Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

George Poppitz April 10, 2023

Decoding the Rules for Biopolymer Mutualism

by

George Poppitz

Dr. David Lynn

Adviser

Chemistry

Dr. David Lynn

Adviser

Dr. Brian Dyer

Committee Member

Dr. Arri Eisen

Committee Member

2023

Decoding the Rules for Biopolymer Mutualism

By

George Poppitz

Dr. David Lynn

Adviser

An abstract of

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Science with Honors

Chemistry

2023

Abstract

Decoding the Rules for Biopolymer Mutualism

By George Poppitz

The interaction between nucleic-acid and amino-acid-based biopolymers is essential for the development of modern cellular life. However, despite decades of investigation, it remains unclear how life originated and how this mutualistic interaction came to exist. Biomolecular condensates are an exciting avenue for exploring this question due to their commonality in biology, but their widely varied composition means new models must be developed. Building on decades of work with peptide models of aggregation, we investigated the co-assembly of peptides and nucleic acids to study how electrostatic interactions underpin their ability of peptides and nucleic acids to assemble. Following this, comparisons between different nucleic acid conformations revealed a bias towards structures with a higher density of negative charges. To confirm this, a cyclic polyanion was used to discover new rules for how the selection of different assembly structures functions.

Decoding the Rules for Biopolymer Mutualism

By

George Poppitz

Dr. David Lynn

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Chemistry

2023

Acknowledgements

Firstly I'd like to thank my adviser Dr. Lynn. I know that without his guidance and mentorship I would not be half the scientist I am today.

Secondly, I'd like to thank the all the graduate students in the Lynn lab who have greatly helped me in my journey. I would especially like to thank Christella Gordon-Kim and Alexis Blake for all the time they have invested in me. I sincerely believe that without your mentorship and assistance I would never have been accepted to graduate programs nor would I be ready to take the next step in my career.

Thirdly, I'd like to thank the many chemistry department professors that have assisted me along the way. All of the professors I have taken classes from have been instrumental to my development but I would especially like to thank Dr. Salaita and Dr. Dyer for introducing me to subjects I love and providing me with a fantastic foundation of knowledge to bring with me into graduate school. Dr. Wuest, Dr. Weinschenk, and Dr. B all helped me greatly in applying to grad school either through writing supporting letters of providing me guidance and helping me refine my list of programs. All of you are fantastic teachers and mentors and I hope to cross paths with you all again.

Fourthly, I'd like to thank all the faculty involved in the Emory-Tibet Science Initiative for allowing me to have such a fantastic opportunity to make a difference. I have enjoyed working alongside Dr. Eisen and Dr. Gray and I deeply appreciate their support. Additionally, I want to thank all of the Tenzin Gyatso Science Scholars of the 5th and 6th cohort: I hope that I was able to teach you even a small fraction of what you have taught me.

Finally, I want to thank all of my friends for their support but especially David Goldberg for constantly pushing me to be better. Come find me at Stanford.

Table of Contents

List of Figures and Tables

Figures

Introduction

A brief review of origin-of-life research

The origin of life on earth has been one of the most enduring scientific questions of all time. Aristotle's original idea, that "lower" life like insects or maggots spontaneously arose by chance due to decaying matter, was broadly accepted up until the 17th century when experiments by Francesco Redi demonstrated that flies laying eggs was necessary for maggots to form (Sheldon 2005; Gottdenker 1979). Additionally, Louis Pasteur demonstrated in the 1860s that microorganisms must also originate from parent organisms, firmly putting to rest ideas of spontaneous generation (Bordenave 2003). While these experiments successfully defeated the incorrect idea of spontaneous generation they left scientists with a new question, one that has continued to challenge scientists even today: if all life requires a progenitor, then how did the first lifeform arise?

This central paradox, "if everything must come from something else then what came first?", frustrated researchers for hundreds of years and many scientists did not bother seriously interrogating this idea. As more and more complexities of cellular life were discovered such as the nucleus or enzymes, disheartened scientists saw the gap between living and non-living matter far too large to bridge. However, in the 1920s Alexander I. Oparin proposed the heterotrophic origin of life theory, reasoning that some synthesis of rudimentary organic compounds would have been possible in the oxygen-deficient atmosphere of early earth and that these early organic structures would be subject to Darwinian natural selection (Lazcano 2016). This theory likely influenced Stanley Miller and Harold Urey when, 20 years later, they demonstrated formation of amino acids from suggested early-earth conditions (Miller 1953).

The Miller-Urey experiment was a key step forward in origins-of-life research and serves as a fantastic example of the bottom-up approach to origins-of-life research.

The two different approaches to origin-of-life research are top-down and bottom-up. Top-down approaches are characterized by finding commonalities amongst lifeforms to try and deduce what the last universal common ancestor (LUCA) looked like. This has largely focused on prokaryotes as the simplest forms of cellular life (Cleland 2013). Earlier work using this approach led to the exciting result that there is a minimal set of genetic information required for cellular life (Mushegian and Koonin 1996), but work attempting to look further back in time than the minimal bacterial genome has largely stalled. In contrast to this, bottom-up approaches generally investigate how the first structures capable of evolution may have formed from pre-Darwinian origins. Here research focuses on prebiotic synthetic pathways to access biomolecules, self-assembly, chemical information storage and transfer, and metabolic networks (Ruiz-Mirazo, Briones, and de la Escosura 2014). Historically the top-down and bottom-up approaches were seen as mutually exclusive and some argue that this view has stymied scientific progress in the field (Preiner et al. 2020). Supporting this assertion that joining top-down and bottom-up approaches is crucial for meaningful research progress, the most significant and enduring hypothesis to arise from origins-of-life research has been the RNA world hypothesis.

First proposed by Alexander Rich in 1962, the RNA world hypothesis posits that RNA acted as both the information-carrying molecule enabling genetics and a catalyst enabling self-replication at the origin of life (Rich 1962; Lazcano 2012). In this paper, Rich uses a top-down approach when he uses the fact that RNA and DNA share structural and functional similarities

or the existence of RNA viruses to suggest that modern, DNA-based life evolved from more primitive RNA organisms. Likewise, Rich's hypothesis shares features of a bottom-up approach in the focus on reactivity and the ability to self-replicate as a molecule subject to Darwinian natural selection. The RNA world has persisted as a hypothesis due to its charming simplicity; only one molecule is responsible for catalysis and information so implausible scenarios relating to the spontaneous occurrence of multiple biopolymers are bypassed (Neveu, Kim, and Benner 2013). Over the years there have been multiple key experiments that lend support to the RNA world hypothesis. Chiefly among these have been experiments with ribozymes, catalytic RNA (Cech, Zaug, and Grabowski 1981). Take for example the breakthrough discovery of a self-replicating ribozyme in 2002 which proved that RNA was capable of creating copies of itself (Paul and Joyce 2002). This was followed by the creation of cross-replicating ribozymes which were able to compete against other ribozymes for shared substrates (Lincoln and Joyce 2009). Taken together, these studies demonstrated the viability of ribozymes as an avenue to gain insight into the origin of life.

However, now greater than 60 years old, the RNA world has had serious detractors over the decades and multiple key questions are still unanswered. For one, RNA is extremely synthetically inaccessible for prebiotic conditions. In fact, the discussion of prebiotic RNA synthesis once famously drew the comment that RNA is a "prebiotic chemists' nightmare" (Joyce and Orgel 1999). Beyond this, criticisms focus on the lack of RNA's stability to hydrolysis, the lack of catalysis on short RNAs, and limited catalytic function, just to name a few of the common accusations (Bernhardt 2012). Although ongoing research is attempting to answer many of these challenges and some progress has been made, (Szostak 2012; Bernhardt and

Tate 2012; Chumachenko, Novikov, and Yarus 2009) many recent works have shifted slightly in focus to consider a slightly modified variation of the RNA world that may be more prebiotically plausible.

