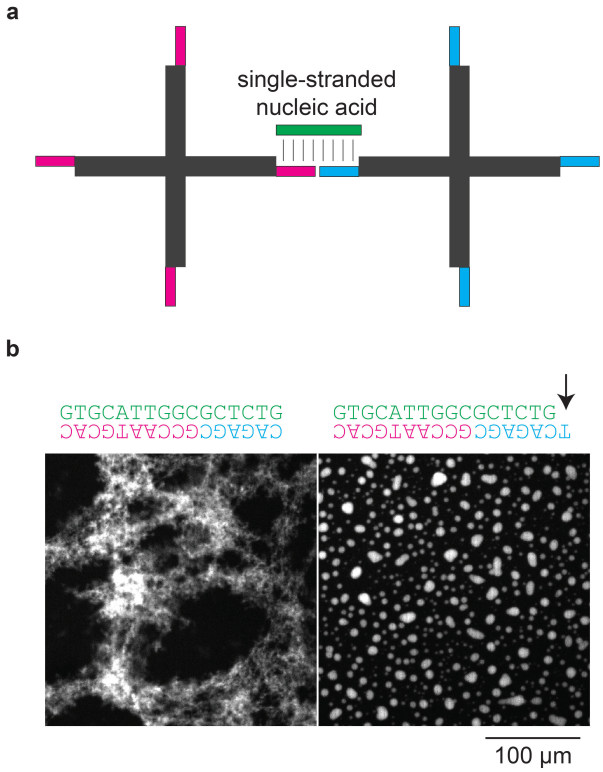


Figure 6.3: **NS self-assembly with a ssDNA bridge. (a)** In a mixture of NSs with 5' overhangs (pink) and 3' overhangs (blue), addition of ssDNA with regions complementary to both NSs enables self-assembly into larger complexes. **(b)** Hybridization strength determines whether the resulting complex is a NS-network (left) or NS-liquid (right). The sequences shown represent overhangs, where the colors match those in (a). Right: A base was intentionally left unpaired upon hybridization, weakening the binding interaction and leading to formation of a NS-liquid.

(or transcription) of an inhibitor strand that binds competitively to the linker ssRNA, leading NSs to dissolve back into solution.

6.3.2 Hybridization chain reaction

Transcription could also be used as a means to trigger formation of NSs themselves. Inspired by the hybridization chain reaction [170], NSs can be designed to form from a system of 4 DNA hairpins that are stable in solution. Introduction of an initiator strand can displace the first hairpin (via a toe-hold mechanism [171]), revealing ssDNA

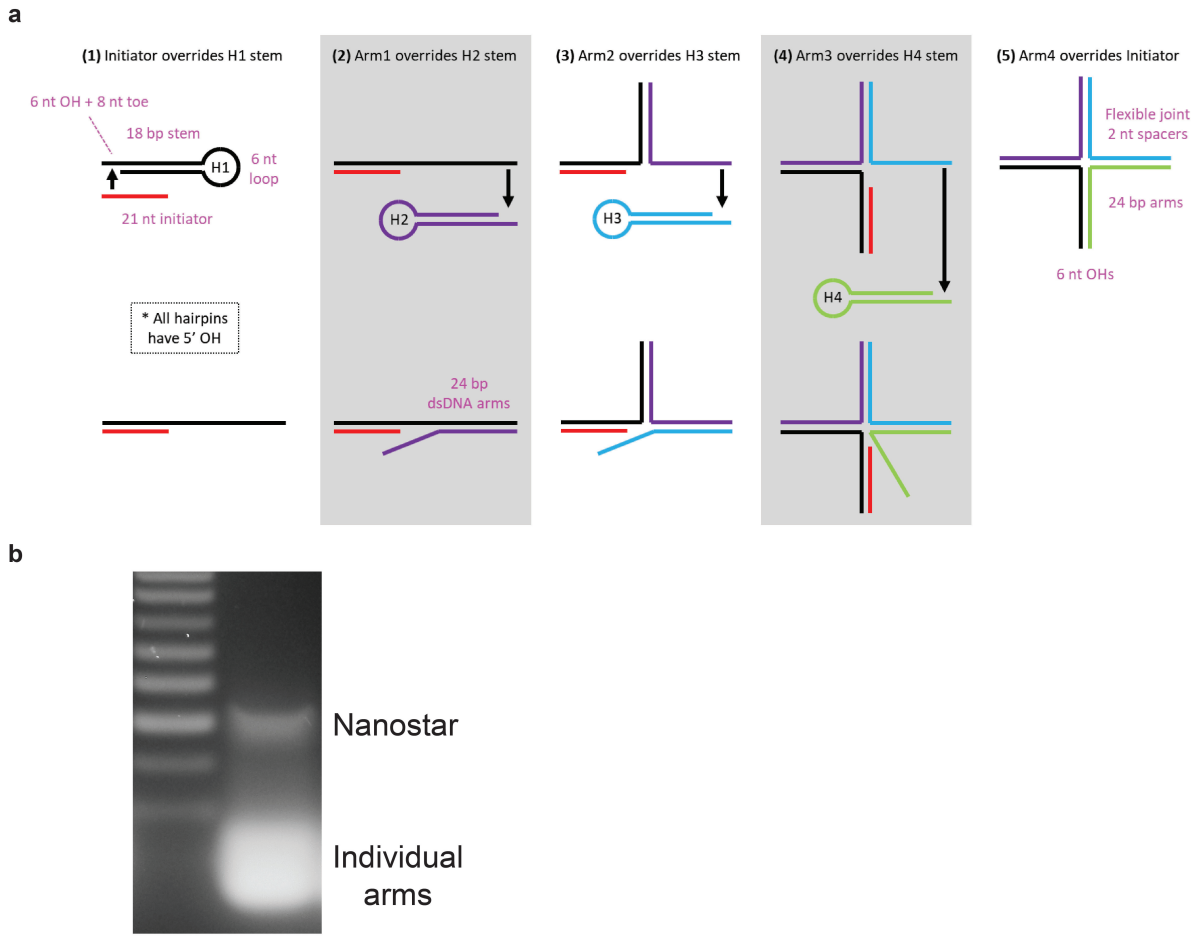


Figure 6.4: **Hybridization chain reaction. (a)** Scheme: Addition (or transcription) of an initiator strand triggers NS self-assembly. **(b)** Agarose gel electrophoresis shows some formation of NSs after initiator strand addition. Efficiency is low, where most of the hairpins remain unopened.

that can displace a second hairpin, and so on (schematically shown in Fig. 6.4a). If sequences are designed appropriately, a 4-armed NS can be formed by simply adding the initiator strand. Preliminary experiments have demonstrated feasibility (Fig. 6.4b), though sequences and conditions need to be tuned to optimize NS formation efficiency. Future work could then verify whether these correctly formed NSs can themselves go on to self-assemble into larger complexes. Eventually, addition of the initiator strand could be controlled through transcription.

6.4 Final word

In this dissertation, I described our efforts to develop an *in vitro* platform to study the interplay between DNA organization and transcription. While much of the work focuses on physical details regarding condensed NS phases, we do demonstrate the functional application of NS-liquids as a model of membraneless organelles. Future work can continue to study how chromatin structure and transcription affect each other, perhaps evolving to integrate increasingly more complex components so as to better mimic chromatin organizational states. Furthermore, the sequence tunability of DNA confers significant adaptability to the NS system, enabling it to potentially be applied to study other biological processes beyond transcription.

Appendix A

Materials and Methods

A.1 Sample preparation

A.1.1 NS annealing

DNA NS were made by annealing four ssDNA oligos following similar protocols for constructing branched DNA nanostructures [172, 100], namely a slow temperature ramp. Oligos were re-suspended to 100 μM in 10 mM Tris buffer (pH 7.5). Equimolar mixtures were prepared (e.g. 25 μM each oligo) and incubated at 95 °C for 5 min and cooled to 4 °C at 0.5 °C per min. Efficient NS formation was verified using agarose gel electrophoresis, where annealed NSs were loaded into 3% agarose gels and run in TAE buffer to confirm the presence of a single dominant NS band. DNA was visualized using SYBR™ Safe dye and a UV transilluminator. Annealed NSs were stored at -20 °C. Before use, NSs were thawed to room temperature and, in the case of strong overhang hybridization, incubated between 40-55 °C for >15 min to ensure that NSs are all dissolved in solution. NSs were eventually diluted/utilized in experiments at a typical working concentration of 10 μM.

To examine potential salt-dependent structural changes of DNA NSs (e.g. Open vs.

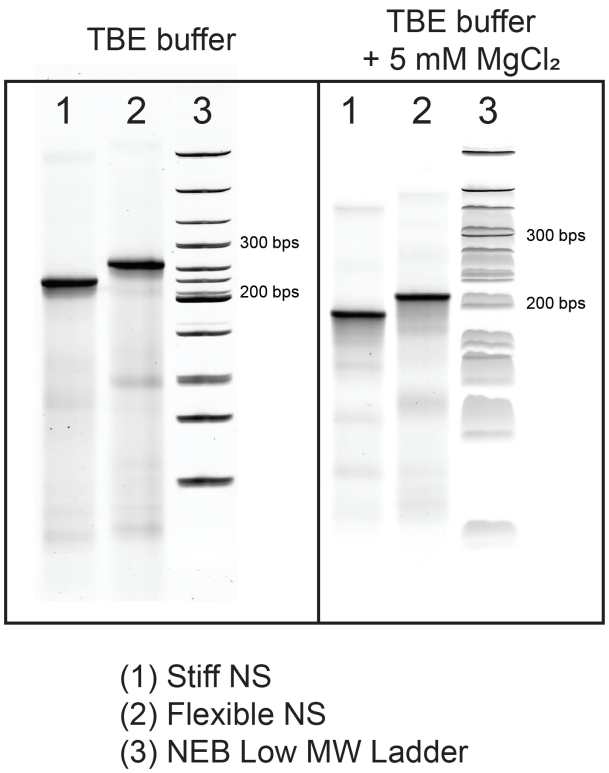


Figure A.1: **DNA nanostars**. NSs were examined using PAGE to probe the efficiency of formation from the annealing procedure. NSs lacking overhangs were used so as to remove the possibility of sticky-end interactions, thereby leaving two samples for analysis: blunt stiff NSs and blunt flexible NSs. Each sample exhibits a single dominant band of correctly formed 4-arm NSs. As expected, flexible NSs, with their additional bases between NS arms, migrate more slowly than stiff NSs. When PAGE is repeated in the presence of 5 mM $MgCl_2$, both stiff and flexible NSs migrate further than in the absence of added salt, relative to the DNA ladder (right). This increased migration supports the notion that NSs adopt a more compact Stacked-X conformation in the presence of $MgCl_2$.

Stacked-X conformations), NSs were examined via PAGE (Fig. A.1). Annealed NSs were loaded into 8% polyacrylamide gels and run under nondenaturing conditions. DNA was stained using SYBR™ Gold dye and gels were imaged using an Amersham Typhoon 5 Biomolecular Imager. When PAGE was performed with added $MgCl_2$, the salt was added to DNA samples, to the polymerized gel, and to the running buffer. Samples were run at 4 °C to minimize additional heating due to the presence of $MgCl_2$.

When fluorescent visualization of NSs was required in microscopy experiments, NSs were modified by either:

- Non-covalent interaction with YOYO-1 dye: The intercalating dye is added at a 1:1 dye to NS ratio and can be visualized through excitation with 488 nm light.
- Fluorescent modification of a:
  - **dsDNA arm:** One of the oligos comprising a NS, specifically oligo v1 (see Appendix B for the DNA sequence), was modified by the manufacturer (Integrated DNA Technologies) so that the third thymidine in the sequence, which resides near the middle of a 20-base pair double-stranded DNA arm in a NS, has fluorescein attached via a 6-carbon spacer to position 5 of the thymine ring. The alteration does not dramatically perturb Watson-Crick base-pairing of the modified thymidine.
  - **ssDNA overhang:** One of the oligos comprising a NS, specifically oligo v1 (see Appendix B for the DNA sequence), was modified by the manufacturer (Integrated DNA technologies) to have a fluorescent dye (e.g. fluorescein or Cy3) positioned adjacent to the unpaired base on the overhang.

Unless stated otherwise, NS samples were prepared in 10 mM Tris buffer (pH 7.5). After the addition of salt, fluorescent dye, and other relevant components (depending on the experiment), samples were mixed via pipette before use. When pre-formation of NS-liquid droplets was required, samples were incubated in the dark at room temperature on a rotator for >1 hr. When high concentrations of salt were utilized, samples were heated at 55-60 °C for 15 min to ensure that NSs remained diffusing in solution before addition to a flow cell (or other sample holder).

A.1.2 dsDNA linker preparation

For dsDNA to integrate into macromolecular assemblies comprised of NSs (i.e. NS-liquids or NS-networks), they must be functionalized with ssDNA overhangs that are complementary to those exhibited by the NSs. This can be accomplished using an adaptation of the polymerase chain reaction (PCR), known as 'auto-sticky PCR' [148], where forward and reverse PCR primers are designed so that each includes an internal abasic site that separates an overhang sequence from the priming sequence. DNA polymerases (DNAps) stall when they encounter abasic sites and thus leave 5' ssDNA overhangs on their dsDNA PCR products.

When determining the annealing temperature for PCR reactions, one can effectively ignore the overhang sequence. However, it is important to check that the overhang sequence, alone or in tandem with the primer sequence, does not exhibit significant homology to unintended regions of the PCR template (e.g. plasmid), as inadvertent DNA amplification can occur. To minimize non-specific binding to unintended regions, a higher annealing temperature (e.g. additional 2 °C) could be utilized.

Regarding the choice of DNAp for PCR, either *Taq* or Phusion®, there are a couple details to consider for generation of dsDNA with ssDNA overhangs.

- Proofreading/Exonuclease activity: Phusion® DNA polymerase, which is a fusion protein comprised of Pfu DNAp and the Sso7d DNA binding protein, exhibits higher fidelity (i.e. fewer errors/mis-incorporations) than *Taq* DNAp [173]. This increased fidelity is a result of having 3' to 5' exonuclease proof reading activity. When performing PCR, one must be careful that this exonuclease activity does not digest PCR primers, which would significantly affect the ability for primers to bind specifically and optimally at a particular annealing temperature. As such, when using Phusion®, always add the enzyme last, and keep everything on ice, so as to

minimize 3' exonuclease activity on the primers.

- Addition of 3' adenosine: *Taq* DNAp, which lacks 3' to 5' proofreading ability, frequently adds a non-templated nucleotide, predominantly adenosine, to the 3' terminus of blunt dsDNA [174]; this feature has been famously exploited in TA cloning [175]. *Taq* DNAp is also known to insert adenosines across from abasic sites [176].

- Insertion across from abasic sites: Behavior across from abasic sites differs between DNAps. The original auto-sticky PCR work utilized *Taq* DNAp, which (as mentioned) inserts a base, typically an adenosine, across from an abasic site but will stall afterwards. That is,

  > *"...the polymerase inserts predominantly dAMP opposite to tetrahydrofuran (1'2'-dideoxyribose) abasic sites, but is unable to continue the synthesis of the complementary strand"* [148]

  Pfu DNAp, on the other hand, stalls before the abasic site; that is, it does not fill the space across from an abasic site [177]. As such, PCR using Phusion® DNAp would leave an overhang that includes the abasic site, while use of *Taq* would leave an overhang that does not include the abasic site (See fig. A.2).

Figure A.2: **DsDNA with ssDNA overhangs.** Different DNAps act differently across from abasic sites, generating dsDNA with slightly different overhangs after PCR with primers containing internal abasic sites.

Upon completion of PCR, DNA was purified using commercial filtration systems (e.g. Zymo DNA clean & concentrator™ or Amicon® Ultra centrifugal filters).