The Ribonucleoprotein (RNP) world can go by many names like RNA-peptide world or oligonucleotide/oligopeptide world, but the central idea is always the same. The unifying theme in every formulation of this hypothesis is that mutualism between polypeptides and RNA was an essential intermediate stage in the origin of life. This idea can be thought of as stemming from the original Miller-Urey discovery of prebiotic amino acid synthesis (Miller 1953) and those original theories from Oparin about heterotrophic origins of life (Hyman and Brangwynne 2012). Amino acids are far more synthetically accessible than RNA so this hypothesis satisfies some bottom-up concerns. Furthermore, many ribozymes absolutely require protein components for stability and function, a fact that has been used to criticize RNA-only approaches to origin-of-life questions (Di Giulio 1997; Le Vay and Mutschler 2019). Some top-down theorists have likewise switched from considering the emergence of the first cellular organisms as where life began and instead began looking at the peptidyl transferase center, which is the catalytic heart of ribosomes, to examine how this mutualism between RNA and protein could hold the key for understanding the emergence of life (Prosdocimi, José, and de Farias 2021). Work has shown that the peptidyl transferase center is likely the most ancient and conserved center of ribosomes, and therefore cellular life (Farias, Rêgo, and José 2014), giving evidence to the theory that mutualistic interactions between RNA and proteins were essential for the development of increasing biological complexity and eventually life. Currently, many bottom-up approaches to the origin-of-life question focus on this aspect of the proposed

evolutionary pathway, striving to demonstrate the prebiotic creation of RNA/peptide materials (Müller et al. 2022). Similarly, top-down work has used the recent discovery of RNA/protein biomolecular condensates in both eukaryotes and prokaryotes to assert that these mutualistic interactions are ubiquitous in modern life because they are an universal, highly-conserved evolutionary adaptation that bridges the RNA world and current cellular life (Ladouceur et al. 2020; Banani et al. 2017). These new top-down and bottom-up approaches have fueled calls to further explore the RNP world as either an alternative hypothesis or a key development of the RNA world (Carter 2015). In light of these ongoing developments in the field of origin-of-life research, it is abundantly clear that further work is needed to understand mutualism between nucleic acids and peptides.

Biopolymer mutualism in biomolecular condensates

As a starting point for the discussion of nucleic acid and protein mutualism, first we must clarify the different categories of ribonucleoproteins (RNPs). Many RNPs form definite structures, the classic example of which is the ribosome. Ribosomes generally compose of 50% or greater RNA by mass which is strongly non-covalently bound to over 50 proteins (van de Waterbeemd et al. 2018; Benjamin et al. 1998). While there are some flexible domains, for the most part the ribosome is highly ordered, forming a definite structure which is essential for its catalytic function. The other large category of RNPs is disordered RNPs. These disordered structures do not take a definite crystalline structure, behaving like liquid droplets albeit with viscosities thousands of times higher than that of water (Brangwynne et al. 2009). These liquid droplets, commonly referred to as biomolecular condensates, are a separate phase from the

surrounding solvent and are biologically useful because they enable cells to organize themselves without relying on the relative complexity of lipid membranes (Hyman, Weber, and Jülicher 2014). For example, the bacterium *Caulobacter crescentus* has evolved to use ATP-concentration-dependent phase separation to control kinase function as well as phase separation of the protein PodJ for control of signaling hubs for asymmetric cell division (Saurabh et al. 2022; Tan et al. 2022). In eukaryotes, the classic examples of biomolecular condensates are germ granules, processing bodies, and stress granules (Gao and Arkov 2013; Guo, Shi, and Wang 2021). The ability of these condensates to perform cellular functions have led to these biomolecular condensates commonly being referred to as membranelles organelles (Lasker et al. 2020; Banani et al. 2017).

Recently great progress has been made to elucidate the physical interactions responsible for liquid-liquid phase separation and gain an understanding of condensates. Peptide models have successfully been employed to gain a deeper understanding of charge-patterning's importance for these interactions and the field has generally matured to have a comprehensive understanding of biomolecular condensate formation (Chang et al. 2017). In fact, recently developed peptide nanomaterials have been used to model biomolecular condensates in bacteria (Tomares et al. 2022). Generally, peptides are a tremendously useful model system to study all manner of supramolecular questions such as condensation or the maturation of condensates into aggregates (Sheehan et al. 2021). This is because peptides afford direct control over sequence which allows for finely tuned system-level behavior. Going beyond condensates, peptides have also been essential for studying the process of fibrilization. A peptide model derived from the nucleating core sequence of amyloid beta 42 has been used

extensively to probe dynamics in the underlying processes of condensate formation and especially for investigating processes where condensates transform into solid, paracrystalline assemblies like nanotubes or amyloid fibers (Hsieh, Lynn, and Grover 2017). The growth of paracrystalline assemblies out of condensate phases is implicated in prion diseases and Alzheimer's disease and often studied in this context (Wang et al. 2021), but studying this interaction can also shed light on the underlying physical rules of biopolymer mutualism in the context of RNPs.

Methods

Synthesis of Ac-KLVIIAG-NH_2 (Pep-KG) and Ac-ELVIIAG-NH_2 (Pep-EG)

Pep-KG was synthesized on Rink amide-MBHA (Anaspec) solid support via a CEM Liberty Blue Automated Microwave Peptide Synthesizer (Serial # LB2447) with 1M N,N'-Diisopropylcarbodiimide (DIC; CAS# 693-13-6 AAPPTEC) as the activator, Oxyma Pure (CAS# 3849-21-6 CEM) as the activator base, and 20% Piperidine (CAS# 110-89-4 Sigma-Aldrich) as the deprotection solution. Amino acids were coupled using 0.25 mmol standard coupling (75°C 210W for 15s followed by 90°C 30W for 110s) and deprotected using 0.25 mmol standard deprotection (75°C 175W 15s, 90°C 30W 50s). The amino acids used were as follows: fmoc-boc-lysine, fmoc-leucine, fmoc-valine, fmoc-isoleucine, fmoc-alanine, fmoc-glycine, fmoc-boc-glutamic acid, and each one dissolved in dimethylformamide (DMF; CAS# 68-12-2 Sigma-Aldrich). The N-terminus of both Pep-KG and Pep-EG were acetylated with a 20% acetic anhydride (CAS# 108-24-7 Sigma-Aldrich) in DMF solution and the 0.25 mmol N-terminal acetylation method (60°C 50W 30s, 25°C 0W 30s, 60°C 50W 30s, 25°C 0W 30s). Upon completion of synthesis, the resin beads were rinsed with dichloromethane (DCM) then let dry via vacuum filtration for cleavage from solid support which was carried out using a cocktail of 9:0.5:0.3:0.2 ratios of trifluoroacetic acid (TFA CAS# 76-05-1 Chem Impex)/thioanisole (CAS# 100-68-5 Sigma-Aldrich)/1,2-ethanedithiol (CAS# 540-63-6 Sigma-Aldrich)/anisole (CAS# 100-66-3 Sigma-Aldrich), where 10 mL of the cocktail was used in each of two vials and the 0.25 mmol of peptide attached to resin was evenly distributed across the two vials. The resin beads were submerged in the cocktail and were continuously perturbed using an orbital shaker at low intensity for to allow for homogeneous coverage of the beads with the cleaving reagents, and

the reaction vessels were left for 3 hours at room temperature. Upon completion, the beads are filtered from the peptides via gravity filtration immediately into cold (-20°C) ether (CAS# 60-29-7 Fischer Scientific). At this point, the ether should become warm and cloudy due to mass precipitation of the peptides, and the mixture is then spun down at 4000 rcf for 15 minutes at 4°C to improve precipitation. The ether supernatant was discarded, and the gel-like pellet was washed with more cold ether for centrifugation, a process that is repeated twice more. The pellet was stored *in vacuo* pending HPLC purification.

Fmoc-OtBu-E: 204251-24-1, 71989-18-9 Chem-Impex Int'l Inc.

Fmoc-Boc-K: 71989-26-9 Chem-Impex Int'l Inc.

Fmoc-L: 35661-60-0 AnaSpec Inc.

Fmoc-V: 68858-20-8 Chem-Impex Int'l Inc.

Fmoc-I: 71989-23-6 AnaSpec Inc.

Fmoc-A: 35661-39-3 Chem-Impex Int'l Inc.

Fmoc-G: 4530-20-5 AnaSpec Inc.

Purification of (Pep-KG) and (Pep-EG)

Purification of both Pep-KG and Pep-EG was carried out using a Waters 2545 Quaternary Gradient Module (Serial # HO845Q016R) and a Waters 2998 Photodiode Array Detector (Serial # M12998467A) on a C18 reverse-phase HPLC column (XSelect CHS Prep. C18 5µm OBD 19x250 mm Column PN:186005492 SN:162|3801711405). Pep-KG was purified with a gradient from 25% to 35% acidified acetonitrile (MeCN , 0.1% TFA) in acidified HPLC water (with 0.1% TFA).

Pep-KG eluted reliably at 31% MeCN. Pep-EG was purified with a 28%-30% gradient of MeCN in 25 mM TEAA (pH 7) HPLC water. Pep-EG eluted as two peaks at 28.6% and 28.7-28.8% MeCN, but only the second peak was considered pure and collected (SI FIGURES). Purified peptides were rotary evaporated and then freeze dried for preparation of the desalting process.

Desalting of Pep-KG

Pep-KG and Pep-EG were each desalted with Sep-Pak C18 3mL cartridges (Waters Co., Milford, M) by first dissolving the purified peptide in HPLC water with 0.1% TFA added, using 6 mL of acidified water per 0.1 mmol peptide and sonicating. Before adding dissolved peptide, each Sep-Pak cartridge was prepared by adding 3 mL of MeCN + 0.1% TFA twice, followed by 2 additions of 3 mL of water + 0.1% TFA. After this preparatory step, 3 mL of dissolved peptide was added to each cartridge and the filtrate from this peptide was collected and passed through the cartridge again. Next, 3 mL of water with 0.1% TFA was added to the cartridges to remove salts from the peptide. After washing, peptides were eluted off the column with two additions of 3 mL 50% MeCN solution with 0.1% TFA. The eluted peptide was rotary evaporated to remove MeCN, frozen, and the water was lyophilized off.

Preparation of pep-KG/GQ conjugate

Desalted Pep-EG (20 mg) and 33.5 mg N-hydroxysuccinimide (CAS# 6066-82-6 Sigma-Aldrich) were dissolved in minimal amount of DMF (approximately 2 mL). 55.8 mg EDC (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide; CAS# 25952-53-8 Sigma-Aldrich) was added afterward, and the solution were left to stir at room temperature overnight, up to 36 hours. Completion of this reaction generated an imide-peptide species, which registered at 851 Da as confirmed by mass

spectrometry (SI). 100 µL 0.2M $NaHCO_3$ was then added to the solution, and 0.15 mmol propargylamine (CAS# 2450-71-1 Sigma-Aldrich) was added immediately after the pH was raised. This reaction was left to stir for 5 hours, and propargylamine coupling, which generates Ac-E_mLVIIAG-NH_2, was confirmed via mass spectrometry (SI). This crude product was used without further purification for conjugation with 5'-azide modified GQ DNA (SEQUENCE) via azide-alkyne click reaction, which utilized 4 nmol GQ DNA, two times excess of Ac-E_mLVIIAG-NH_2 (4 µL), 4 µL 2M Triethylamine-acetate buffer at pH 7, 4 µL saturated ascorbic acid for activation of 10mM Cu(II)-TBTA in 55% aqueous DMSO (Lumiprobe, MD, USA), which was added last (4 µL). This reaction was run at room temperature for up to 48 hours and was followed by ethanol precipitation of the peptide/DNA conjugate. Due to addition of excess peptide, only peptide/DNA conjugates were be pulled down using this protocol, which involved the addition of 2 mL 2M $MgCl_2$ solution and four times excess 200 proof ethanol. The solution was shaken vigorously prior to resting in an ice bath for one hour. The solution became turbid, and after this incubation period, the sample was spun down at 13,000 rpm for 20 minutes, generating a gel-like pellet. After discarding the supernatant, the pellet was resuspended and washed with 70% ethanol and then spun down once more at 13,000 rpm for 10 minutes, generating a pellet consisting of the peptide/DNA conjugate. The ethanol was discarded, and the sample was stored overnight *in vacuo* awaiting assembly with Pep-KG.

Assembly of Pep-KG with and without templates

All assembled samples were prepared using desalted Pep-KG following HPLC purification. Pep-KG was assembled at a final concentration of 1 mM both in the presence and the absence of polyanionic templates. Templates were added to satisfy a 1:1 charge ratio with Pep-KG, which

contributes a single positive charge. Assemblies were done with the following charge assumptions: the Drew-Dickerson dodecamer single stranded sequence d(CGCGAATTGCGC) contributed a charge of −12, whereas the double stranded counterpart of the same sequence had −24 charges. Both dA3 and trimetaphosphate contributed a −3 charge each. Poly-adenosine templates contain an equivalent number of negative charges to adenosines. A solution of double the concentration of each template, relative to the template's final concentration, was prepared using 40% MeCN, where an equal volume of 2 mM Pep-KG in 40% MeCN was then added to the template solutions that were pre-incubated at 37°C. As an example, an initial solution of 400 µl 0.66 mM trimetaphosphate was prepared in 40% MeCN, where 400 µl 2 mM Pep-KG in 40% MeCN was added to it. All samples were incubated at 37°C throughout the assembly times of up to 2 weeks.

Thioflavin T fluorescence assays

Samples for Thioflavin T fluorescence assays were prepared by combining 74 µL of sample, 100 µL of 40% acetonitrile, and 1 µL of 10 mM Thioflavin T (CAS# 2390-54-7 purchased from Sigma-Aldrich) in 40% acetonitrile in the wells of a 96 well plate (Microplate, 96 well, PS, F-bottom, µCLEAR, black, med. binding, Greiner Bio-one). Thioflavin T fluorescence was read with a BioTek Synergy Mx plate reader (Serial# 250843) every 15 minutes for 24 hours, with short shaking before each read. The excitation wavelength was 444 nm and fluorescence was measured at 484 nm. A well containing 1 µL 10 mM ThT in 40% MeCN and 174 uL 40% MeCN was used as a baseline. The plate was held at 37°C for all 24 hours.

Transmission Electron Microscopy

8 µL of each sample was pipetted onto a carbon-film coated, 200 mesh copper grid (CF200-Cu purchased from Electron Microscopy Services) The sample was then negative stained with 8 µL of the supernatant of a 2% w/v Uranyl acetate in water solution that had been centrifuged at 12,000 rcf for 10 minutes (CAS# 541-09-3 purchased from Electron Microscopy Solutions). Loaded electron microscopy grids were visualized with a Hitachi HT7700 transmission electron microscope at 80 kV.

Circular Dichroism

Circular Dichroism spectra (CD) were recorded on a Jasco-810 Spectropolarimeter. Samples were micro-pipetted onto a 50µL Hellma Analytics quartz cell with a 0.1 mm path length (Model # 106-0.10-40). Spectra were measured by averaging 3 scans from 260-190 nm with a 0.2 nm data pitch and 100 nms^{-1} scanning speed. Molar ellipticity was calculated with the equation $[\theta] = \theta/(10 \times c \times l)$ where c is the peptide concentration in moles/L and l is the pathlength of the cuvette in cm.