A.1.3 Protein purification

T7 RNAp

T7 RNAp purification was performed in collaboration with Greg Ekberg and Professor Christopher S. Hayes. Overnight LB cultures of *E. coli* strain X90 [178], transformed with plasmid pCH3943-pT7-911Q [179] (see Appendix B for DNA sequences), were diluted into 100 ml of fresh LB to an OD_{600} of 0.05. Cultures were incubated at 37 °C with shaking at 250 rpm to an OD_{600} of 0.6 before inducing protein expression with 1.5 mM IPTG for 1 hr. Cells were harvested over ice and centrifuged for 10 min at 4,000 rpm (Sorvall GSA rotor). Cell pellets were resuspended in 10 ml extraction buffer (20 mM sodium phosphate, pH 7.0; 150 mM sodium chloride; 0.05% Triton X-100; 10 mM DTT; 1 mM PMSF) and lysed by french press at 20,000 psi. Lysates were cleared by two

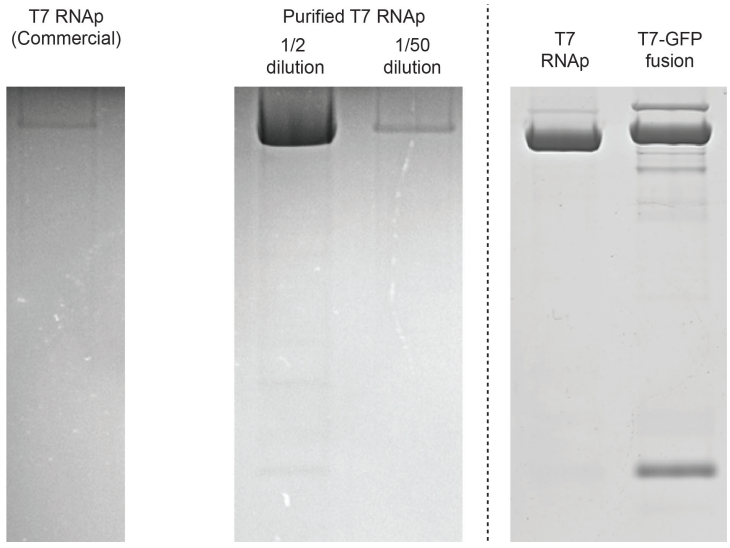


Figure A.3: **T7 RNAp purification.** Samples are identified with a label at the top of each lane.

consecutive rounds of centrifugation at 14,000 rpm for 10 min (Sorvall SS-34 rotor) and supernatants were incubated with Ni^{2+}-NTA (Qiagen) for 3 hr at 4 °C. Ni^{2+}-NTA was washed with wash buffer (20 mM sodium phosphate, pH 7.0; 150 mM sodium chloride; 0.05% Triton X-100; 10 mM β-mercaptoethanol; 20 mM imidazole) and loaded onto a Poly-Prep column (Bio-Rad) at 4 °C. T7 RNAp-His 6 was then eluted from Ni^{2+}-NTA with native elution buffer (20 mM Tris-HCl, pH 7.9; 10 mM DTT, 250 mM imidazole). Purified T7 RNAp were dialyzed twice against 1 L of dialysis buffer (10 mM Tris, pH 7.9; 100 mM sodium chloride; 10 mM DTT, 1mM EDTA). Purified protein was mixed with an equal volume of ice cold glycerol and vortexed thoroughly. Samples were filtered with a 0.22 μm filter and stored at -20 °C. PAGE verified the purity of the T7 RNAp preparation (Fig. A.3).

T7 RNAp-GFP fusion

Cloning was performed to fuse Superfolder GFP (sfGFP) [180] to the C-terminus of T7 RNAp, which already exhibits an N-terminal His-tag. The primers utilized for cloning are listed in the following table; for plasmid sequences, see Appendix B. N-terminal T7 RNAP was amplified off of plasmid pCH3943-pT7-911Q using primers 4489/4490. sfGFP was amplified off of plasmid pCH762-pBSK-sfGFP using primers 4491/4492. Fragments were fused by overlapping PCR using primers 4489/4492. Fragments were digested and cloned into plasmid pCH3943-pT7-911Q, generating plasmid pCH14252-pT7-911Q-sfGFP. The fusion was purified using the same protocol applied to T7 RNAp. While a notable contaminant of low molecular weight is observed via PAGE (Fig. A.3 right), GFP-fused T7 RNAp is present at a far higher concentrations. These samples were only utilized for fluorescent visualization and not to perform transcription. Notably, the fusion is indeed able to perform T7 promoter-dependent transcription.

Primer #	Oligo name	Sequence
4489	Eco-T7pol-Fwd	CACACAGAATTCATTAAAGAGG
4490	T7-gfp-fus-rvs	GCCCCCGGACCCGCCACTAGTCGCGAACGCGAAG
4491	T7-gfp-fus-fwd	GGCGGGTCCGGGGGCTCTAAAGGTGAAGAACTG
4492	Nhe-GFP-Rvs	GACGCTAGCTTATTTGTAGAGCTCATCC

GFP

Emerald GFP [181], containing an N-terminal His-tag and purified via Ni^{2+}-NTA chromatography, was graciously provided by Thomas Nguyen, Claire Tran, and Professor Irene A. Chen.

A.2 Microscopy

A.2.1 Flow cell preparation

For microscopic visualization, samples were prepared in one of two types of sample holders.

Capillary tubes: Borosilicate capillary glass slides (0.1 x 1.0 mm inner dimensions; Electron Microscopy Sciences) were washed in a series of solutions (D.I. water, acetone, isopropanol, and D.I. water), dried with nitrogen gas, and plasma cleaned before DNA samples were wicked into the tubes. The ends were immediately sealed with Devcon® 5 minute epoxy and subsequently incubated in a dark, humid chamber. The DNA samples were visualized using an inverted fluorescence microscope (Nikon Eclipse TE2000-U) at 10x magnification. Large composite images were constructed using the Stitching plugin in FIJI ImageJ [182, 183]. Confocal microscopy was performed using the Olympus Fluoview 1000 Spectral Confocal and 3D images were rendered using the Imaris software.

Sealed flow cell: Glass coverslips were washed in a series of solutions (0.1 M NaOH, deionized water, acetone, isopropanol, and deionized water), dried with nitrogen gas, and plasma cleaned. When performing microscopy on inverted microscopes with high

magnification and immersion objectives, no. 1.5 coverslips were used for the bottom coverslip to minimize aberrations. Parafilm with an internal channel (cut via razor by hand or via laser cutter) was sandwiched between the cleaned coverslips and melted with a soldering iron. The channel was filled with sample via pipette and subsequently sealed using Devcon® 5 minute epoxy (Fig. A.4).

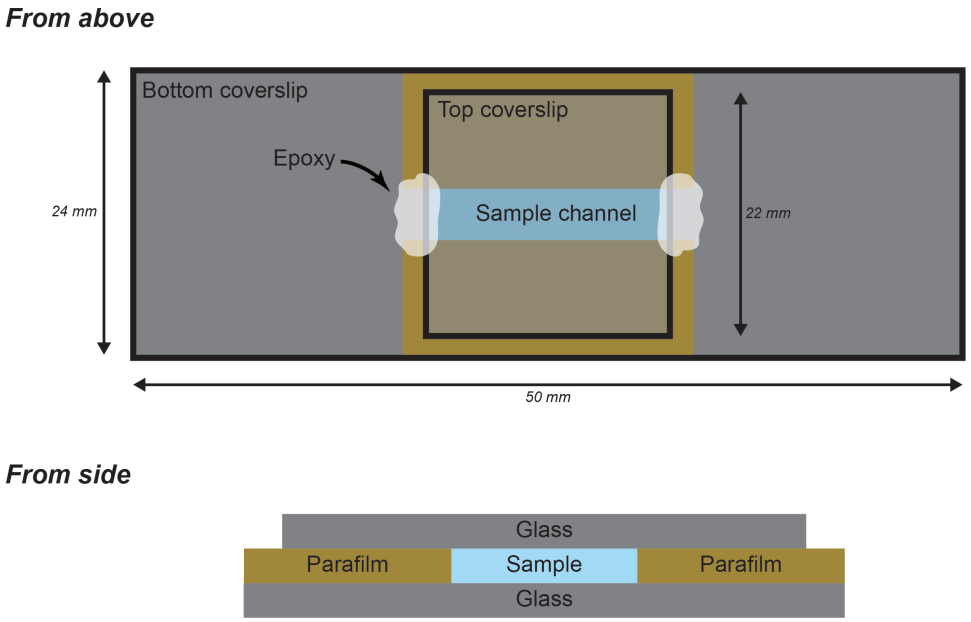


Figure A.4: **Typical flow cell.** Flow cells were utilized as simple sample holders that allowed microscopic visualization.

Flow cells w/ permeable membranes: Permeable-membrane flow cells were prepared as previously described [83]. Briefly, glass coverslips (no. 1) were washed in a series of solutions (0.1 M NaOH, deionized water, acetone, isopropanol, and deionized water), dried with nitrogen gas, and plasma cleaned. Coverslips were then incubated with a 0.5% (v/v) solution of 3-(trimethoxysilyl)propyl methacrylate in isopropanol at 80 °C for 1 hour. A piece of parafilm was laser-cut with an internal cut out and sandwiched between the chemically-treated coverslips. A soldering iron was used to melt the parafilm to form a sealed channel with two openings. A pre-mix of polyethylene glycol-diacrylate (PEG-DA, 400 Da, Polyscience, Inc.) and a photo-initiator, 2-hydroxy-2-methylpropiophenone

(Sigma-Aldrich), was prepared in a 50:1 ratio. The premix was diluted with deionized water to obtain a 20% (v/v) solution, injected into the channel via a pipette, and exposed to UV radiation through a mask so as to generate two PEG-DA permeable membranes. The resulting flow cell consists of a central channel separated from two buffer exchange channels by permeable membranes. The DNA sample was flowed into the central channel via a pipette and capillary tubes (Paradigm Optics, Inc.) were inserted into both ends of the exchange channels. The other ends of the capillary tubes were encased with Tygon® tubing, before the entire flow cell was sealed with Devcon® 5 minute epoxy. A syringe pump was used to withdraw solutions via the Tygon® tubing to exchange buffers in the exchange channels. DNA samples were visualized using an inverted fluorescence microscope (IX71, Olympus) at 10x magnification.

A.2.2 Microscope systems

Several microscopes were used to conduct the work described in this dissertation. While specific systems are referenced in previous sections, I summarize them all here.

- Nikon Eclipse TE2000-U inverted microscope: Chapter 2
- Olympus Fluoview 1000 spectral confocal microscope: Chapter 2
- Olympus IX71 inverted microscope: Chapter 2
- Leica SP8 resonant scanning confocal microscope: Chapter 3, 4, & 6
- Visitech VT-Infinity3 confocal scanning module: Chapter 3, 4, 5, & 6

Generally speaking, microscope settings (e.g. excitation/emission wavelengths, exposure time, gain, etc.) were selected so as to minimize fluorophore bleaching and to produce a high signal intensity without pixel saturation.

A.2.3 Fluorescence recovery after photobleaching (FRAP)

Total intensity (Chapter 2): Regions of interest (ROIs) in DNA hydrogel samples were bleached using 488 nm excitation and imaged every 10 seconds for 15 minutes. For experiments examining recovery at longer time scales, samples were incubated in a dark, humid chamber when not being imaged. Fluorescence intensities within ROIs were corrected for bleaching during image acquisition by simultaneously imaging un-bleached regions. The corrected intensities were normalized as a fraction of pre-bleach values.

Radial profile (Chapter 3): Fluorescent NS-liquid droplets, containing a 99:1 mixture of untagged:Cy3-tagged NSs, were imaged using a Leica SP8 confocal microscope. Image acquisition and data analysis was carried out following the method described by Seiffert and Oppermann [116]. Briefly, samples were imaged (512 x 512 pixels) using an objective with low numerical aperture (10x; 0.3 NA) with 20x zoom. A point was bleached with 550 nm excitation (10 sec at 50% laser power, which was measured to be about 0.5 mW at the sample level) and fluorescent recovery was examined over $558-717$ nm emission at 2, 5, and 5 sec intervals for 0.25, 0.5, and 1 M [NaCl], respectively. The average

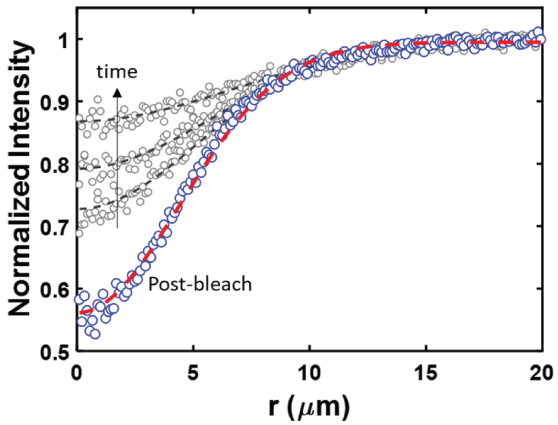


Figure A.5: **Recovery of radial profile.** Normalized averaged intensities at radial distance r from the bleach spot at 0 (blue), 1, 2, and 5 min after bleaching at [NaCl] = 0.25 M. Experiments are carried out at 20 °C. At each time point, the data are fit to $I(t,r) = I_0(t) - A(t)exp(-r^2/2v(t)^2)$.

fluorescent intensity, $I(r)$, at a given radius from the bleached spot, r, was calculated at each recovery time point and fit to the equation $I(t,r) = I_0(t) - A(t)exp(-r^2/2v(t)^2)$ (Fig. A.5), where v^2 is the time-dependent variance of the Gaussian function. The NS diffusion coefficients at different [NaCl] are obtained from plots of v^2 vs. time as shown in Fig. 3.4.

A.3 NS-liquid characterization

A.3.1 Density & Volume fraction

100 μL solutions of 5–15 mg/mL DNA NS were prepared with 0.25, 0.5, or 1 M NaCl, with 1:100 molar ratio of YOYO-1:NS. The NS solutions were incubated at 50 °C for 30 min and centrifuged at 3,000 x g for 100 min. Using a transilluminator (Invitrogen) for visualization (Fig. 3.1 right), the DNA-dilute phase was removed via pipette. The DNA concentration of the resulting NS-liquid was then measured via absorbance at 260 nm using a Nanodrop One spectrophotometer (Thermo Fisher Scientific). Volume fraction was calculated as $\phi = c_{DNA}/(1.7\ \mathrm{g/cm^3})$ [184].

We confirm that the effect of YOYO-1 dye on DNA NS phase separation is negligible. With YOYO-1, at 0.5 M NaCl, the [DNA] in the dense NS-liquid is measured to be 25 ± 0.3 mg/mL; without YOYO-1, at the same 0.5 M NaCl, [DNA] is measured to be 25 ± 0.4 mg/mL. We also confirm that [DNA] in the NS-liquid phase is independent of the overall [DNA] in the NS solution (Fig. 3.2) for all three explored NaCl concentrations.

Sedimentation Velocity

To confirm that the DNA concentration in NS-liquid droplets matches that of the bulk NS-liquid phase, we measured density by examining the sedimentation velocity of NS-

liquid droplets. Stokes Law dictates the drag force, F_{drag}, exerted on hard spheres when moving through a viscous fluid with low Reynolds number [185] as $F_{drag} = 6\pi\eta_{fluid}Rv$, where η_{fluid} is the fluid viscosity, R is the radius of the spherical object, and v is the velocity. For a spherical particle sedimenting at terminal velocity, the drag force is equal to the force due to gravity, $F_g = \frac{4}{3}\pi R^3\Delta\rho g$, where $\Delta\rho$ is the density difference between the particle and the surrounding fluid and g is gravitational acceleration (g=9.8 m/s^2), yielding:

$$\Delta\rho = \frac{9}{2}\frac{v\eta_{fluid}}{gR^2} \tag{A.1}$$

We took confocal volumes, with a depth of 150 μm, of sedimenting NS-liquid droplets and extracted terminal sedimentation velocities at [NaCl] = 0.25, 0.5, and 1 M, as shown in Fig. A.6, to calculate $\Delta\rho$ (Eq. A.1). Assuming that NS-liquid droplets are composed only of DNA nanostars and the surrounding fluid, one can obtain the DNA concentration in the NS-droplets as:

$$\phi_{DNA} = \frac{\rho_{droplet} - \rho_{fluid}}{\rho_{DNA} - \rho_{fluid}} \tag{A.2}$$

where ρ_{DNA} is taken to be 1700 mg/mL and ρ_{fluid} is 1008, 1019, and 1040 mg/mL for 0.25, 0.5, and 1 M NaCl, respectively [186].

The concentration of DNA in the droplet phase increases from 12 ($\pm$ 5) to 24.0 ($\pm$ 9) mg/mL with an increase in [NaCl] from 0.25 to 1 M, consistent with the DNA concentrations measured from bulk phase separation. Slightly lower values from the sedimentation method could be attributed to the limited z-resolution in tracking sedimenting NS droplets, our assumption that the surrounding solution lacks NS (whose presence would slightly alter η_{fluid} and ρ_{fluid}), and possible deformation of the sedimenting droplets.