CD melting

43 µL of each sample was loaded onto a 50 µL Hellma Analytics quartz cell with 0.1 mm path length (Model # 106-0.10-40) while taking care to minimize bubbles inside the cuvette, and the edge of the cuvette was wrapped in parafilm to minimize evaporation. Each melting trial was conducted using a Jasco J-1500 (Serial # B043361638) Spectropolarimeter starting at 37°C and ending at 92°, changing 1°C every minute and accumulating 3 scans from 260-190 nm every 5°C interval. The data pitch was 0.2 nm, the scanning speed used was 200 nm min^{-1}.

Results and Discussion

The launching point for the investigation of template control of peptide assemblies began with further exploring the electrostatic interactions implicated in previous findings from the lab. Initial work began with using simple salts to “screen” electrostatic interactions between the DNA phosphates and the positive charges on the peptide. The logic was very clear, previous work has demonstrated that condensate formation is sensitive to both simple and complex ions (Saurabh) so varying concentration of salt in the buffer should affect the co-assembly of DNA and peptide. To assess this, Pep-KG was incubated with either single- or double-stranded DNA at 37°C in 40% MeCN buffers that contained varying levels of salt. As with all experiments presented here, the peptide and template were co-assembled in a 1:1 charge ratio so each negative charge on the template corresponded to a positively charged peptide. See methods for detailed description of assemblies.

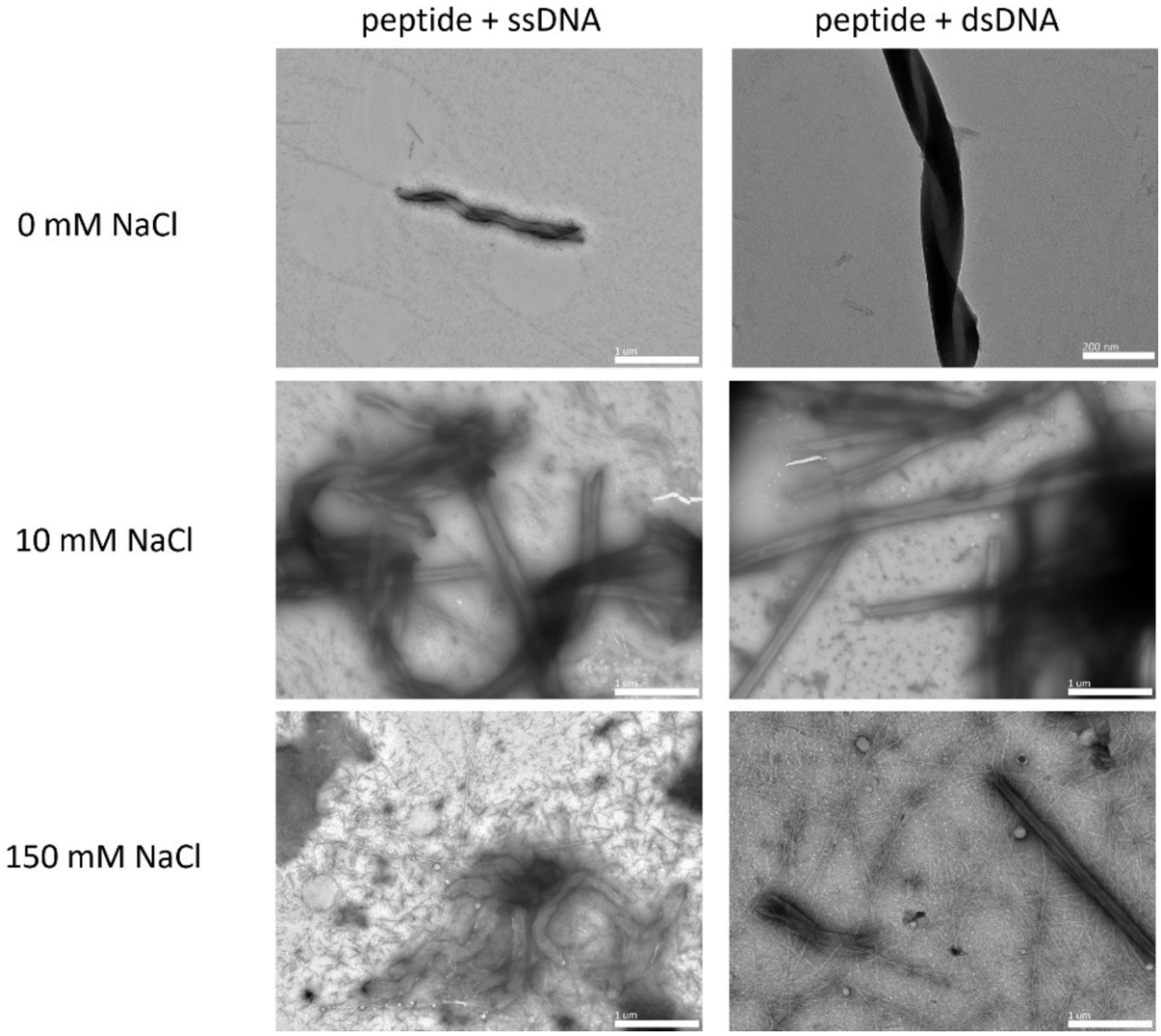

Figure 1: Electrostatic screening by salts interrupts Pep-KG and DNA assemblies. All assemblies were incubated with the Dickerson dodecamer in either single-stranded or double stranded form for 1 week at 37°C in 40% MeCN before imaging via TEM. Figure reproduced from (Kumar Bandela et al. 2022) with permission.

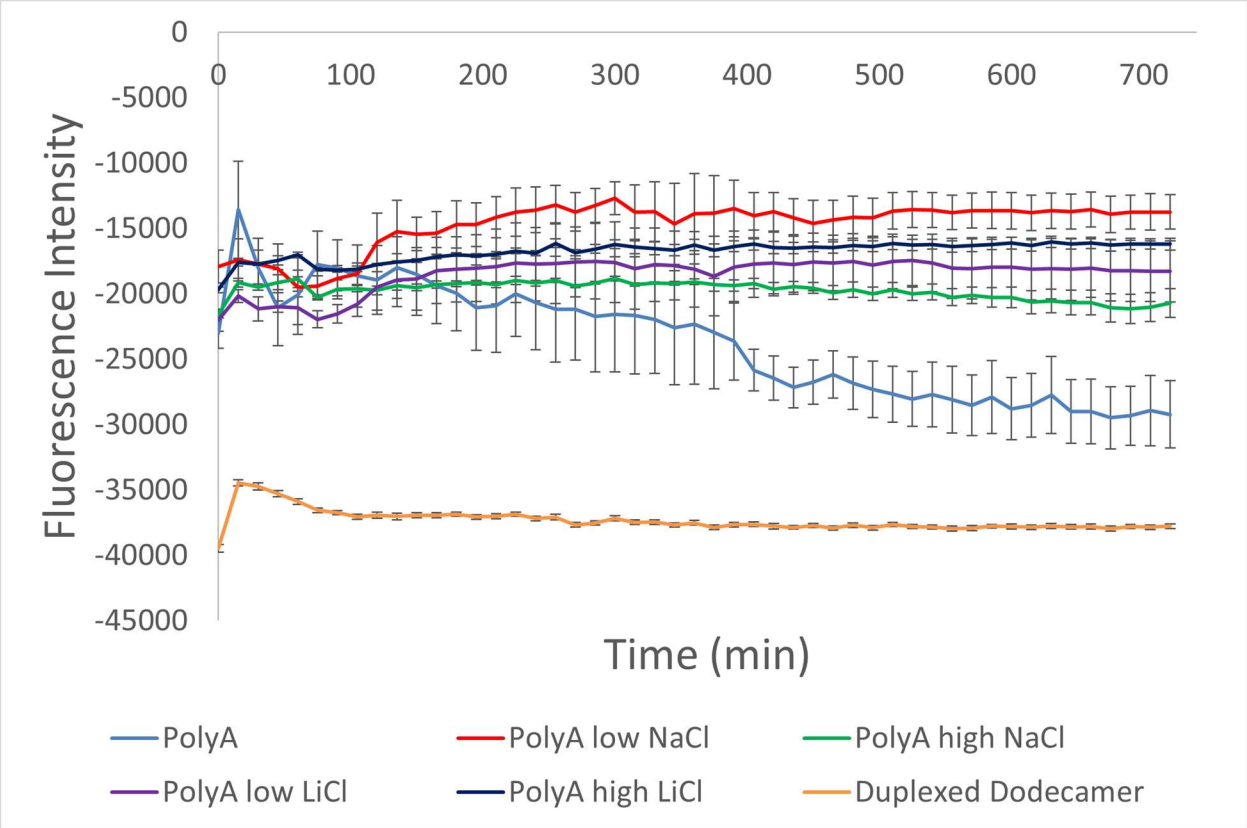

Figure 2: 2-Aminopurine (2AP) fluorescence quenching indicates duplexing of single-stranded DNA upon incubation with Pep-KG as well as supports electrostatic screening of ionic interactions. 0 fluorescence intensity is baseline 2AP fluorescence, negative values indicate quenching. DNA incubated with Pep-KG at 37°C in 40% MeCN. Low salt concentration is 10 mM and high is 150 mM. All error bars are 95% CI values and all samples are n = 3.

TEM imaging revealed that incubation with high concentrations of NaCl morphologically disrupted the assemblies (Fig 1). This disruption to the assemblies was taken as evidence of the importance of electrostatic interactions between the template and the peptide, although given that TEM images were taken after 1 week we were unable to gain insight onto the time

dependence of this. Given that peptide assembly is first passing through a condensate phase before forming paracrystalline structures, what stage of system maturation was the salt affecting? Furthermore, salts are known to significantly affect DNA conformation so more insight into template conformation was desired.

To answer this question, a modified DNA template containing 2-Aminopurine (2AP) was used. 2AP is a fluorescent nucleotide analogue that is quenched upon duplexing. With 2AP incorporated into the middle of the probe, 2AP quenching serves as an indicator of DNA conformation regarding single vs double-stranded conformations. Interestingly, even with poly-adenosine DNA templates which we assumed would be single-stranded, the 2AP exhibited significant quenching (Fig 2, blue line). While the quenching was not to the same degree as the 2AP in the Dickerson dodecamer sequence, which started double-stranded and remained double-stranded so therefore acted as a positive control, it appears that incubation with the Pep-KG caused single-stranded DNA to adopt a double stranded conformation. Also interesting was that various concentrations of salts decreased this quenching effect, so this seems to confirm the effect observed in earlier TEM studies (Fig 1).

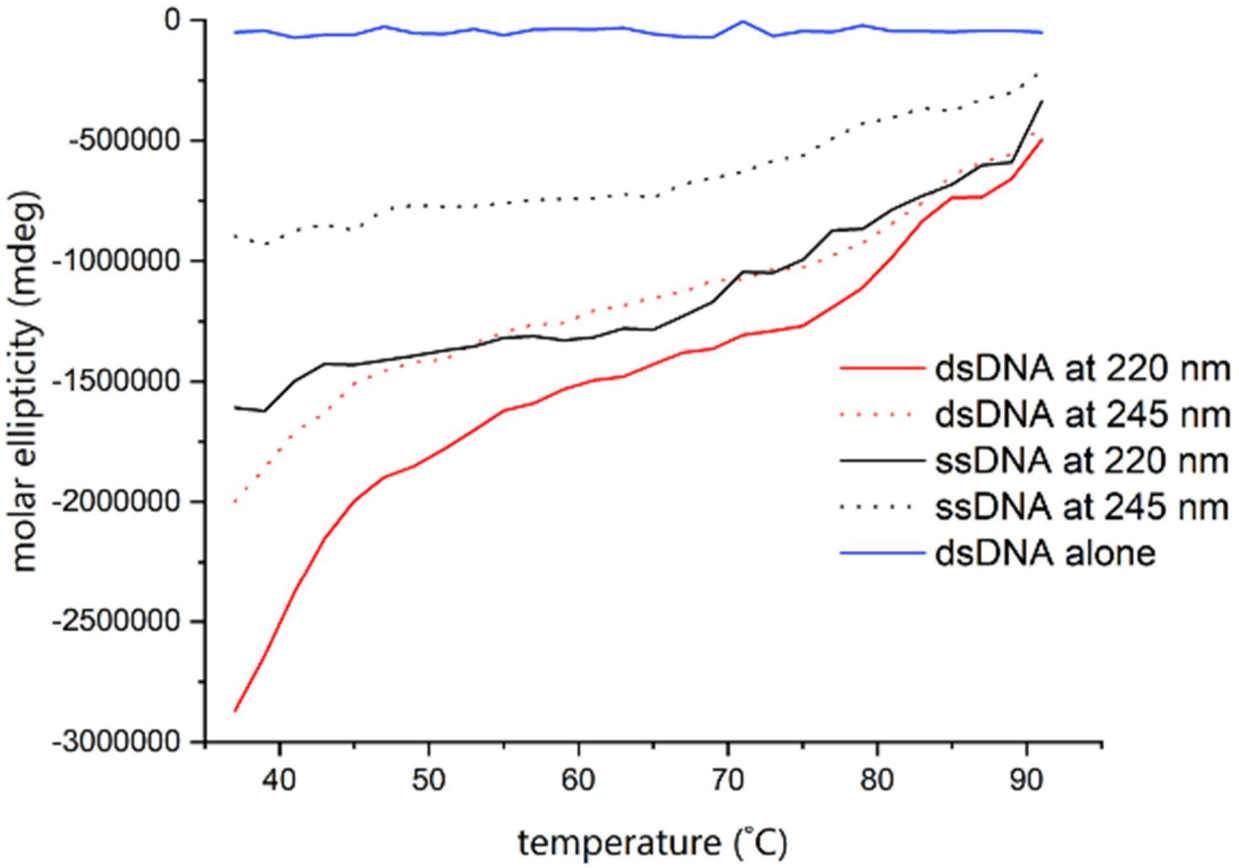

Figure 3: CD melting studies demonstrate increased stability of DNA structure in cross-β peptide assembly. 220 nm corresponds to β-sheet signature, 245 nm corresponds to DNA duplex.

Further investigation of DNA conformation was undertaken with Circular Dichroism (CD) Melting. Cross-β assemblies exhibit a CD minimum at 220 nm and B-form DNA has a strong negative minimum at 245 nm, so by tracking the disappearance of these different minima at different temperatures we could measure the melting of these structures and therefore compare relative thermal stabilities. Peptide assemblies templated with both the double-stranded and single-stranded DNA templates exhibited strong β-sheet signals although the double-stranded template led to a stronger signal than the single-stranded template (Fig 3,

solid lines). Most interestingly, however, was the finding that the Dickerson dodecamer control had already melted at or below 37 °C whereas both of the template DNA strands still showed clear B-form DNA signatures on CD which progressively disappeared at higher temperatures likely due to melting or possibly evaporation from the cuvette (Fig 3). This provides compelling evidence that the peptide assemblies strengthen the DNA structure and provide resistance to thermal denaturation.