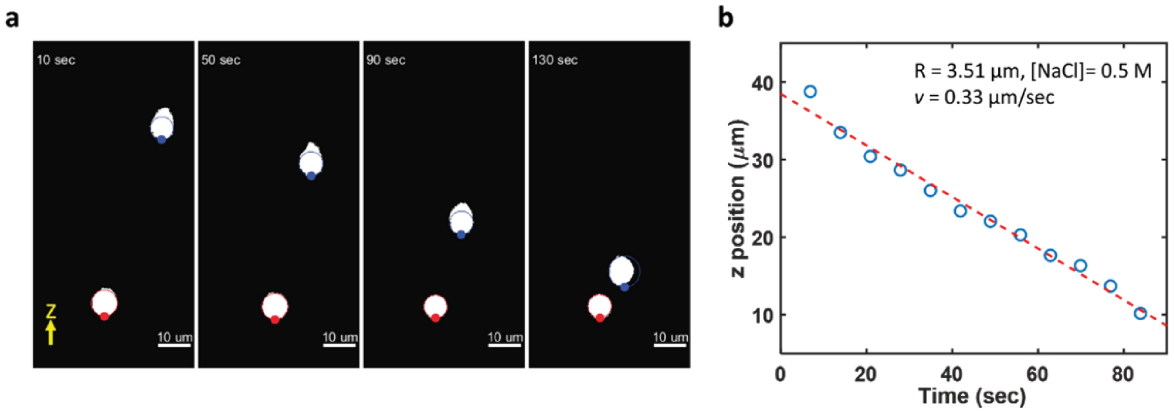


Figure A.6: **Sedimentation velocity.** **(a)** Binarized images of projections onto xz-plane of a sedimenting droplet. The droplet encircled in red is settled on the surface while the droplet encircled in blue is sedimenting. Panels correspond to 10, 50, 90, and 130 sec, respectively. Experiments are carried out at 20 °C. **(b)** A plot of z position of a sedimenting droplet vs. time at [NaCl]=0.5 M.

A.3.2 Viscosity measurements via microrheology

Mixtures of NS and 200 nm fluorescent beads (540 ex/560 em, FluoSpheres^{TM} from Thermo Fisher Scientific) were added into 10 μL of tris buffer at [NaCl] of 0.25, 0.5, and 1 M and incubated overnight at room temperature on a rotator, thereby allowing formation of large (>50 μm) spherical NS droplets. NS-liquid droplets were introduced into a flowcell (i.e. a coverslip-parafilm-coverslip sandwich with a channel cut in the parafilm) using end-cut pipette tips, and subsequently sealed with epoxy. The mobility of fluorescent beads in NS-liquids were tracked with confocal microscopy for >100 min with 1 sec intervals. MSDs were calculated and fit to the form $MSD(\tau) \sim 4D_{probe}\tau^{\alpha}$, where α is the diffusive exponent, to estimate the diffusion coefficient D_{probe}. Viscosities η of DNA NS-liquids at different [NaCl] were then calculated using the Stokes-Einstein equation, $D_{probe} = k_B T/6\pi\eta r$, where k_B is the Boltzmann constant, $T = 293$ K is temperature, and $r = 100$ nm is the probe radius.

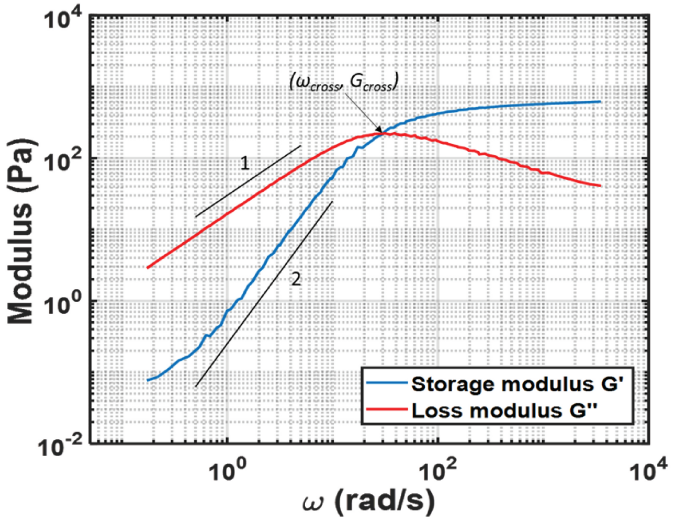


Figure A.7: **Bulk rheology.** Storage and loss moduli (G' and G") of the NS-liquid phase at 0.5 M NaCl at 20 °C from bulk rheology. The viscosity η is calculated from the crossover modulus G_{cross} and frequency ω_{cross} as $\eta = 2\pi G_{cross}/\omega_{cross} \approx 45.86$ Pa·sec. The solid lines are guides for slopes of 1 and 2.

Bulk rheology

To validate the viscosity results from microrheology, we measured viscosity of the NS-liquid phase using a conventional rheometer (ARG2, TA instruments). NS-liquid (>0.5 mL) was extracted from bulk scale phase separation at [NaCl] = 0.5 M and transferred to a rheometer with 20 mm diameter flat plates and a 0.2 mm gap. The storage and loss moduli were measured over a range of frequencies at 5% strain, which is confirmed to be in the linear regime. The viscosity was calculated using $\eta = 2\pi G_{cross}/\omega_{cross}$, where G_{cross} and ω_{cross} are the modulus and frequency, respectively, of the point at which the storage and loss modulus curves intersect, as shown in Fig. A.7. At 20 °C, the viscosity was calculated to be about 46 Pa·sec, very close to the viscosity measured from microrheology, which was about 45 Pa·sec at [NaCl] = 0.5 M.

A.3.3 Surface tension measurements

Coalescence was observed for fluorescent NS-liquid droplets on a flat oil/water interface [187]. A flow cell was prepared where the surface of the bottom coverslip was made hydrophobic with Sigmacote® (Sigma-Aldrich). 2% of PFPE-PEG-PFPE Triblock-copolymer surfactant (E2K0660; RAN Biotechnologies, Inc.) in 3M™ Novec™ 7500 Engineered Fluid (3M) was prepared and wicked into the flow cell. Then, an aqueous solution of pre-formed NS-liquid droplets in the salt of interest was flowed into the flow cell. The result is a flat oil/water interface where individual NS-liquid droplets can diffuse laterally, with minimal friction, to encounter and coalesce with other droplets. Coalescence events were recorded with confocal microscopy (Visitech), where 3D volumes were captured with 10 sec intervals at 561 nm excitation.

To calculate surface tensions, we use a formulation by Leal that shows, at late stages, the relaxation timescale τ of the transformation of an elliptical droplet into a spherical one is dictated by the ratio of the internal and external viscosities ($\lambda = \eta_{int}/\eta_{ext}$), the radius of the droplet R, and the surface tension σ [117].

$$\tau = \frac{(2\lambda + 3)(19\lambda + 16)}{40(\lambda + 1)} \frac{\eta_{ext} R}{\sigma} \tag{A.3}$$

When $\lambda \gg 1$ (i.e. $\eta_{int} \gg \eta_{ext}$), the above equation is reduced to equation 3.1. We calculate τ as the decay time scale of the dimensionless parameter $A \equiv (L-W)/(L+W)$, which is a ratio of the difference and sum of the length (L) and width (W) of a deformed droplet during coalescence (Fig. 3.5). Combining those results with measured values of droplet radii and estimated viscosity values obtained from microrheology, surface tensions of NS-liquid droplets in various [NaCl] can be calculated from equation 3.1.

A.3.4 Overhang binding probabilities

To estimate the effect of salt on NS interactions, we calculate the binding probabilities of overhangs. To do this, we use a mean-field picture in which we ignore connectivity of the overhangs within a NS. That is, individual overhangs are assumed to be freely translating at a concentration set by the measured c_{DNA} in the DNA liquid phase. Using the c_{DNA} data from Table 3.1, the NS overhang concentrations are 1.28, 1.66, and 2.34 mM for [NaCl]=0.25, 0.5, and 1 M, respectively. Then, given the SantaLucia parameter set [118], and T = 293 K, we find the binding probabilities for this NS overhang sequence (5'-*CGATCGA*-3) to be 98.7, 99.2, and 99.5% at [NaCl]=0.25, 0.5, and 1 M, respectively. We note that this mean-field picture will overestimate the binding probability: the actual connectivity (4 overhangs per NS) will lead to geometric constraints that will likely decrease the actual achieved binding probability.

A.4 Quantification of transcriptional output

A.4.1 Microplate reader

Transcription reactions were performed at 25 °C in standard transcription conditions: 40 mM Tris buffer (pH 8), 8 mM $MgCl_2$, 500 μM rATP/rUTP/rCTP, 1 mM rGTP, 1 U/μL SUPERase-In^{TM} RNAse inhibitor, and 20 μM DFHBI-1T. 50-100 mM NaCl was also added, depending on the experiment. For experiments with NS-liquid droplets, droplets were allowed to form in the absence of T7 RNAp, RNAse inhibitor, and rNTPs for >2 hr in a 384 micro-well plate (Corning #3544). T7 RNAp and RNase inhibitor were subsequently added without mixing and allowed to incubate for >30 min. Finally, rNTPs were added to all samples simultaneously using a multichannel pipette.

Fluorescent quantification of transcriptional output, as reported by the *Spinach*/DFHBI-

1T complex, was achieved using a Tecan Infinite 200 Pro plate reader with an excitation wavelength of 480 nm and emission wavelength of 513 nm, where bandwidth was 9 nm and 20 nm, respectively.

A.4.2 Polyacrylamide gel electrophoresis

RNA products from T7 RNAp transcription reactions with iSpinach dsDNA templates were examined using PAGE. Transcription reactions were frozen upon completion and thawed just before electrophoresis. Samples were heated at 75 °C for 15 min, cold shocked on ice for 15 min, then mixed with formamide to remove/minimize RNA secondary

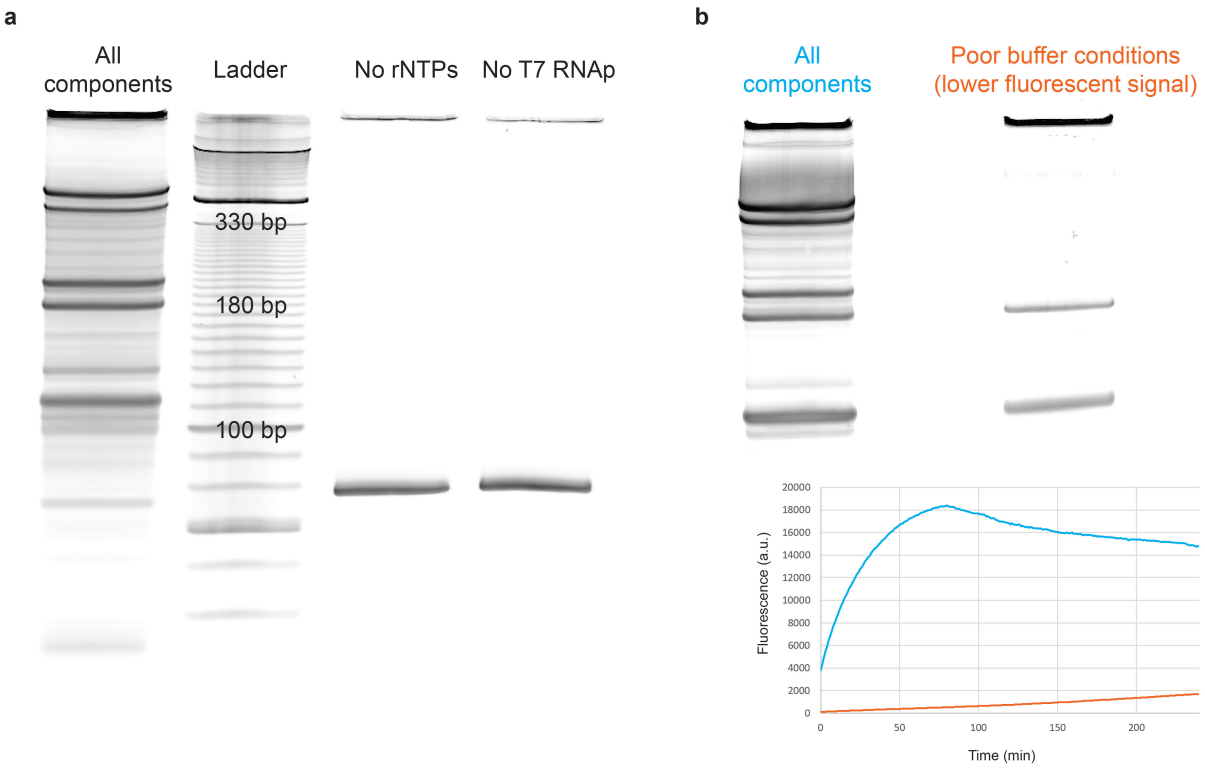


Figure A.8: **Denaturing PAGE gels of iSpinach transcription products. (a)** Comparison of RNA products when transcription reactions contained all components (left) to when a major component is removed (right). **(b)** Fewer RNA products (top) correspond to a lower iSpinach signal in fluorescent experiments (bottom).

structures. Samples were loaded into 8% polyacrylamide gels and run in denaturing conditions (7.5 M urea). DNA was stained using SYBR™ Gold and gels were imaged using an Amersham Typhoon 5 Biomolecular Imager.

As iSpinach contains three terminators, four dominant RNA bands were observed, corresponding to 119, 177, 364, and 419 nt for the three terminators and enzyme run-off, respectively (Fig. A.8). When a primary component of transcription is removed, none of the RNA products are observed. In conditions that hinder T7 RNAp transcription (e.g. high salt or depleted Mg^{2+}), fewer RNA products are observed, corresponding to a lower fluorescence signal in microplate reader experiments (Fig. A.8b).

Appendix B

DNA sequences

B.1 Nanostars

SsDNA were synthesized and desalted by Integrated DNA Technologies. Underlined bases refer to nucleotides that remain unpaired upon NS formation/hybridization and can be removed to generate the 4 different NS architectures. Colored bases highlight the differences between the families of oligos used in each chapter. iFluorT = internal fluorescein-modified thymine; iCy3 = internal Cy3 dye.

Oligo name	Sequence
Ch.2	3' overhang
v1	CTACTATGGCGGGTGATAAAAACGGGAAGAGCATGCCCATCCACGATCG
v1-Fluor	CTACTA|iFluorT|GGCGGG[continued as above]
v2	GGATGGGCATGCTCTTCCCGAACTCAACTGCCTGGTGATACGACGATCG
v3	CGTATCACCAGGCAGTTGAGAACATGCGAGGGTCCAATACCGACGATCG
s4	CGGTATTGGACCCTCGCATGAATTTATCACCCGCCATAGTAGACGATCG
Ch.3	5' overhang
v1	CGATCGACTACTATGGCGGGTGATAAAAACGGGAAGAGCATGCCCATCC
v1-Cy3	CGATCGA|iCy3|CTACTA[continued as above]
v2	CGATCGAGGATGGGCATGCTCTTCCCGAACTCAACTGCCTGGTGATACG
v3	CGATCGACGTATCACCAGGCAGTTGAGAACATGCGAGGGTCCAATACCG
s4	CGATCGACGGTATTGGACCCTCGCATGAATTTATCACCCGCCATAGTAG

B.2 dsDNA linkers

B.2.1 Primers

DsDNA linkers were generated using either standard PCR, when generating blunt dsDNA, or 'auto-sticky' PCR, when generating dsDNA template exhibiting 5' overhangs (see Appendix A.1.2). FP or ForP = Forward primer. RP or RevP = Reverse primer. idSp = internal abasic site. iCy3 = internal Cy3 dye.