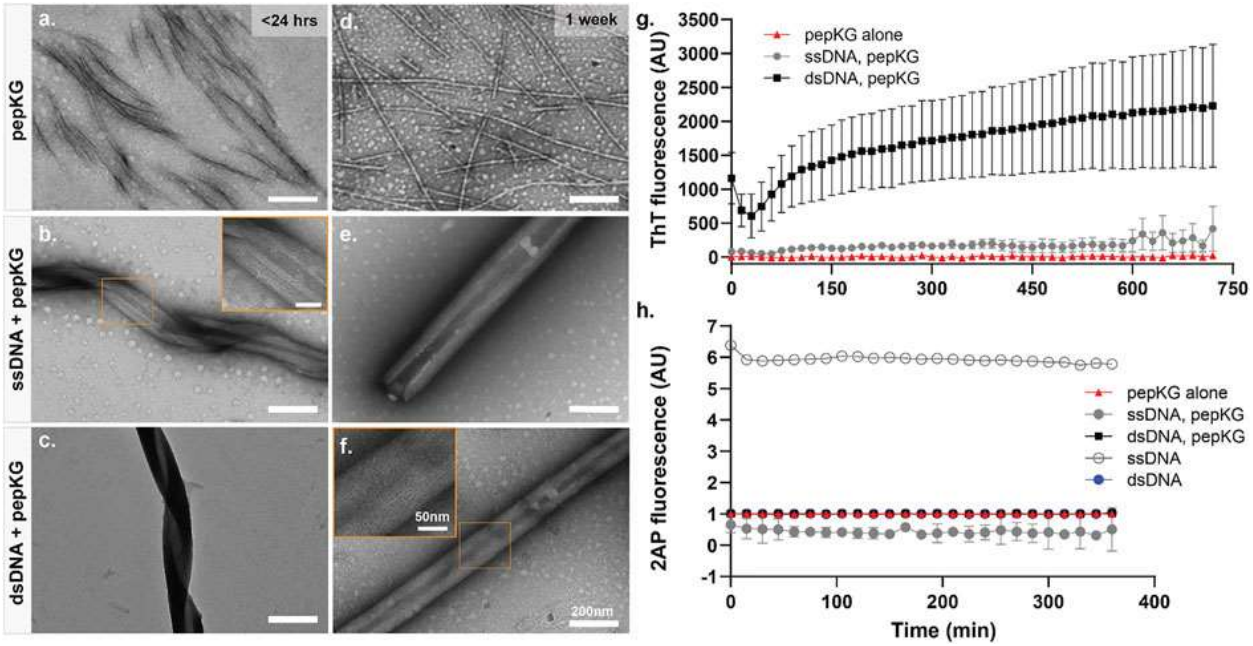

Figure 4: Pep-KG assembly rate is accelerated when templated with dsDNA. TEM image panels (A–C): Pep-KG assembled alone, Pep-KG assembled with ssDNA, and Pep-KG assembled with dsDNA, respectively, within 1 day of initial dissolution. Panels (D–F) were taken after 1 week of incubation. All peptides and DNA templates were dissolved in 40% MeCN at 1 mM Pep-KG concentration and continuously incubated at 37°C. Measurements of ThT fluorescence (G) indicates cross-β growth of the assembly within the first 12 h and (H) 2AP fluorescence within the first 6 h of assembly. Each fluorescence measurement was done at 37°C incubation across

all timepoints. All error bars are 95% CI values and all samples for 2AP and ThT fluorescence are n = 3. Image reproduced from (Gordon-Kim et al. 2022) et al. with permission.

Following this, kinetic experiments were performed to assess whether there was a difference in assembly rate between peptide assemblies templated by single-stranded DNA or double-stranded DNA. TEM was used to image assemblies at multiple times as they formed mature structures. Fig 4a,d demonstrates the maturation of peptide assemblies in the absence of polyanionic templates, transitioning through a condensate phase to heterogeneous fragments within the first 24 hours before eventually transitioning to the characteristic fibers. After 1 week, the mature peptide/nucleic acid co-assemblies both are multilamellar nanotubes (Fig 4e,f) however the ssDNA templated assembly shows greater heterogeneity at earlier time points. Supporting the assertion that there is a difference in assemblies at early time points, there is a significantly higher amount of Thioflavin T fluorescence within the first 12 hours (Fig 4g). Thioflavin T (ThT) is a fluorescent dye that emits upon intercalation with the leaflets of a cross-β assembly, so a higher degree of ThT fluorescence is used to indicate a higher degree of paracrystalline assembly formation. The far higher degree of ThT fluorescence with the double-stranded DNA template within the first hours but the apparent similarities in the overall assembly pathway as assessed by TEM indicate that dsDNA is nucleating the same assembly pathway more rapidly than the single-stranded conformation does. This is a very interesting result because previous work demonstrated sequence-independent templating as well as successful templating interactions for DNA sequences above 6 nucleotides in length (Rha et al. 2020).

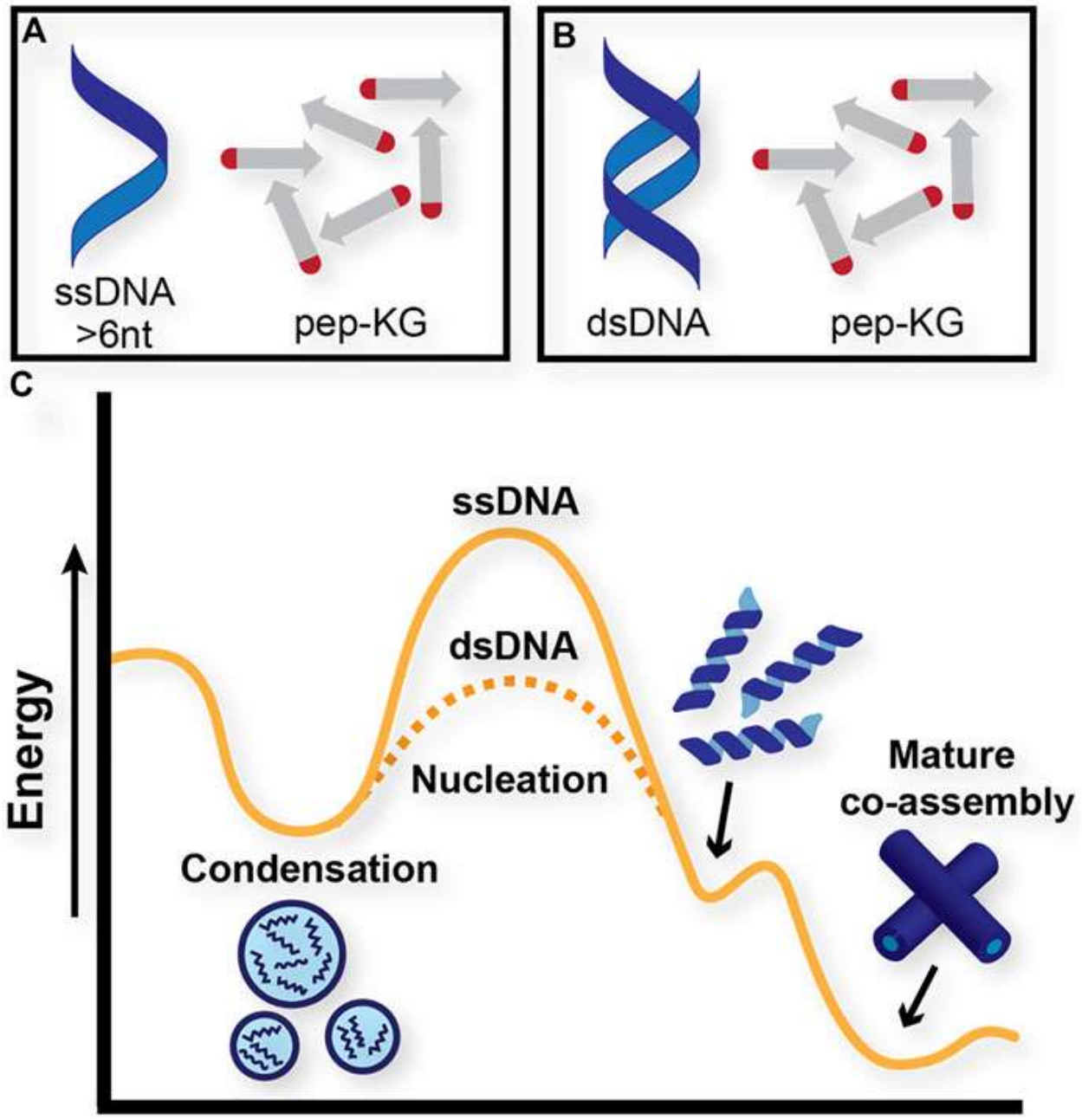

Figure 5: Proposed energy diagram of distinct pathways that Pep-KG may undergo through context-dependent co-assemblies. (A) and (B) denote Pep-KG assembly templated by ssDNA and dsDNA, respectively. Proposed energy diagram (C) illustrates dsDNA templates lower the energetics of nanotube assembly relative to ssDNA templates, as demonstrated by ThT

fluorescence experiments. Image reproduced from (Gordon-Kim et al. 2022) et al. with permission.