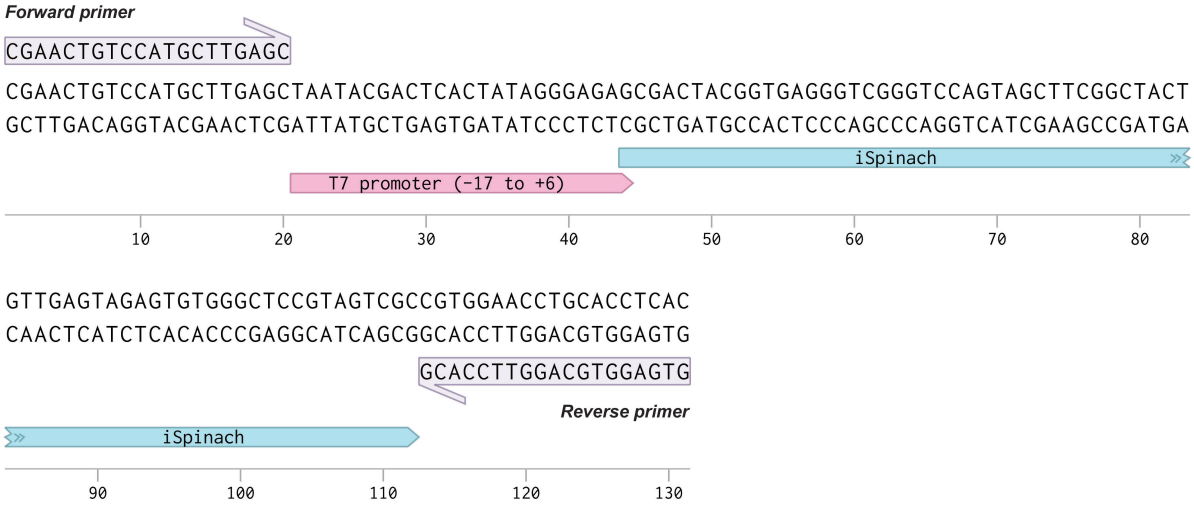
Oligo name	Sequence	PCR template
CGATCG-ForP-Sp2	CGATCG|idSp|TCATCACTAAAGAAATCCCGAGT	pDAN-Sp2
CGATCG-ForP-Sp2-iCy3	CGATCG|iCy3|TCATCACTAAAGAAATCCCGAGT	pDAN-Sp2
CGATCG-RevP-Sp2	CGATCG|idSp|ATGATTTCCTCGGGCAAAAA	pDAN-Sp2
GeneFP-iSp-OH6	CGATCG|idSp|CGAACTGTCCATGCTTGAGC	pDAN
GeneFP-iSp-OH6-Cy3	CGATCG|iCy3|CGAACTGTCCATGCTTGAGC	pDAN
GeneRP-iSp-OH6	CGATCG|idSp|GTGAGGTGCAGGTTCCACG	pDAN
T7-Ext-131-FP	CGAACTGTCCATGCTTGAGCTAATACGACTCACTATAGGGAG	iSp-short
T7-Ext-131-RP	GTGAGGTGCAGGTTCCACGGCGACTACGGAGCC	iSp-short
GalE-Ext-FP	CGAACTGTCCATGCTTGAGCCGGCGTAGAGGATCGAGATCTC	pET-21b-GalE
GalE-Ext270-RP	GTGAGGTGCAGGTTCCACGCTGCGCTTACTGTTACAGAGGTTATC	pET-21b-GalE
GalE-Ext350-RP	GTGAGGTGCAGGTTCCACGGCTTCGTTACGAATATCGCCTTCAAC	pET-21b-GalE

Extension PCR was utilized to produce the dsDNA fragments of 131 bp, 278 bp, and 350 bp. Extension primers were designed such that the PCR products contained regions allowing subsequent PCR with pDAN-specific primers. For the 131 bp product, a dsDNA PCR template was first generated by annealing and polymerizing oligos (in the table below) so as to produce the dsDNA PCR template iSp-short.

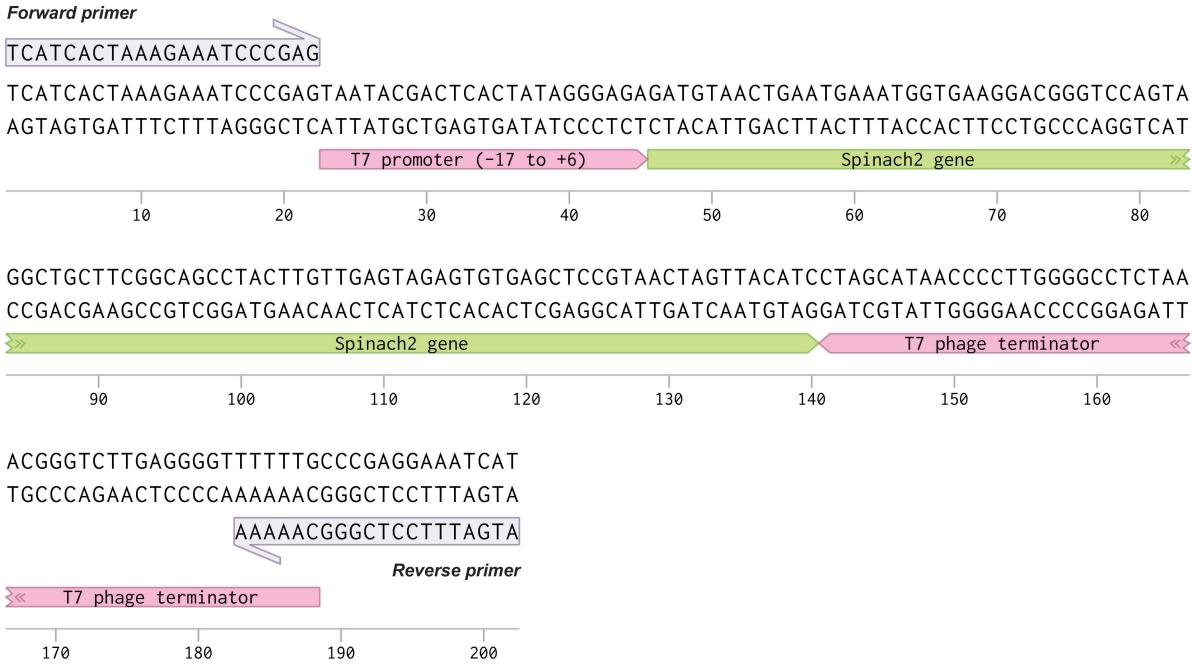
Oligo name	Sequence
T7-ext-Top	TAATACGACTCACTATAGGGAGAGCGACTACGGTGAGGGTCGGGTCCAGTAGCTTCGGCT
T7-ext-Bottom	GCGACTACGGAGCCCACACTCTACTCAACAGTAGCCGAAGCTACTGGACCCGACCCTCAC

B.2.2 Annotated dsDNA

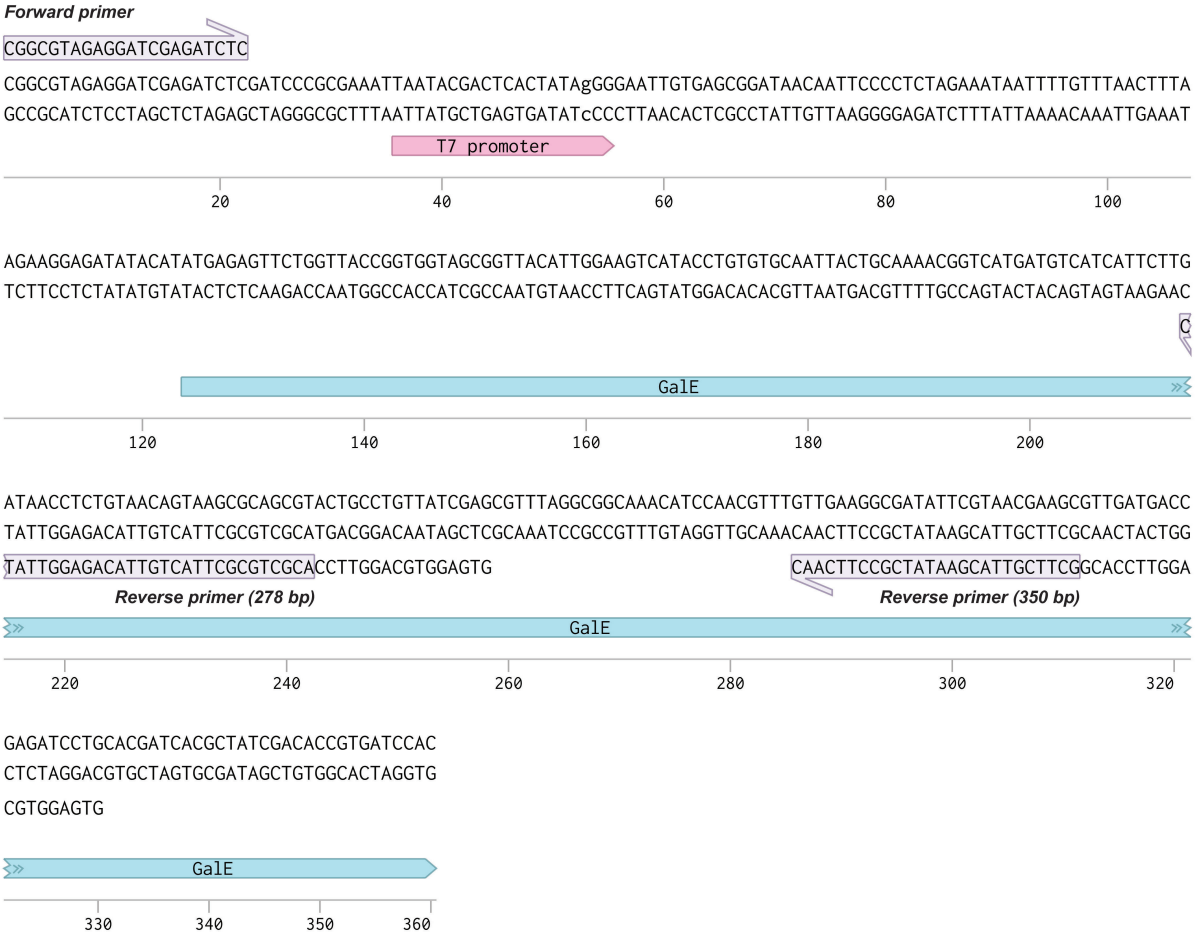
- iSpinach-No Terminators (131 bp)



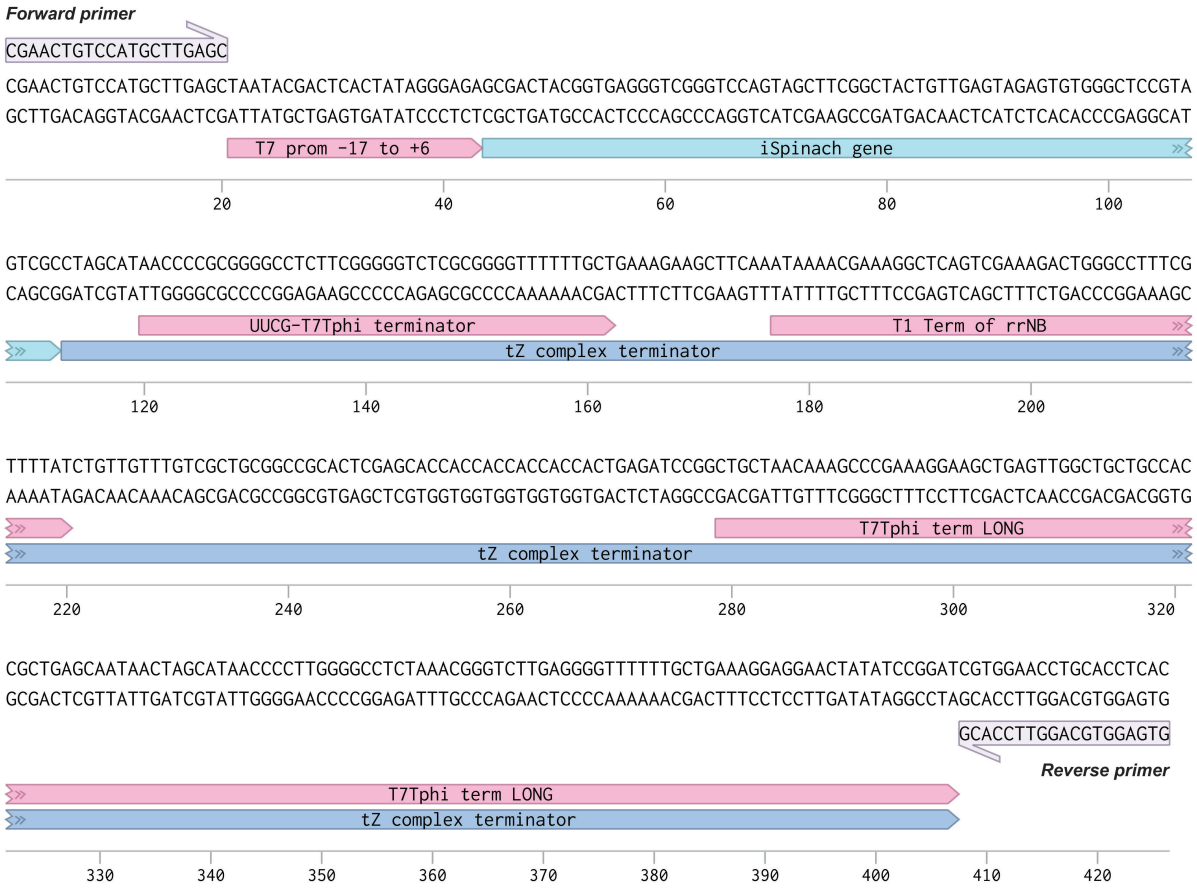
- Spinach2 (202 bp)



- GalE (278 and 350 bp)



- iSpinach (426 bp)



B.3 Plasmids

- **pCH762-pBSK-sfGFP**:

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTT
TCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGAT
AAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATT
TCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTT
TTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGT
TGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAG
ATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACT
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGG
CAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTT
GAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTA
AGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
AACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTT
TTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCG
GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCT
GTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTT
ACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAA
GTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATT
GCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGC
AGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC
GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTG
AGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTT
ACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAA
GGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTT

AACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA

AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGC

AAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAG

AGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGA

TACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCA

AGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTAC

CAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT

CAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGG

GGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTG

AGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGG

AGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGA

GCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCC

TGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTC

GTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTT

TACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTG

CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG

CTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGA

GCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGC

GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA

AAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT

AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG

GAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATG

ATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGG

AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGACC

TCGAGGGTGGCGCGGATCCGGCTGGTCTGTCTAAAGGTGAAGAACTG

TTCACCGGTGTTGTTCCGATCCTGGTTGAACTGGATGGTGATGTTAAC

GGCCACAAATTCTCTGTTCGTGGTGAAGGTGAAGGTGATGCAACCAAC

GGTAAACTGACCCTGAAATTCATCTGCACTACCGGTAAACTGCCGGTT

CCATGGCCGACTCTGGTGACTACCCTGACCTATGGTGTTCAGTGTTTT

TCTCGTTACCCGGATCACATGAAGCAGCATGATTTCTTCAAATCTGCA

ATGCCGGAAGGTTATGTACAGGAGCGCACCATTTCTTTCAAAGACGAT

GGCACCTACAAAACCCGTGCAGAGGTTAAATTTGAAGGTGATACTCTG

GTGAACCGTATTGAACTGAAAGGCATTGATTTCAAAGAGGACGGCAA

CATCCTGGGCCACAAACTGGAATATAACTTCAACTCCCATAACGTTTA

CATCACCGCAGACAAACAGAAGAACGGTATCAAAGCTAACTTCAAAAT

TCGCCATAACGTTGAAGACGGTAGCGTACAGCTGGCGGACCACTACCA

GCAGAACACTCCGATCGGTGATGGTCCGGTTCTGCTGCCGGATAACCA

CTACCTGTCCACCCAGTCTGTTCTGTCCAAAGACCCGAACGAAAAGCG

CGACCACATGGTGCTGCTGGAGTTCGTTACTGCAGCAGGTATCACGCA

CGGCATGGATGAGCTCTACAAATAAGTCGACAAGCTTGTCGGGCTGCA

GGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC

CGGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGG

CCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC

TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCG

AAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG

GCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTG

GTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCC

GCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTC

CCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTG

CTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCAC

GTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGG

AGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACAC

TCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGA
TTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACG
CGAATTTTAACAAAATATTAACGCTTACAATTTAG

- **pCH3943-pT7-911Q**:

GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC
GGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT
CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT
TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA
TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTT
GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGA
TCTGTCTATTTCGTTCATCCATAGCTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGA
TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC
AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCC

GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGT

TCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC

GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCC

CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCG

GTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCA

GTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC

ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG

TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCG

TCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTC

ATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCG

CTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT

TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA

AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTG

AATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG

TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAA

ACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGT

CTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTAT

CACGAGGCCCTTTCGTCTTCACCTCGAGAAATCATAAAAAATTTATTT

GCTTTGTGAGCGGATAACAATTATAATAGATTCAATTGTGAGCGGATA

ACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGAGG

ATCGCATCACCATCACCATCACGGATCCATGAACACGATTAACATCGC

TAAGAACGACTTCTCTGACATCGAACTGGCTGCTATCCCGTTCAACAC

TCTGGCTGACCATTACGGTGAGCGTTTAGCTCGCGAACAGTTGGCCCT

TGAGCATGAGTCTTACGAGATGGGTGAAGCACGCTTCCGCAAGATGTT

TGAGCGTCAACTTAAAGCTGGTGAGGTTGCGGATAACGCTGCCGCCAA

GCCTCTCATCACTACCCTACTCCCTAAGATGATTGCACGCATCAACGA

CTGGTTTGAGGAAGTGAAAGCTAAGCGCGGCAAGCGCCCGACAGCCT
TCCAGTTCCTGCAAGAAATCAAGCCGGAAGCCGTAGCGTACATCACCA
TTAAGACCACTCTGGCTTGCCTAACCAGTGCTGACAATACAACCGTTC
AGGCTGTAGCAAGCGCAATCGGTCGGGCCATTGAGGACGAGGCTCGC
TTCGGTCGTATCCGTGACCTTGAAGCTAAGCACTTCAAGAAAAACGTT
GAGGAACAACTCAACAAGCGCGTAGGGCACGTCTACAAGAAAGCATTT
ATGCAAGTTGTCGAGGCTGACATGCTCTCTAAGGGTCTACTCGGTGGC
GAGGCGTGGTCTTCGTGGCATAAGGAAGACTCTATTCATGTAGGAGT
ACGCTGCATCGAGATGCTCATTGAGTCAACCGGAATGGTTAGCTTACA
CCGCCAAAATGCTGGCGTAGTAGGTCAAGACTCTGAGACTATCGAACT
CGCACCTGAATACGCTGAGGCTATCGCAACCCGTGCAGGTGCGCTGGC
TGGCATCTCTCCGATGTTCCAACCTTGCGTAGTTCCTCCTAAGCCGTG
GACTGGCATTACTGGTGGTGGCTATTGGGCTAACGGTCGTCGTCCTCT
GGCGCTGGTGCGTACTCACAGTAAGAAAGCACTGATGCGCTACGAAG
ACGTTTACATGCCTGAGGTGTACAAAGCGATTAACATTGCGCAAAACA
CCGCATGGAAAATCAACAAGAAAGTCCTAGCGGTCGCCAACGTAATCA
CCAAGTGGAAGCATTGTCCGGTCGAGGACATCCCTGCGATTGAGCGTG
AAGAACTCCCGATGAAACCGGAAGACATCGACATGAATCCTGAGGCTC
TCACCGCGTGGAAACGTGCTGCCGCTGCTGTGTACCGCAAGGACAAGG
CTCGCAAGTCTCGCCGTATCAGCCTTGAGTTCATGCTTGAGCAAGCCA
ATAAGTTTGCTAACCATAAGGCCATCTGGTTCCCTTACAACATGGACT
GGCGCGGTCGTGTTTACGCTGTGTCAATGTTCAACCCGCAAGGTAACG
ATATGACCAAAGGACTGCTTACGCTGGCGAAAGGTAAACCAATCGGTA
AGGAAGGTTACTACTGGCTGAAAATCCACGGTGCAAACTGTGCGGGT
GTCGATAAGGTTCCGTTCCCTGAGCGCATCAAGTTCATTGAGGAAAAC
CACGAGAACATCATGGCTTGCGCTAAGTCTCCACTGGAGAACACTTGG

TGGGCTGAGCAAGATTCTCCGTTCTGCTTCCTTGCGTTCTGCTTTGAG
TACGCTGGGGTACAGCACCACGGCCTGAGCTATAACTGCTCCCTTCCG
CTGGCGTTTGACGGGTCTTGCTCTGGCATCCAGCACTTCTCCGCGATG
CTCCGAGATGAGGTAGGTGGTCGCGCGGTTAACTTGCTTCCTAGTGAA
ACCGTTCAGGACATCTACGGGATTGTTGCTAAGAAAGTCAACGAGATT
CTACAAGCAGACGCAATCAATGGGACCGATAACGAAGTAGTTACCGTG
ACCGATGAGAACACTGGTGAAATCTCTGAGAAAGTCAAGCTGGGCAC
TAAGGCACTGGCTGGTCAATGGCTGGCTTACGGTGTTACTCGCAGTGT
GACTAAGCGTTCAGTCATGACGCTGGCTTACGGGTCCAAAGAGTTCGG
CTTCCGTCAACAAGTGCTGGAAGATACCATTCAGCCAGCTATTGATTC
CGGCAAGGGTCTGATGTTCACTCAGCCGAATCAGGCTGCTGGATACAT
GGCTAAGCTGATTTGGGAATCTGTGAGCGTGACGGTGGTAGCTGCGG
TTGAAGCAATGAACTGGCTTAAGTCTGCTGCTAAGCTGCTGGCTGCTG
AGGTCAAAGATAAGAAGACTGGAGAGATTCTTCGCAAGCGTTGCGCT
GTGCATTGGGTAACTCCTGATGGTTTCCCTGTGTGGCAGGAATACAAG
AAGCCTATTCAGACGCGCTTGAACCTGATGTTCCTCGGTCAGTTCCGC
TTACAGCCTACCATTAACACCAACAAAGATAGCGAGATTGATGCACAC
AAACAGGAGTCTGGTATCGCTCCTAACTTTGTACACAGCCAAGACGGT
AGCCACCTTCGTAAGACTGTAGTGTGGGCACACGAGAAGTACGGAATC
GAATCTTTTGCACTGATTCACGACTCCTTCGGTACCATTCCGGCTGAC
GCTGCGAACCTGTTCAAAGCAGTGCGCGAAACTATGGTTGACACATAT
GAGTCTTGTGATGTACTGGCTGATTTCTACGACCAGTTCGCTGACCAG
TTGCACGAGTCTCAATTGGACAAAATGCCAGCACTTCCGGCTAAAGGT
AACTTGAACCTCCGTGACATCTTAGAGTCGGACTTCGCGTTCGCGTAA
CGCCGGATCCTAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGA
TCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGG

TTGCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGA
GATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGA
TATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAG
GCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTG
GATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTT
TATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCG
GAATTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAG
TGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTC
ATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACAT
ATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCC
TAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGT
GAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTT
CGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGT
GCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTGATGGCTT
CCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTG
GCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAAC
GCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGG
CTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGA
ACGCTCTCCTGAGTAGGACAAATCCGCCGCTCTAGAGCTGCCTCGCGC
GTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGA
CGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTC
AGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACC
CAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCAT
CAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGC
ACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCT
CGCTCACTGACTCGCTGCGCTCGGTCTGTCGGCTGCGGCGAGCGGTAT

CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAA

- **pCH14252-pT7-911Q-sfGFP (*T7 RNAp-sfGFP gene only*)**:

...ATGAACACGATTAACATCGCTAAGAACGACTTCTCTGACATCGAACT
GGCTGCTATCCCGTTCAACACTCTGGCTGACCATTACGGTGAGCGTTT
AGCTCGCGAACAGTTGGCCCTTGAGCATGAGTCTTACGAGATGGGTG
AAGCACGCTTCCGCAAGATGTTTGAGCGTCAACTTAAAGCTGGTGAGG
TTGCGGATAACGCTGCCGCCAAGCCTCTCATCACTACCCTACTCCCTA
AGATGATTGCACGCATCAACGACTGGTTTGAGGAAGTGAAAGCTAAG
CGCGGCAAGCGCCCGACAGCCTTCCAGTTCCTGCAAGAAATCAAGCCG
GAAGCCGTAGCGTACATCACCATTAAGACCACTCTGGCTTGCCTAACC
AGTGCTGACAATACAACCGTTCAGGCTGTAGCAAGCGCAATCGGTCGG
GCCATTGAGGACGAGGCTCGCTTCGGTCGTATCCGTGACCTTGAAGCT
AAGCACTTCAAGAAAAACGTTGAGGAACAACTCAACAAGCGCGTAGG
GCACGTCTACAAGAAAGCATTTATGCAAGTTGTCGAGGCTGACATGCT
CTCTAAGGGTCTACTCGGTGGCGAGGCGTGGTCTTCGTGGCATAAGG
AAGACTCTATTCATGTAGGAGTACGCTGCATCGAGATGCTCATTGAGT
CAACCGGAATGGTTAGCTTACACCGCCAAAATGCTGGCGTAGTAGGTC
AAGACTCTGAGACTATCGAACTCGCACCTGAATACGCTGAGGCTATCG
CAACCCGTGCAGGTGCGCTGGCTGGCATCTCTCCGATGTTCCAACCTT
GCGTAGTTCCTCCTAAGCCGTGGACTGGCATTACTGGTGGTGGCTATT
GGGCTAACGGTCGTCGTCCTCTGGCGCTGGTGCGTACTCACAGTAAGA
AAGCACTGATGCGCTACGAAGACGTTTACATGCCTGAGGTGTACAAAG
CGATTAACATTGCGCAAAACACCGCATGGAAAATCAACAAGAAAGTCC

TAGCGGTCGCCAACGTAATCACCAAGTGGAAGCATTGTCCGGTCGAGG

ACATCCCTGCGATTGAGCGTGAAGAACTCCCGATGAAACCGGAAGACA

TCGACATGAATCCTGAGGCTCTCACCGCGTGGAAACGTGCTGCCGCTG

CTGTGTACCGCAAGGACAAGGCTCGCAAGTCTCGCCGTATCAGCCTTG

AGTTCATGCTTGAGCAAGCCAATAAGTTTGCTAACCATAAGGCCATCT

GGTTCCCTTACAACATGGACTGGCGCGGTCGTGTTTACGCTGTGTCAA

TGTTCAACCCGCAAGGTAACGATATGACCAAAGGACTGCTTACGCTGG

CGAAAGGTAAACCAATCGGTAAGGAAGGTTACTACTGGCTGAAAATCC

ACGGTGCAAACTGTGCGGGTGTCGATAAGGTTCCGTTCCCTGAGCGCA

TCAAGTTCATTGAGGAAAACCACGAGAACATCATGGCTTGCGCTAAGT

CTCCACTGGAGAACACTTGGTGGGCTGAGCAAGATTCTCCGTTCTGCT

TCCTTGCGTTCTGCTTTGAGTACGCTGGGGTACAGCACCACGGCCTGA

GCTATAACTGCTCCCTTCCGCTGGCGTTTGACGGGTCTTGCTCTGGCA

TCCAGCACTTCTCCGCGATGCTCCGAGATGAGGTAGGTGGTCGCGCGG

TTAACTTGCTTCCTAGTGAAACCGTTCAGGACATCTACGGGATTGTTG

CTAAGAAAGTCAACGAGATTCTACAAGCAGACGCAATCAATGGGACCG

ATAACGAAGTAGTTACCGTGACCGATGAGAACACTGGTGAAATCTCTG

AGAAAGTCAAGCTGGGCACTAAGGCACTGGCTGGTCAATGGCTGGCT

TACGGTGTTACTCGCAGTGTGACTAAGCGTTCAGTCATGACGCTGGCT

TACGGGTCCAAAGAGTTCGGCTTCCGTCAACAAGTGCTGGAAGATACC

ATTCAGCCAGCTATTGATTCCGGCAAGGGTCTGATGTTCACTCAGCCG

AATCAGGCTGCTGGATACATGGCTAAGCTGATTTGGGAATCTGTGAGC

GTGACGGTGGTAGCTGCGGTTGAAGCAATGAACTGGCTTAAGTCTGC

TGCTAAGCTGCTGGCTGCTGAGGTCAAAGATAAGAAGACTGGAGAGA

TTCTTCGCAAGCGTTGCGCTGTGCATTGGGTAACTCCTGATGGTTTCC

CTGTGTGGCAGGAATACAAGAAGCCTATTCAGACGCGCTTGAACCTGA

TGTTCCTCGGTCAGTTCCGCTTACAGCCTACCATTAACACCAACAAAG
ATAGCGAGATTGATGCACACAAACAGGAGTCTGGTATCGCTCCTAACT
TTGTACACAGCCAAGACGGTAGCCACCTTCGTAAGACTGTAGTGTGGG
CACACGAGAAGTACGGAATCGAATCTTTTGCACTGATTCACGACTCCT
TCGGTACCATTCCGGCTGACGCTGCGAACCTGTTCAAAGCAGTGCGCG
AAACTATGGTTGACACATATGAGTCTTGTGATGTACTGGCTGATTTCT
ACGACCAGTTCGCTGACCAGTTGCACGAGTCTCAATTGGACAAAATGC
CAGCACTTCCGGCTAAAGGTAACTTGAACCTCCGTGACATCTTAGAGT
CGGACTTCGCGTTCGCGACTAGTGGCGGGTCCGGGGGCTCTAAAGGT
GAAGAACTGTTCACCGGTGTTGTTCCGATCCTGGTTGAACTGGATGGT
GATGTTAACGGCCACAAATTCTCTGTTCGTGGTGAAGGTGAAGGTGAT
GCAACCAACGGTAAACTGACCCTGAAATTCATCTGCACTACCGGTAAA
CTGCCGGTTCCATGGCCGACTCTGGTGACTACCCTGACCTATGGTGTT
CAGTGTTTTTCTCGTTACCCGGATCACATGAAGCAGCATGATTTCTTC
AAATCTGCAATGCCGGAAGGTTATGTACAGGAGCGCACCATTTCTTTC
AAAGACGATGGCACCTACAAAACCCGTGCAGAGGTTAAATTTGAAGGT
GATACTCTGGTGAACCGTATTGAACTGAAAGGCATTGATTTCAAAGAG
GACGGCAACATCCTGGGCCACAAACTGGAATATAACTTCAACTCCCAT
AACGTTTACATCACCGCAGACAAACAGAAGAACGGTATCAAAGCTAAC
TTCAAAATTCGCCATAACGTTGAAGACGGTAGCGTACAGCTGGCGGAC
CACTACCAGCAGAACACTCCGATCGGTGATGGTCCGGTTCTGCTGCCG
GATAACCACTACCTGTCCACCCAGTCTGTTCTGTCCAAAGACCCGAAC
GAAAAGCGCGACCACATGGTGCTGCTGGAGTTCGTTACTGCAGCAGG
TATCACGCACGGCATGGATGAGCTCTACAAATAA...