The proposed explanation for the results observed in Fig 4 are as follows: the single-stranded and double-stranded DNA templates are progressing on the same assembly pathway as assessed via TEM. However, a significant rate difference is observed in the first hours of the assembly, with dsDNA rapidly accelerating assembly. We attribute this to a more rapid nucleation step. We hypothesize that dsDNA is more effectively recruiting peptides to form a stable nuclei which is then capable of propagating out of the particle phase and forming a mature assembly.

To probe that nucleation step, work was undertaken to find other templates with unusual charge ordering that might avoid the previously understood rule that DNA sequences shorter than 6 nucleobases were incapable of templating assembly. Trimetaphosphate (TMP; Fig 6a) was found to successfully template the formation of multilamellar nanotubes which is in contrast to untemplated assemblies which form fibers (Fig 4d). TMP is a cyclic polyphosphate with an overall negative charge of -3 so it was compared against tri-deoxyadenoside (dA3; Fig 6a) which also has an overall negative charge of -3. Assemblies with TMP form the same nanotube morphology previously observed with longer single-stranded and double-stranded templates (Fig 6c) whereas dA3 leads to heterogenous fibers similar to un-templated Pep-KG assemblies (Fig 6b). These TEM images were taken 1 day after co-assembly so ThT kinetics was done to analyze the first hours after dissolution. Follow-up work with ThT fluorescence also

indicates significant differences in assemblies; there is significantly higher ThT fluorescence in TMP co-assemblies than dA3 co-assemblies at around 5 hours. However, note that for both the TMP and dA3, the lag time appears to be similar. After around 200 minutes of no increase in ThT fluorescence, the ThT fluorescence begins to increase. To understand this in terms of what is happening to the peptide and the template, we believe that for around the first 200 minutes the system is going through a condensate phase so there are no paracrystalline structures for the ThT to stain. After this initial condensate phase, nuclei for the paracrystalline structures are selected within the condensates and these structures begin to propagate; it is these solid aggregate structures that can bind ThT and therefore be measured. The ThT data indicates that both the TMP and dA3 assembly pathways have similar lag times before this liquid to solid transition, so we can surmise that the difference between TMP and dA3 is simply which nuclei can form inside of the condensates. TMP, presumably due to its unique ordering of charges into a more dense ring as opposed to the linear arrangement of dA3, can access or stabilize different nuclei.

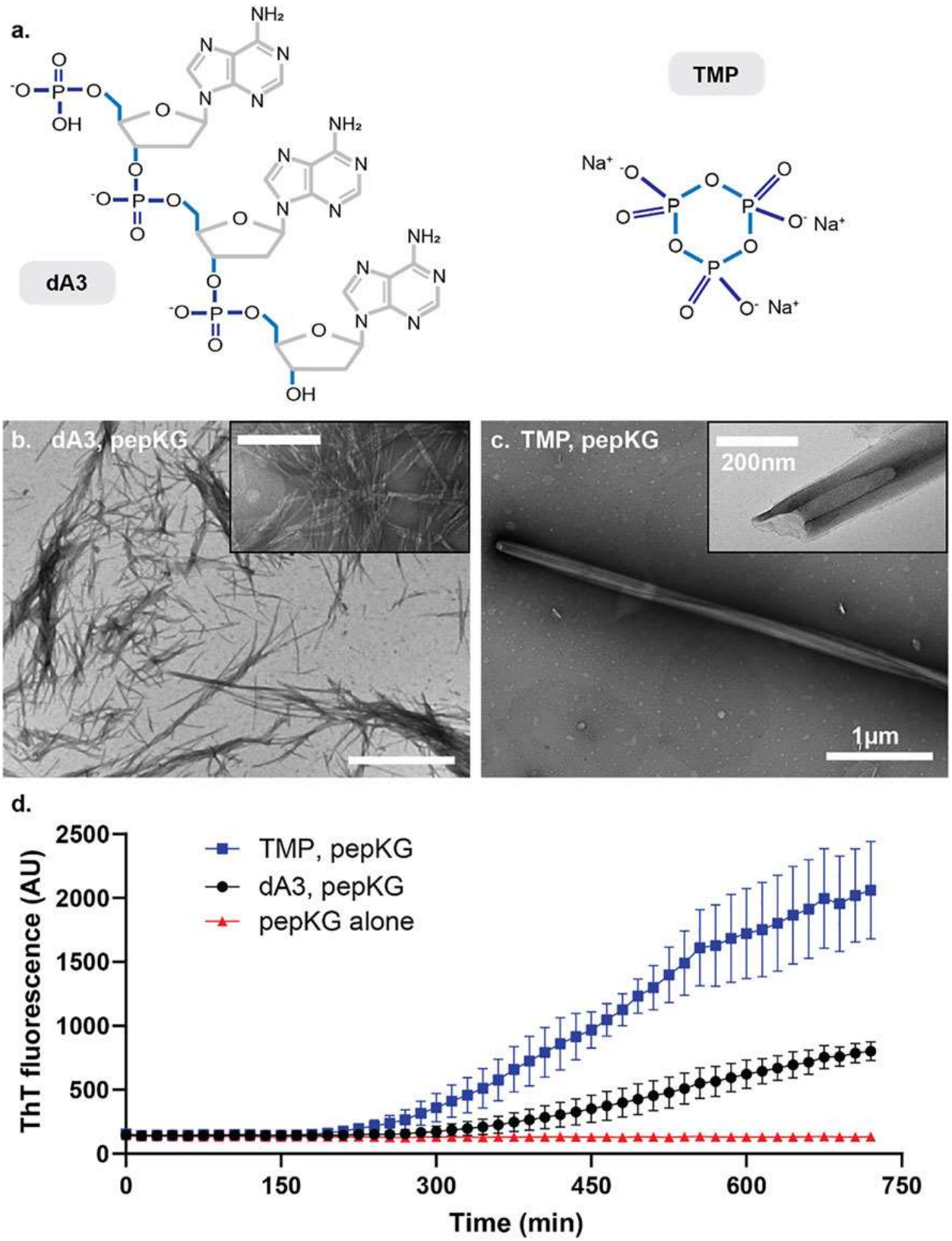
a.
dA3
TMP
b. dA3, pepKG
c. TMP, pepKG
200nm
1µm
d.
TMP, pepKG
dA3, pepKG
pepKG alone
ThT fluorescence (AU)
Time (min)
0
500
1000
1500
2000
2500
0
150
300
450
600
750

Figure 6: (A) Structures of A3 (left) and TMP (right). Micrographs of Pep-KG assembled in the presence of A3 (B) and in the presence of TMP (C). Both images were taken within 1 day of dissolution in 40% MeCN at 37°C. Both samples consisted of 1 mM Pep-KG and 333 µM dsDNA and 333 µM TMP. (D) ThT fluorescence of Pep-KG/dA3 and Pep-KG/TMP samples within the first 12 h. Error bars represent 95% CI values, and all samples are n = 3. Image reproduced from (Gordon-Kim et al. 2022) et al. with permission.