- **pDAN**:

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCC
CGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAA
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCT
TAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCAAATGCG
GTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGC
GCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
CGGGCCTCATCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGC
AAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG
TAAAACGACGGCCAGTGCAACGCGATGACGATGGATAGCGATTCATC
GATGAGCTGACCCGATCGCCGCCGCCGGAGGGTTGCGTTTGAGACGG
GCGACAGATCGAACTGTCCATGCTTGAGCTAATACGACTCACTATAGG
GAGAGCGACTACGGTGAGGGTCGGGTCCAGTAGCTTCGGCTACTGTT
GAGTAGAGTGTGGGCTCCGTAGTCGCCTAGCATAACCCCGCGGGGCC
TCTTCGGGGGTCTCGCGGGGTTTTTTGCTGAAAGAAGCTTCAAATAAA
ACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTT
GTCGCTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCG
GCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCT
GAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGG
GGTTTTTTGCTGAAAGGAGGAACTATATCCGGATCGTGGAACCTGCAC
CTCACATCAGTTCTGGACCAGCGAGCTGTGCTGCGACTCGTGGCGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT
TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAA

CGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCA
GCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC
GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGA
CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC
AGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG
CTCTCCTGTTCCGACCCTGTCGCTTACCGGATACCTGTCCGCCTTTCT
CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT
CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACC
CCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG
TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT
GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTT
TGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCC
TTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGT
AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCT
CAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCG
TGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAA
TAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT
TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA

AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACA
GGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCC
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAA
AAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTG
GCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTT
ACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCA
ACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGC
CCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAC
TGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAA
ACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA
ATGTTGAATACTCATACTCTACCTTTTTCAATATTATTGAAGCATTTA
TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA
AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC
TGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAG
GCGTATCACGAGGCCCTTTCGTC

- **pDAN-Sp2**:

CCCGTGTAAAACGACGGCCAGTTTATCTAGTCAGCTTGATTCTAGCTG
ATCGTGGACCGGAAGGTGAGCCAGTGAGTTGATTGCAGTCCAGTTAC
GCTGGAGTCTGAGGCTCGTCCTGAATGATATGCGGCCTCTCATCACTA
AAGAAATCCCGAGTAATACGACTCACTATAGGGAGAGATGTAACTGAA
TGAAATGGTGAAGGACGGGTCCAGTAGGCTGCTTCGGCAGCCTACTT
GTTGAGTAGAGTGTGAGCTCCGTAACTAGTTACATCCTAGCATAACCC
CTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCCCGAGGAAATC

ATGCGCGTGATCTTACGGCATTATACGTATGATCGGTCCACGATCAGC
TAGATTATCTAGTCAGCTTGATGTCATAGCTGTTTCCTGAGGCTCAAT
ACTGACCATTTAAATCATACCTGACCTCCATAGCAGAAAGTCAAAAGC
CTCCGACCGGAGGCTTTTGACTTGATCGGCACGTAAGAGGTTCCAACT
TTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTTAT
CGAGATTTTCAGGAGCTAAGGAAGCTAAAATGAGTATTCAACATTTCC
GTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTG
CTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTG
GGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT
CCTTGAGAGTTTACGCCCCGAAGAACGTTTTCCAATGATGAGCACTTT
TAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGA
GTACTCACCAGTCACAGAAAAGCATCTCACGGATGGCATGACAGTAAG
AGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAA
CTTACTTCTGGCAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTT
GCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGA
GCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGT
AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTAC
TCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGT
TGCAGGATCACTTCTGCGCTCGGCCCTCCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAG
CACTGGGGCCAGATGGTAAGCCCTCCCGCATCGTAGTTATCTACACGA
CGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAG
ATAGGTGCCTCACTGATTAAGCATTGGTAATGAGGGCCCAAATGTAAT
CACCTGGCTCACCTTCGGGTGGGCCTTTCTTGAGGACCTAAATGTAAT
CACCTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTGCTGGCGTTTTTC

CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGATGCTCAAGT
CAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
CCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC
GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT
AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG
CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCG
CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC
TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT
AGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATTTTCTACCGAAGAAAGGCCC

- **pET-21b-GalE**:

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTG
TGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC
CCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTT
TCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAG
TGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTC
ACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTT
GGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAAC
ACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCC
GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAA
CGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTT
TCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACA

TTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCA
ATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGC
CCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCA
GAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACG
AGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAG
TTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCT
GCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACT
CGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACC
AGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG
CAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCT
GACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACAT
GGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATG
AAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGG
CAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTT
CCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGA
CCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAA
TCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGG
GCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAG
TCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG
CCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATA
TACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGG
TGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGT
TTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA
AACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAA
CTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATA

CTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTG
TAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTG
CTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT
AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGC
ACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTA
CAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGC
GGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGA
GGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGG
TTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGG
GGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTC
CTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCC
CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACC
GCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGA
AGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGG
TATTTCACACCGCATATATGGTGCACTCTCAGTACAATCTGCTCTGAT
GCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGG
TCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGAC
GGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCT
CCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG
CGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCA
CAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGA
AGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTT
TTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTC
ATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACG
GGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTA
AACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAG

GGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGT
AGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGG
CGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGAC
CATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCT
TCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCA
ACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCG
CACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAA
ACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCA
AGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGC
GAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTC
CTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATA
GTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTC
AAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT
GCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCA
ACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGC
GGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGG
TTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCA
CTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGC
GCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGG
GAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACA
TGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGC
GAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACA
GAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCG
ACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATA

ATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGG
AACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTG
CACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACAC
CACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGAC
AATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAA
TCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAA
TGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCG
CAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAA
GAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACA
TTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCG
CGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCC
CTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCC
GTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGC
CCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAAC
AAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGA
TGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATG
CCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGA
AATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCC
CTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAG
AGTTCTGGTTACCGGTGGTAGCGGTTACATTGGAAGTCATACCTGTGT
GCAATTACTGCAAAACGGTCATGATGTCATCATTCTTGATAACCTCTG
TAACAGTAAGCGCAGCGTACTGCCTGTTATCGAGCGTTTAGGCGGCAA
ACATCCAACGTTTGTTGAAGGCGATATTCGTAACGAAGCGTTGATGAC
CGAGATCCTGCACGATCACGCTATCGACACCGTGATCCACTTCGCCGG
GCTGAAAGCCGTGGGCGAATCGGTACAAAAACCGCTGGAATATTACG

ACAACAATGTCAACGGCACTCTGCGCCTGATTAGCGCCATGCGCGCCG
CTAACGTCAAAAACTTTATTTTTAGCTCCTCCGCCACCGTTTATGGCG
ATCAGCCCAAAATTCCATACGTTGAAAGCTTCCCGACCGGCACACCGC
AAAGCCCTTACGGCAAAAGCAAGCTGATGGTGGAACAGATCCTCACCG
ATCTGCAAAAAGCCCAGCCGGACTGGAGCATTGCCCTGCTGCGCTACT
TCAACCCGGTTGGCGCGCATCCGTCGGGCGATATGGGCGAAGATCCG
CAAGGCATTCCGAATAACCTGATGCCATACATCGCCCAGGTTGCTGTA
GGCCGTCGCGACTCGCTGGCGATTTTTGGTAACGATTATCCGACCGAA
GATGGTACTGGCGTACGCGATTACATCCACGTAATGGATCTGGCGGAC
GGTCACGTCGTGGCGATGGAAAAACTGGCGAACAAGCCAGGCGTACA
CATCTACAACCTCGGCGCTGGCGTAGGCAACAGCGTGCTGGACGTGGT
TAATGCCTTCAGCAAAGCCTGCGGCAAACCGGTTAATTATCATTTTGC
ACCGCGTCGCGAGGGCGACCTTCCGGCCTACTGGGCGGACGCCAGCA
AAGCCGACCGTGAACTGAACTGGCGCGTAACGCGCACACTCGATGAAA
TGGCGCAGGACACCTGGCACTGGCAGTCACGCCATCCACAGGGATATC
CCGATTAAGGAACGACCATGACGCAATTTAATCCCGTTGATCATCCAC
ATCGCCGCTACAACCCGCTCACCGGGCAATGGATTCTGGTTTCACCGC
ACCGCGCTAAGCGCCCCTGGCAGGGGGCGCAGGAAACGCCAGCCAAA
CAGGTGTTACCTGCGCACTCGAGCACCACCACCACCACCACTGAGATC
CGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA
GGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGAT

Appendix C

Abbreviations

bp	Base pair(s)
CFGE	Cell-free gene expression
CID	Chromosomal interaction domain
CT	Chromosome territory
DFHBI	3,5-difluoro-4-hydroxybenzylidene imidazolinone
DFHBI-1T	(Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1H-imidazol-5(4H)-one
DNA	Deoxyribonucleic acid
DNAp	DNA polymerase
dsDNA	Double-stranded DNA
FITC	Fluorescein isothiocyanate
FRAP	Fluorescence recovery after photobleaching
GFP	Green fluorescent protein
kb	Kilobase
LLPS	Liquid-liquid phase separation
lncRNA	Long non-coding RNA

mRNA	Messenger RNA
miRNA	MicroRNA
MSD	Mean squared displacement
NAP	Nucleoid-associated protein
ncRNA	Non-coding RNA
NS	Nanostar
nt	Nucleotide(s)
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
rasiRNA	Repeat-associated short interfering RNA
RNA	Ribonucleic acid
RNAp	RNA polymerase
roi	Region of interest
SD	Supercoil domain
sfGFP	Superfolder green fluorescent protein
siRNA	Short interfering RNA
ssDNA	Single-stranded DNA
TAD	Topologically-associated domain
TF	Transcription factor
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Bibliography

[1] F. H. C. Crick, *On protein synthesis*, in *Symp Soc Exp Biol*, vol. 12, pp. 138–163, 1958.

[2] F. H. C. Crick, *Central dogma of molecular biology*, *Nature* **227** (1970), no. 5258 561–563.

[3] S. Prabakaran, G. Lippens, H. Steen, and J. Gunawardena, *Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding*, *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **4** (2012), no. 6 565–583.

[4] R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, *et. al.*, *How many human proteoforms are there?*, *Nature Chemical Biology* **14** (2018), no. 3 206–214.

[5] D. Baltimore, *Viral rna-dependent dna polymerase: Rna-dependent dna polymerase in virions of rna tumour viruses*, *Nature* **226** (1970), no. 5252 1209–1211.

[6] H. M. Temin and S. Mizutani, *Viral rna-dependent dna polymerase: Rna-dependent dna polymerase in virions of rous sarcoma virus*, *Nature* **226** (1970), no. 5252 1211–1213.

[7] W. Gilbert, *Origin of life: The rna world*, *Nature* **319** (1986), no. 6055 618.

[8] T. R. Cech, *A model for the rna-catalyzed replication of rna*, *Proceedings of the National Academy of Sciences* **83** (1986), no. 12 4360–4363.

[9] M. P. Robertson and G. F. Joyce, *Highly efficient self-replicating rna enzymes*, *Chemistry & Biology* **21** (2014), no. 2 238–245.

[10] D. P. Horning and G. F. Joyce, *Amplification of rna by an rna polymerase ribozyme*, *Proceedings of the National Academy of Sciences* **113** (2016), no. 35 9786–9791.

[11] J. Attwater, A. Raguram, A. S. Morgunov, E. Gianni, and P. Holliger, *Ribozyme-catalysed rna synthesis using triplet building blocks*, *eLife* **7** (2018) e35255.

[12] S. B. Prusiner, *Prions*, *Proceedings of the National Academy of Sciences* **95** (1998), no. 23 13363–13383.

[13] A. Šarić, A. K. Buell, G. Meisl, T. C. Michaels, C. M. Dobson, S. Linse, T. P. Knowles, and D. Frenkel, *Physical determinants of the self-replication of protein fibrils*, *Nature Physics* **12** (2016), no. 9 874–880.

[14] J. Harrow, A. Frankish, J. M. Gonzalez, E. Tapanari, M. Diekhans, F. Kokocinski, B. L. Aken, D. Barrell, A. Zadissa, S. Searle, *et. al.*, *Gencode: the reference human genome annotation for the encode project*, *Genome Research* **22** (2012), no. 9 1760–1774.

[15] M. Pertea, A. Shumate, G. Pertea, A. Varabyou, Y.-C. Chang, A. K. Madugundu, A. Pandey, and S. Salzberg, *Thousands of large-scale rna sequencing experiments yield a comprehensive new human gene list and reveal extensive transcriptional noise*, *bioRxiv* (2018) [https://www.biorxiv.org/content/early/2018/05/29/332825.full.pdf].

[16] H. Siomi and M. C. Siomi, *On the road to reading the rna-interference code*, *Nature* **457** (2009), no. 7228 396–404.

[17] D. Moazed, *Small rnas in transcriptional gene silencing and genome defence*, *Nature* **457** (2009), no. 7228 413–420.

[18] D. Holoch and D. Moazed, *Rna-mediated epigenetic regulation of gene expression*, *Nature Reviews Genetics* **16** (2015), no. 2 71–84.

[19] C. H. Waddington, *The epigenotype*, *Endeavour* **1** (1942) 18–20.

[20] C. Deans and K. A. Maggert, *What do you mean,epigenetic?*, *Genetics* **199** (2015), no. 4 887–896.

[21] A. Annunziato, *Dna packaging: nucleosomes and chromatin*, *Nature Education* **1** (2008), no. 1 26.

[22] S. John, P. J. Sabo, R. E. Thurman, M.-H. Sung, S. C. Biddie, T. A. Johnson, G. L. Hager, and J. A. Stamatoyannopoulos, *Chromatin accessibility pre-determines glucocorticoid receptor binding patterns.*, *Nature Genetics* **43** (2011), no. 3 264–268.

[23] K. L. Jost, B. Bertulat, and M. C. Cardoso, *Heterochromatin and gene positioning: inside, outside, any side?*, *Chromosoma* **121** (2012), no. 6 555–563.

[24] O. Cuvier and B. Fierz, *Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells*, *Nature Reviews Genetics* **18** (2017), no. 8 457–472.

[25] J. R. Dixon, D. U. Gorkin, and B. Ren, *Chromatin domains: the unit of chromosome organization*, *Molecular cell* **62** (2016), no. 5 668–680.

[26] S. Sati and G. Cavalli, *Chromosome conformation capture technologies and their impact in understanding genome function*, *Chromosoma* **126** (2017), no. 1 33–44.

[27] N. Harmston, E. Ing-Simmons, G. Tan, M. Perry, M. Merkenschlager, and B. Lenhard, *Topologically associating domains are ancient features that coincide with metazoan clusters of extreme noncoding conservation*, *Nature Communications* **8** (2017), no. 441 1–13.

[28] J. Dekker, K. Rippe, M. Dekker, and N. Kleckner, *Capturing chromosome conformation*, *Science* **295** (2002), no. 5558 1306–1311.

[29] T. Sexton and G. Cavalli, *The role of chromosome domains in shaping the functional genome*, *Cell* **160** (2015), no. 6 1049–1059.

[30] E. Heitz, *Das heterochromatin der moose*, *Jahrb Wiss Botanik* **69** (1928) 762–818.

[31] B. Li, M. Carey, and J. L. Workman, *The role of chromatin during transcription.*, *Cell* **128** (2007), no. 4 707–719.

[32] C. D. Allis and T. Jenuwein, *The molecular hallmarks of epigenetic control*, *Nature Reviews Genetics* **17** (2016), no. 8 487–500.

[33] R. E. Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, *et. al.*, *Integrative analysis of 111 reference human epigenomes*, *Nature* **518** (2015), no. 7539 317–330.

[34] B. D. Strahl and C. D. Allis, *The language of covalent histone modifications*, *Nature* **403** (2000), no. 6765 41–45.

[35] T. Jenuwein and C. D. Allis, *Translating the histone code*, *Science* **293** (2001), no. 5532 1074–1080.

[36] E. Guccione, F. Martinato, G. Finocchiaro, L. Luzi, L. Tizzoni, V. Dall' Olio, G. Zardo, C. Nervi, L. Bernard, and B. Amati, *Myc-binding-site recognition in the human genome is determined by chromatin context.*, *Nature Cell Biology* **8** (2006), no. 7 764–770.

[37] W. Akhtar, J. de Jong, A. V. Pindyurin, L. Pagie, W. Meuleman, J. de Ridder, A. Berns, L. F. Wessels, M. van Lohuizen, and B. van Steensel, *Chromatin position effects assayed by thousands of reporters integrated in parallel*, *Cell* **154** (2013), no. 4 914–927.

[38] H. J. Muller, *Types of visible variations induced by x-rays indrosophila*, *Journal of Genetics* **22** (1930), no. 3 299–334.

[39] J. R. Girton and K. M. Johansen, *Chromatin structure and the regulation of gene expression: the lessons of pev in drosophila*, *Advances in Genetics* **61** (2008) 1–43.

[40] S. C. Elgin and G. Reuter, *Position-effect variegation, heterochromatin formation, and gene silencing in drosophila*, *Cold Spring Harbor Perspectives in Biology* **5** (2013), no. 8 a017780.

[41] S. C. Dillon and C. J. Dorman, *Bacterial nucleoid-associated proteins, nucleoid structure and gene expression.*, *Nature Reviews Microbiology* **8** (2010), no. 3 185–195.

[42] C. J. Dorman, *Genome architecture and global gene regulation in bacteria: making progress towards a unified model?*, *Nature Reviews Microbiology* **11** (2013), no. 5 349–355.

[43] A. Badrinarayanan, T. B. Le, and M. T. Laub, *Bacterial chromosome organization and segregation*, *Annual Review of Cell and Developmental Biology* **31** (November, 2015) 171–199.

[44] J. Stavans and A. Oppenheim, *DNA-protein interactions and bacterial chromosome architecture.*, *Physical Biology* **3** (2006), no. 4 R1–R10.

[45] M. Valens, S. Penaud, M. Rossignol, F. Cornet, and F. Boccard, *Macrodomain organization of the Escherichia coli chromosome.*, *The EMBO Journal* **23** (2004), no. 21 4330–4341.

[46] V. S. Lioy, A. Cournac, M. Marbouty, S. Duigou, J. Mozziconacci, O. Espéli, F. Boccard, and R. Koszul, *Multiscale structuring of the e. coli chromosome by nucleoid-associated and condensin proteins*, *Cell* **172** (2018), no. 4 771–783.

[47] L. Postow, C. D. Hardy, J. Arsuaga, and N. R. Cozzarelli, *Topological domain structure of the Escherichia coli chromosome.*, *Genes & Development* **18** (2004), no. 14 1766–1779.

[48] D. F. Browning, D. C. Grainger, and S. J. Busby, *Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression.*, *Current Opinion in Microbiology* **13** (2010), no. 6 773–780.

[49] S. Kar, R. Edgar, and S. Adhya, *Nucleoid remodeling by an altered HU protein: reorganization of the transcription program.*, *Proceedings of the National Academy of Sciences of the United States of America* **102** (2005), no. 45 16397–16402.

[50] D. J. Jin and J. E. Cabrera, *Coupling the distribution of RNA polymerase to global gene regulation and the dynamic structure of the bacterial nucleoid in Escherichia coli.*, *Journal of Structural Biology* **156** (2006), no. 2 284–291.

[51] M. Berger, A. Farcas, M. Geertz, P. Zhelyazkova, K. Brix, A. Travers, and G. Muskhelishvili, *Coordination of genomic structure and transcription by the main bacterial nucleoid-associated protein HU.*, *EMBO Reports* **11** (2010), no. 1 59–64.

[52] T. A. Azam, A. Iwata, A. Nishimura, T. A. L. I. Azam, and S. Ueda, *Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid*, *Journal of Bacteriology* **181** (1999), no. 20 6361–6370.

[53] S. B. Zimmerman, *Shape and compaction of Escherichia coli nucleoids.*, *Journal of Structural Biology* **156** (2006), no. 2 255–261.

[54] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Stochastic gene expression in a single cell.*, *Science (New York, N.Y.)* **297** (2002), no. 5584 1183–1186.

[55] A. Raj and A. van Oudenaarden, *Nature, nurture, or chance: stochastic gene expression and its consequences.*, *Cell* **135** (2008), no. 2 216–226.

[56] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, *Real-time kinetics of gene activity in individual bacteria.*, *Cell* **123** (2005), no. 6 1025–1036.

[57] S. Chong, C. Chen, H. Ge, and X. S. Xie, *Mechanism of Transcriptional Bursting in Bacteria*, *Cell* **158** (2014), no. 2 314–326.

[58] P. Bordes, A. Conter, V. Morales, J. Bouvier, A. Kolb, and C. Gutierrez, *DNA supercoiling contributes to disconnect sigmaS accumulation from sigmaS-dependent transcription in Escherichia coli.*, *Molecular Microbiology* **48** (2003), no. 2 561–571.

[59] D. a. Koster, A. Crut, S. Shuman, M.-A. Bjornsti, and N. H. Dekker, *Cellular strategies for regulating DNA supercoiling: a single-molecule perspective.*, *Cell* **142** (2010), no. 4 519–530.

[60] A. Travers and G. Muskhelishvili, *Bacterial chromatin.*, *Current Opinion in Genetics & Development* **15** (2005), no. 5 507–514.

[61] S. Deng, R. a. Stein, and N. P. Higgins, *Organization of supercoil domains and their reorganization by transcription.*, *Molecular Microbiology* **57** (2005), no. 6 1511–1521.

[62] T. B. Le, M. V. Imakaev, L. A. Mirny, and M. T. Laub, *High-resolution mapping of the spatial organization of a bacterial chromosome*, *Science* **342** (2013), no. 6159 731–734.

[63] T. B. Le and M. T. Laub, *Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries*, *The EMBO Journal* (2016) e201593561.

[64] C. R. Clapier and B. R. Cairns, *The biology of chromatin remodeling complexes.*, *Annual Review of Biochemistry* **78** (2009) 273–304.

[65] T. C. Voss and G. L. Hager, *Dynamic regulation of transcriptional states by chromatin and transcription factors.*, *Nature Reviews Genetics* **15** (2014), no. 2 69–81.

[66] D. Zink, M. D. Amaral, A. Englmann, S. Lang, L. A. Clarke, C. Rudolph, F. Alt, K. Luther, C. Braz, N. Sadoni, *et. al.*, *Transcription-dependent spatial arrangements of cftr and adjacent genes in human cell nuclei*, *Journal of Cell Biology* **166** (2004), no. 6 815–825.

[67] T. Ragoczy, M. Bender, A. Telling, R. Byron, and M. Groudine, *The locus control region is required for association of the murine β-globin locus with engaged transcription factories during erythroid maturation*, *Genes & Development* **20** (2006), no. 11 1447–1457.

[68] M. Melé and J. L. Rinn, *cats cradling the 3d genome by the act of lncrna transcription*, *Molecular Cell* **62** (2016), no. 5 657–664.

[69] P. A. Latos, F. M. Pauler, M. V. Koerner, H. B. Şenergin, Q. J. Hudson, R. R. Stocsits, W. Allhoff, S. H. Stricker, R. M. Klement, K. E. Warczok, *et. al.*, *Airn transcriptional overlap, but not its lncrna products, induces imprinted igf2r silencing*, *Science* **338** (2012), no. 6113 1469–1472.

[70] J. Dekker, A. S. Belmont, M. Guttman, V. O. Leshyk, J. T. Lis, S. Lomvardas, L. A. Mirny, C. C. Oshea, P. J. Park, B. Ren, *et. al.*, *The 4d nucleome project*, *Nature* **549** (2017), no. 7671 219–226.

[71] I. R. Lehman, M. J. Bessman, E. S. Simms, and A. Kornberg, *Enzymatic synthesis of deoxyribonucleic acid i. preparation of substrates and partial purification of an enzyme from escherichia coli*, *Journal of Biological Chemistry* **233** (1958), no. 1 163–170.

[72] M. J. Bessman, I. Lehman, E. S. Simms, and A. Kornberg, *Enzymatic synthesis of deoxyribonucleic acid ii. general properties of the reaction*, *Journal of Biological Chemistry* **233** (1958), no. 1 171–177.

[73] N. Kresge, R. D. Simoni, and R. L. Hill, *Arthur kornberg's discovery of dna polymerase i*, *Journal of Biological Chemistry* **280** (2005), no. 49 e46–e48.

[74] M. Grunberg-Manago and S. Ochoa, *Enzymatic synthesis and breakdown of polynucleotides; polynucleotide phosphorylase*, *Journal of the American Chemical Society* **77** (1955), no. 11 3165–3166.

[75] A. Kornberg, *Remembering our teachers*, *Journal of Biological Chemistry* **276** (2001), no. 1 3–11.

[76] M. W. Nirenberg and J. H. Matthaei, *The dependence of cell-free protein synthesis in e. coli upon naturally occurring or synthetic polyribonucleotides*, *Proceedings of the National Academy of Sciences* **47** (1961), no. 10 1588–1602.

[77] M. Salas, M. A. Smith, W. M. Stanley, A. J. Wahba, and S. Ochoa, *Direction of reading of the genetic message*, *Journal of Biological Chemistry* **240** (1965), no. 10 3988–3995.

[78] N. Kresge, R. D. Simoni, and R. L. Hill, *How severo ochoa determined the direction of the genetic code*, *Journal of Biological Chemistry* **281** (2006), no. 21 e16–e17.

[79] N. F. Lue and R. D. Kornberg, *Accurate initiation at rna polymerase ii promoters in extracts from saccharomyces cerevisiae*, *Proceedings of the National Academy of Sciences* **84** (1987), no. 24 8839–8843.

[80] R. D. Kornberg, *The molecular basis of eukaryotic transcription*, *Proceedings of the National Academy of Sciences* **104** (2007), no. 32 12955–12961.

[81] D. T. Nguyen and O. A. Saleh, *Tuning phase and aging of dna hydrogels through molecular design*, *Soft Matter* **13** (2017), no. 32 5421–5427.

[82] B.-j. Jeon, D. T. Nguyen, G. R. Abraham, N. Conrad, D. K. Fygenson, and O. A. Saleh, *Salt-dependent properties of a coacervate-like, self-assembled dna liquid*, *Soft Matter* **14** (2018), no. 34 7009–7015.

[83] C.-Y. Park, D. R. Jacobson, D. T. Nguyen, S. Willardson, and O. A. Saleh, *A thin permeable-membrane device for single-molecule manipulation*, *Review of Scientific Instruments* **87** (2016), no. 1 014301.

[84] F. A. Aldaye, A. L. Palmer, and H. F. Sleiman, *Assembling materials with dna as the guide*, *science* **321** (2008), no. 5897 1795–1799.

[85] M. R. Jones, N. C. Seeman, and C. A. Mirkin, *Programmable materials and the nature of the dna bond*, *Science* **347** (2015), no. 6224 1260901–1–1260901–11, [http://science.sciencemag.org/content/347/6224/1260901.full.pdf].

[86] N. C. Seeman, *Nucleic acid junctions and lattices.*, *Journal of Theoretical Biology* **99** (1982), no. 2 237–247.

[87] S. M. Douglas, I. Bachelet, and G. M. Church, *A logic-gated nanorobot for targeted transport of molecular payloads*, *Science* **335** (2012), no. 6070 831–834.

[88] S. Raniolo, G. Vindigni, V. Unida, A. Ottaviani, E. Romano, A. Desideri, and S. Biocca, *Entry, fate and degradation of dna nanocages in mammalian cells: a matter of receptors*, *Nanoscale* **10** (2018), no. 25 12078–12086.

[89] S. H. Um, J. B. Lee, N. Park, S. Y. Kwon, C. C. Umbach, and D. Luo, *Enzyme-catalysed assembly of DNA hydrogel.*, *Nature Materials* **5** (2006), no. 10 797–801.

[90] J. Li, C. Zheng, S. Cansiz, C. Wu, J. Xu, C. Cui, Y. Liu, W. Hou, Y. Wang, L. Zhang, I.-t. Teng, H.-H. Yang, and W. Tan, *Self-assembly of DNA Nanohydrogels with Controllable Size and Stimuli-Responsive Property for Targeted Gene Regulation Therapy*, *Journal of the American Chemical Society* **137** (2015), no. 4 1412–1415.

[91] L. D. Shea, E. Smiley, J. Bonadio, and D. J. Mooney, *DNA delivery from polymer matrices for tissue engineering.*, *Nature Biotechnology* **17** (1999), no. 6 551–554.

[92] Z. Zhu, C. Wu, H. Liu, Y. Zou, X. Zhang, H. Kang, C. J. Yang, and W. Tan, *An aptamer cross-linked hydrogel as a colorimetric platform for visual detection*, *Angewandte Chemie* **122** (2010), no. 6 1070–1074.

[93] X. He, B. Wei, and Y. Mi, *Aptamer based reversible DNA induced hydrogel system for molecular recognition and separation*, *Chemical Communications* **46** (2010), no. 34 6308–6310.

[94] S. Nagahara and T. Matsuda, *Hydrogel formation via hybridization of oligonucleotides derivatized in water-soluble vinyl polymers*, *Polymer Gels and Networks* **4** (1996), no. 2 111–127.

[95] E. Cheng, Y. Xing, P. Chen, Yang, Y. Sun, D. Zhou, T. Xu, Q. Fan, and D. Liu, *A pH-triggered, fast-responding DNA hydrogel*, *Angewandte Chemie* **48** (2009), no. 41 7660–7663.

[96] H. Kang, H. Liu, X. Zhang, J. Yan, Z. Zhu, L. Peng, H. Yang, Y. Kim, and W. Tan, *Photoresponsive DNA-Cross-Linked Hydrogels for Controllable Release and Cancer Therapy*, *Langmuir* **27** (2011), no. 1 399–408.

[97] D. C. Lin, B. Yurke, and N. A. Langrana, *Mechanical Properties of a Reversible, DNA-Crosslinked Polyacrylamide Hydrogel*, *Journal of Biomechanical Engineering* **126** (2004), no. 1 104–110.

[98] Y. Xing, E. Cheng, Y. Yang, P. Chen, T. Zhang, Y. Sun, Z. Yang, and D. Liu, *Self-Assembled DNA Hydrogels with Designable Thermal and Enzymatic Responsiveness*, *Advanced Materials* **23** (2011), no. 9 1117–1121.

[99] D. Wang, Y. Hu, P. Liu, and D. Luo, *Bioresponsive DNA Hydrogels: Beyond the Conventional Stimuli Responsiveness*, *Accounts of Chemical Research* **50** (2017), no. 4 733–739.

[100] S. Biffi, R. Cerbino, F. Bomboi, E. M. Paraboschi, R. Asselta, F. Sciortino, and T. Bellini, *Phase behavior and critical activated dynamics of limited-valence DNA nanostars.*, *Proceedings of the National Academy of Sciences of the United States of America* **110** (2013), no. 39 15633–15637.

[101] J. Berry, S. C. Weber, N. Vaidya, M. Haataja, and C. P. Brangwynne, *Rna transcription modulates phase transition-driven nuclear body assembly*, *Proceedings of the National Academy of Sciences* **112** (2015), no. 38 E5237–E5245.

[102] H. Zhang, S. Elbaum-Garfinkle, E. M. Langdon, N. Taylor, P. Occhipinti, A. A. Bridges, C. P. Brangwynne, and A. S. Gladfelter, *RNA Controls PolyQ Protein Phase Transitions*, *Molecular Cell* **60** (2015), no. 2 220–230.

[103] M. Verheul and S. P. F. M. Roefs, *Structure of Particulate Whey Protein Gels: Effect of NaCl Concentration, pH, Heating Temperature, and Protein Composition*, *Journal of Agricultural and Food Chemistry* **46** (1998), no. 12 4909–4916.

[104] N. R. Markham and M. Zuker, *DINAMelt web server for nucleic acid melting prediction*, *Nucleic Acids Research* **33** (2005) W577–W581.

[105] D. Costa, M. G. Miguel, and B. Lindman, *Swelling properties of cross-linked DNA gels*, *Advances in Colloid and Interface Science* **158** (2010), no. 1 21–31.

[106] A. R. Mount, C. P. Mountford, S. A. G. Evans, T. J. Su, A. H. Buck, P. Dickinson, C. J. Campbell, L. M. Keane, J. G. Terry, J. S. Beattie, A. J. Walton, P. Ghazal, and J. Crain, *The stability and characteristics of a DNA Holliday junction switch*, *Biophysical Chemistry* **124** (2006), no. 3 214–221.

[107] P. Shrestha, T. Emura, D. Koirala, Y. Cui, K. Hidaka, W. J. Maximuck, M. Endo, H. Sugiyama, and H. Mao, *Mechanical properties of DNA origami nanoassemblies are determined by Holliday junction mechanophores*, *Nucleic Acids Research* **44** (2016), no. 14 6574–6582.

[108] K. T. Delaney and G. H. Fredrickson, *Theory of polyelectrolyte complexationcomplex coacervates are self-coacervates*, *The Journal of Chemical Physics* **146** (2017), no. 22 224902, [http://dx.doi.org/10.1063/1.4985568].

[109] C. G. De Kruif, F. Weinbreck, and R. de Vries, *Complex coacervation of proteins and anionic polysaccharides*, *Current Opinion in Colloid & Interface Science* **9** (2004), no. 5 340–349.

[110] D. Priftis and M. Tirrell, *Phase behaviour and complex coacervation of aqueous polypeptide solutions*, *Soft Matter* **8** (2012), no. 36 9396–9405.

[111] Z. Xiao, W. Liu, G. Zhu, R. Zhou, and Y. Niu, *A review of the preparation and application of flavour and essential oils microcapsules based on complex coacervation technology*, *Journal of the Science of Food and Agriculture* **94** (2014), no. 8 1482–1494.

[112] A. A. Hyman, C. A. Weber, and F. Jülicher, *Liquid-liquid phase separation in biology*, *Annual Review of Cell and Developmental Biology* **30** (2014) 39–58.

[113] C. P. Brangwynne, P. Tompa, and R. V. Pappu, *Polymer physics of intracellular phase transitions*, *Nature Physics* **11** (2015), no. 11 899–904.

[114] L. L. Josephson, E. M. Furst, and W. J. Galush, *Particle tracking microrheology of protein solutions*, *Journal of Rheology* **60** (2016), no. 4 531–540.

[115] S. Elbaum-Garfinkle, Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, S. Myong, and C. P. Brangwynne, *The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics*, *Proceedings of the National Academy of Sciences* **112** (2015), no. 23 7189–7194.

[116] S. Seiffert and W. Oppermann, *Systematic evaluation of frap experiments performed in a confocal laser scanning microscope*, *Journal of Microscopy* **220** (2005), no. 1 20–30.

[117] L. G. Leal, *Advanced Transport Phenomena: Fluid Mechanics and Convective Transport Processes*. Cambridge University Press, Cambridge, 1st ed., 2007.

[118] J. SantaLucia, *A unified view of polymer, dumbbell, and oligonucleotide dna nearest-neighbor thermodynamics*, *Proceedings of the National Academy of Sciences* **95** (1998), no. 4 1460–1465.

[119] H. Eyring, *Viscosity, plasticty, and diffusion as examples of absolute reaction rates*, *The Journal of Chemical Physics* **4** (1936), no. 4 283–291.

[120] F. Bomboi, S. Biffi, R. Cerbino, T. Bellini, F. Bordi, and F. Sciortino, *Equilibrium gels of trivalent DNA-nanostars: Effect of the ionic strength on the dynamics*, *The European Physical Journal E* **38** (2015), no. 64.

[121] J. Fernandez-Castanon, F. Bomboi, L. Rovigatti, M. Zanatta, A. Paciaroni, L. Comez, L. Porcar, C. J. Jafta, G. C. Fadda, T. Bellini, *et. al.*, *Small-angle neutron scattering and molecular dynamics structural study of gelling dna nanostars*, *The Journal of Chemical Physics* **145** (2016), no. 8 084910.

[122] M. Ortiz-Lombardía, A. González, R. Eritja, J. Aymamí, F. Azorín, and M. Coll, *Crystal structure of a DNA Holliday junction.*, *Nature Structural Biology* **6** (1999), no. 10 913–917.

[123] P. A. Egelstaff, *An introduction to the liquid state*. Clarendon Press, Oxford, 2nd ed., 1994.

[124] T. Keyes and I. Oppenheim, *Bilinear hydrodynamics and the Stokes-Einstein law*, *Physical Review A* **8** (1973), no. 2 937–949.

[125] A. Jaiswal, T. Egami, K. F. Kelton, K. S. Schweizer, and Y. Zhang, *Correlation between fragility and the Arrhenius crossover phenomenon in metallic, molecular, and network liquids*, *Physical Review Letters* **117** (2016), no. 20 205701, [arXiv:1604.0892].

[126] H. Bohidar, P. L. Dubin, P. R. Majhi, C. Tribet, and W. Jaeger, *Effects of protein-polyelectrolyte affinity and polyelectrolyte molecular weight on dynamic properties of bovine serum albumin-poly(diallyldimethylammonium chloride) coacervates*, *Biomacromolecules* **6** (2005), no. 3 1573–1585.

[127] E. Spruijt, J. Sprakel, M. A. C. Stuart, and J. van der Gucht, *Interfacial tension between a complex coacervate phase and its coexisting aqueous phase*, *Soft Matter* **6** (2010), no. 1 172–178.

[128] E. Spruijt, A. H. Westphal, J. W. Borst, M. A. Cohen Stuart, and J. van der Gucht, *Binodal compositions of polyelectrolyte complexes*, *Macromolecules* **43** (2010), no. 15 6476–6484.

[129] M. Lemmers, J. Sprakel, I. K. Voets, J. van der Gucht, and M. A. C. Stuart, *Multiresponsive reversible gels based on charge-driven assembly*, *Angewandte Chemie* **49** (2010), no. 4 708–711.

[130] D. Priftis, R. Farina, and M. Tirrell, *Interfacial energy of polypeptide complex coacervates measured via capillary adhesion*, *Langmuir* **28** (2012), no. 23 8721–8729.

[131] J. Van der Gucht, E. Spruijt, M. Lemmers, and M. A. C. Stuart, *Polyelectrolyte complexes: bulk phases and colloidal systems*, *Journal of Colloid and Interface Science* **361** (2011), no. 2 407–422.

[132] E. Spruijt, M. A. C. Stuart, and J. van der Gucht, *Effects of protein-polyelectrolyte affinity and polyelectrolyte molecular weight on dynamic properties of bovine serum albumin-poly(diallyldimethylammonium chloride) coacervates*, *Macromolecules* **46** (2013), no. 4 1633–1641.

[133] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, *Germline p granules are liquid droplets that localize by controlled dissolution/condensation*, *Science* **324** (2009), no. 5935 1729–1732.

[134] M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, and C. P. Brangwynne, *Coexisting liquid phases underlie nucleolar subcompartments*, *Cell* **165** (2016), no. 7 1686–1697.

[135] M. T. Wei, S. Elbaum-Garfinkle, A. S. Holehouse, C. C. H. Chen, M. Feric, C. B. Arnold, R. D. Priestley, R. V. Pappu, and C. P. Brangwynne, *Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles*, *Nature Chemistry* **9** (2017), no. 11 1118–1125.

[136] Y. Shin, J. Berry, N. Pannucci, M. P. Haataja, J. E. Toettcher, and C. P. Brangwynne, *Spatiotemporal control of intracellular phase transitions using light-activated optodroplets*, *Cell* **168** (2017), no. 1-2 159–171.

[137] D. M. Mitrea and R. W. Kriwacki, *Phase separation in biology; functional organization of a higher order*, *Cell Communication and Signaling* **14** (2016), no. 1.

[138] F.-M. Boisvert, S. van Koningsbruggen, J. Navascués, and A. I. Lamond, *The multifunctional nucleolus*, *Nature Reviews Molecular Cell Biology* **8** (2007), no. 7 574–585.

[139] A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, and G. J. Narlikar, *Liquid droplet formation by hp1α suggests a role for phase separation in heterochromatin*, *Nature* **547** (2017), no. 7662 236–240.

[140] A. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, and G. H. Karpen, *Phase separation drives heterochromatin domain formation*, *Nature* **547** (2017), no. 7662 241–245.

[141] T. Misteli and E. Soutoglou, *The emerging role of nuclear architecture in dna repair and genome maintenance*, *Nature Reviews Molecular Cell Biology* **10** (2009), no. 4 243–254.

[142] I. Chiolo, A. Minoda, S. U. Colmenares, A. Polyzos, S. V. Costes, and G. H. Karpen, *Double-strand breaks in heterochromatin move outside of a dynamic hp1a domain to complete recombinational repair*, *Cell* **144** (2011), no. 5 732–744.

[143] R. Martienssen and D. Moazed, *Rnai and heterochromatin assembly*, *Cold Spring Harbor Perspectives in Biology* **7** (2015), no. 8 a019323.

[144] J. C. Peng and G. H. Karpen, *Epigenetic regulation of heterochromatic dna stability*, *Current Opinion in Genetics & Development* **18** (2008), no. 2 204–211.

[145] A. Klosin and A. A. Hyman, *A liquid reservoir for silent chromatin*, *Nature* **547** (2017), no. 7662 168–170.

[146] B. Button, L.-H. Cai, C. Ehre, M. Kesimer, D. B. Hill, J. K. Sheehan, R. C. Boucher, and M. Rubinstein, *A periciliary brush promotes the lung health by separating the mucus layer from airway epithelia*, *Science* **337** (2012), no. 6097 937–941.

[147] *Fluorescein isothiocyanate dextran - technical report*, https://www.sigmaaldrich.com/technical-documents/articles/chemistry/fluorescently-labeled-dextrane.html.

[148] J. Gal, R. Schnell, S. Szekeres, and M. Kalman, *Directional cloning of native pcr products with preformed sticky ends (autosticky pcr)*, *Molecular and General Genetics* **260** (1999), no. 6 569–573.

[149] S. I. Gunderson, K. A. Chapman, and R. R. Burgess, *Interactions of t7 rna polymerase with t7 late promoters measured by footprinting with methidiumpropyl-edta-iron (ii)*, *Biochemistry* **26** (1987), no. 6 1539–1546.

[150] A. Zemella, L. Thoring, C. Hoffmeister, and S. Kubick, *Cell-free protein synthesis: Pros and cons of prokaryotic and eukaryotic systems*, *ChemBioChem* **16** (2015), no. 17 2420–2431.

[151] T. U. Arumugam, D. Ito, E. Takashima, M. Tachibana, T. Ishino, M. Torii, and T. Tsuboi, *Application of wheat germ cell-free protein expression system for novel malaria vaccine candidate discovery*, *Expert Review of Vaccines* **13** (2014), no. 1 75–85.

[152] A. S. Salehi, M. T. Smith, A. M. Bennett, J. B. Williams, W. G. Pitt, and B. C. Bundy, *Cell-free protein synthesis of a cytotoxic cancer therapeutic: Onconase production and a just-add-water cell-free system*, *Biotechnology Journal* **11** (2016), no. 2 274–281.

[153] X. Ge, D. Luo, and J. Xu, *Cell-free protein expression under macromolecular crowding conditions*, *PLoS One* **6** (2011), no. 12 e28707.

[154] E. Sokolova, E. Spruijt, M. M. Hansen, E. Dubuc, J. Groen, V. Chokkalingam, A. Piruska, H. A. Heus, and W. T. Huck, *Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate*, *Proceedings of the National Academy of Sciences* **110** (2013), no. 29 11692–11697.

[155] N. Park, S. H. Um, H. Funabashi, J. Xu, and D. Luo, *A cell-free protein-producing gel.*, *Nature Materials* **8** (may, 2009) 432–437.

[156] S. S. Daube, D. Bracha, A. Buxboim, and R. H. Bar-Ziv, *Compartmentalization by directional gene expression*, *Proceedings of the National Academy of Sciences* **107** (2010), no. 7 2836–2841.

[157] K. Maeshima, S. Ide, K. Hibino, and M. Sasai, *Liquid-like behavior of chromatin*, *Current Opinion in Genetics & Development* **37** (2016) 36–45.

[158] M. Chamberlin and J. Ring, *Characterization of t7-specific ribonucleic acid polymerase i. general properties of the enzymatic reaction and the template specificity of the enzyme*, *Journal of Biological Chemistry* **248** (1973), no. 6 2235–2244.

[159] W. Song, R. L. Strack, N. Svensen, and S. R. Jaffrey, *Plug-and-play fluorophores extend the spectral properties of spinach*, *Journal of the American Chemical Society* **136** (2014), no. 4 1198–1201.

[160] R. J. Trachman III, L. Truong, and A. R. Ferré-D'Amaré, *Structural principles of fluorescent rna aptamers*, *Trends in Pharmacological Sciences* **38** (2017), no. 10 928–939.

[161] R. A. Ikeda, A. C. Lin, and J. Clarke, *Initiation of transcription by t7 rna polymerase as its natural promoters.*, *Journal of Biological Chemistry* **267** (1992), no. 4 2640–2649.

[162] R. L. Strack, M. D. Disney, and S. R. Jaffrey, *A superfolding spinach2 reveals the dynamic nature of trinucleotide repeat–containing rna*, *Nature Methods* **10** (2013), no. 12 1219–1224.

[163] J. J. Dunn, F. W. Studier, and M. Gottesman, *Complete nucleotide sequence of bacteriophage t7 dna and the locations of t7 genetic elements*, *Journal of Molecular Biology* **166** (1983), no. 4 477–535.

[164] L. Hartvig and J. Christiansen, *Intrinsic termination of t7 rna polymerase mediated by either rna or dna.*, *The EMBO Journal* **15** (1996), no. 17 4767–4774.

[165] J. Mairhofer, A. Wittwer, M. Cserjan-Puschmann, and G. Striedner, *Preventing t7 rna polymerase read-through transcription a synthetic termination signal capable of improving bioprocess stability*, *ACS Synthetic Biology* **4** (2014), no. 3 265–273.

[166] M. Chamberlin and J. Ring, *Characterization of t7-specific ribonucleic acid polymerase ii. inhibitors of the enzyme and their application to the study of the enzymatic reaction*, *Journal of Biological Chemistry* **248** (1973), no. 6 2245–2250.

[167] J. Smrek and K. Kremer, *Small activity differences drive phase separation in active-passive polymer mixtures*, *Physical Review Letters* **118** (2017), no. 9 098002.

[168] H. Sutherland and W. A. Bickmore, *Transcription factories: gene expression in unions?*, *Nature Reviews Genetics* **10** (2009), no. 7 457–466.

[169] W.-K. Cho, J.-H. Spille, M. Hecht, C. Lee, C. Li, V. Grube, and I. I. Cisse, *Mediator and rna polymerase ii clusters associate in transcription-dependent condensates*, *Science* (2018) eaar4199.

[170] R. M. Dirks and N. A. Pierce, *Triggered amplification by hybridization chain reaction*, *Proceedings of the National Academy of Sciences* **101** (2004), no. 43 15275–15278.

[171] D. Y. Zhang and E. Winfree, *Control of dna strand displacement kinetics using toehold exchange*, *Journal of the American Chemical Society* **131** (2009), no. 47 17303–17314.

[172] N. Park, J. S. Kahn, E. J. Rice, M. R. Hartman, H. Funabashi, J. Xu, S. H. Um, and D. Luo, *High-yield cell-free protein production from P-gel.*, *Nature Protocols* **4** (2009), no. 12 1759–1770.

[173] S. Ishino and Y. Ishino, *Dna polymerases as useful reagents for biotechnology–the history of developmental research in the field*, *Frontiers in Microbiology* **5** (2014) 465.

[174] J. M. Clark, *Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic dna polymerases*, *Nucleic Acids Research* **16** (1988), no. 20 9677–9686.

[175] T. Holton and M. Graham, *A simple and efficient method for direct cloning of pcr products using ddt-tailed vectors.*, *Nucleic Acids Research* **19** (1991), no. 5 1156.

[176] J. A. Sikorsky, D. A. Primerano, T. W. Fenger, and J. Denvir, *Dna damage reduces taq dna polymerase fidelity and pcr amplification efficiency*, *Biochemical and Biophysical Research Communications* **355** (2007), no. 2 431–437.

[177] J. Gál, R. Schnell, and M. Kálmán, *Polymerase dependence of autosticky polymerase chain reaction*, *Analytical Biochemistry* **282** (2000), no. 1 156–158.

[178] J. R. Beckwith and E. R. Signer, *Transposition of the lac region of escherichia coli: I. inversion of the lac operon and transduction of lac by $\phi 80$*, *Journal of Molecular Biology* **19** (1966), no. 2 254–265.

[179] I. E. Ichetovkin, G. Abramochkin, and T. E. Shrader, *Substrate recognition by the leucyl/phenylalanyl-trna-protein transferase conservation within the enzyme family and localization to the trypsin-resistant domain*, *Journal of Biological Chemistry* **272** (1997), no. 52 33009–33014.

[180] J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, and G. S. Waldo, *Engineering and characterization of a superfolder green fluorescent protein*, *Nature Biotechnology* **24** (2006), no. 1 79–88.

[181] R. N. Day and M. W. Davidson, *The fluorescent protein palette: tools for cellular imaging*, *Chemical Society Reviews* **38** (2009), no. 10 2887–2921.

[182] S. Preibisch, S. Saalfeld, and P. Tomancak, *Globally optimal stitching of tiled 3D microscopic image acquisitions*, *Bioinformatics* **25** (2009), no. 11 1463–1465.

[183] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, *Fiji: an open-source platform for biological-image analysis*, *Nature Methods* **9** (2012), no. 7 676–682, [1081-8693].

[184] J.-P. Thiery, G. Macaya, and G. Bernardi, *An analysis of eukaryotic genomes by density gradient centrifugation*, *Journal of Molecular Biology* **108** (1976), no. 1 219–235.

[185] G. K. Batchelor, *An introduction to fluid dynamics*. Cambridge university press, 2000.

[186] D. W. Green, *Perrys chemical engineers*. McGraw Hill, 2008.

[187] S. J. DeCamp, G. S. Redner, A. Baskaran, M. F. Hagan, and Z. Dogic, *Orientational order of motile defects in active nematics*, *Nature Materials* **14** (2015), no. 11 1110–1115, [arXiv:1501.0622].