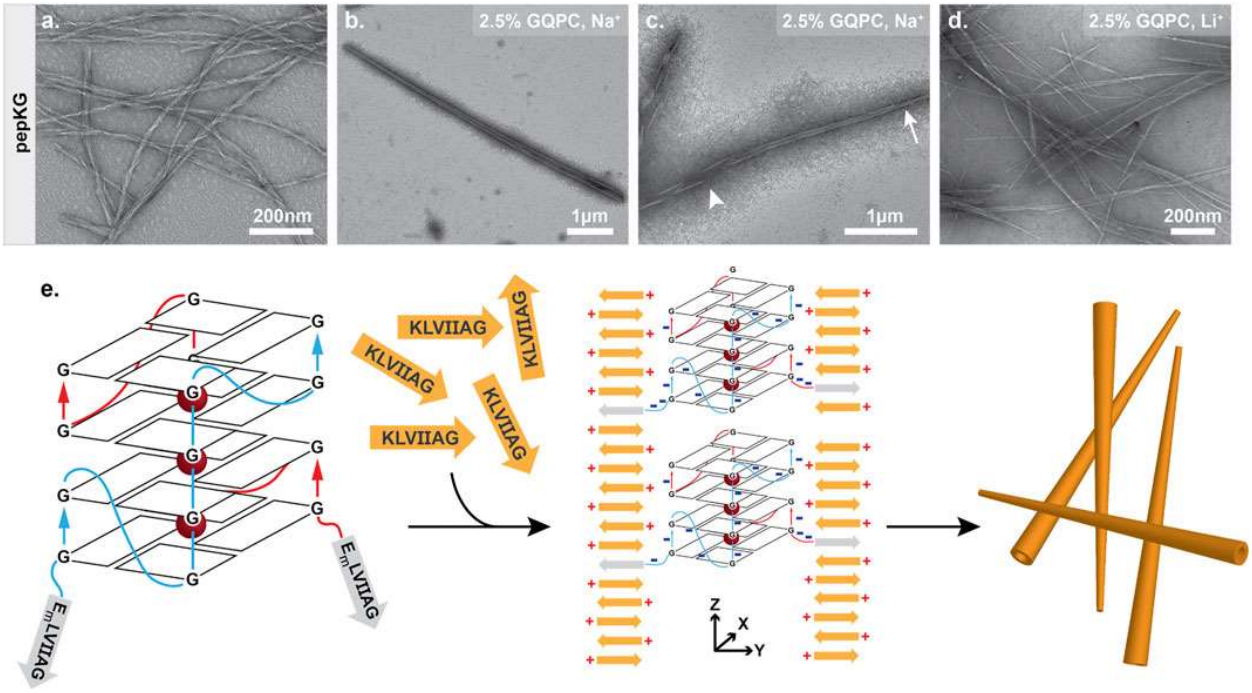

Figure 7: Guanine quadruplex folding is necessary for tapered nanotube morphology. Pep-KG (0.4 mM) assembled in sodium phosphate buffer results in bundled fiber formation (A). Co-assembly of GQPC (2.5% molar equivalent) and pep-KG (0.4 mM) yields tapered conical assemblies resembling nanotubes in sodium phosphate buffer (B). Like previously described nanotubes, GQPC/Pep-KG co-assemblies form ribbons first. The tapered ends of GQPC/pep-KG

co-assemblies close first (tapered end, arrow; wide end, arrowhead) (C). LiPO4 buffer used in place of NaPO4 resulted in a mixture of both non-tapered nanotubes and fibers (D), suggesting that GQ formation is essential for conical tube morphology as depicted in the proposed co-assembly mechanism (E). Figure reproduced from (Gordon-Kim et al. 2022) with permission.

Conclusion

The studies presented here elucidate some of the underlying principles regarding electrostatic control of liquid-to-solid phase transitions in the development from condensate phases to paracrystalline assemblies. Firstly, the sensitivity of the electrostatic interactions was demonstrated by screening the key electrostatic interactions with various concentrations of salts. This led to the novel discovery that the paracrystalline assembly appears to change the conformation of single-stranded DNA templates, which eventually gave rise to the discovery that single-stranded DNA and double-stranded DNA templated assemblies apparently follow the exact same assembly pathway and give rise to identical final morphologies. The key difference between ssDNA and dsDNA was that dsDNA nucleates assembly more effectively, a fact we attribute to the higher density of negative charges. The knowledge that ordering of charges is essential for effective recruitment of peptides into stable nuclei within the condensate phase led us to explore TMP. Studies contrasting TMP with dA3 reveal that ordering of negative charges enables nucleation events that were previously considered impossible. Interestingly this appears to be a matter of nuclei selection within condensates during the liquid-to-solid phase transition.

Altogether, this work sheds more light onto the physical rules regarding the electrostatic interactions governing co-assembly of nucleic acids and peptide or protein materials. Currently, this type of biopolymer mutualism is hypothesized to be crucial to the origin of life. Interactions between RNA and proteins might have been essential for the development of greater biological complexity especially for the evolution of highly conserved cellular machinery like the peptidyl transferase center. The studies presented here should not be understood as recreating

prebiotic evolution, instead they should be used as a model to understand the rules underlying the mechanisms of possible prebiotic evolution. Armed with this understanding, future work can move towards recreating possible prebiotic conditions to demonstrate the evolution of the biopolymer mutualism studied here. Development of a nucleic-acid/peptide network wherein peptides are essential for the replication of the nucleic-acid would be compelling evidence for biopolymer mutualism at the heart of the origin of life and this work would pave the way for work focusing on evolution of information carried on the nucleic acid. Additionally, the discovery that TMP enables an entirely separate assembly pathway due to differences in the ability to form stable nuclei provides an interesting launching point for future work investigating nuclei selection. Future work investigating this could focus on demonstrating competition of different nuclei within condensates and this could either be accomplished with nucleic-acid/peptide models or dynamic covalent networks. There are many unanswered questions in both origins-of-life research and biomolecular condensate research, and it is my sincere hope that work will continue to illuminate these complex and challenging questions and therefore bring us closer to understanding the origin of life.

References

Banani, Salman F., Hyun O. Lee, Anthony A. Hyman, and Michael K. Rosen. 2017. 'Biomolecular condensates: organizers of cellular biochemistry', *Nature Reviews. Molecular Cell Biology*, 18: 285-98.

Benjamin, Dennis R., Carol V. Robinson, Joseph P. Hendrick, F. Ulrich Hartl, and Christopher M. Dobson. 1998. 'Mass spectrometry of ribosomes and ribosomal subunits', *Proceedings of the National Academy of Sciences of the United States of America*, 95: 7391-95.

Bernhardt, Harold S. 2012. 'The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others)a', *Biology Direct*, 7: 23.

Bernhardt, Harold S., and Warren P. Tate. 2012. 'Primordial soup or vinaigrette: did the RNA world evolve at acidic pH?', *Biology Direct*, 7: 4.

Bordenave, Guy. 2003. 'Louis Pasteur (1822–1895)', *Microbes and Infection*, 5: 553-60.

Brangwynne, Clifford P., Christian R. Eckmann, David S. Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A. Hyman. 2009. 'Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation', *Science*, 324: 1729-32.

Carter, Charles W. 2015. 'What RNA World? Why a Peptide/RNA Partnership Merits Renewed Experimental Attention', *Life*, 5: 294-320.

Cech, Thomas R., Arthur J. Zaug, and Paula J. Grabowski. 1981. 'In vitro splicing of the ribosomal RNA precursor of tetrahymena: Involvement of a guanosine nucleotide in the excision of the intervening sequence', *Cell*, 27: 487-96.

Chang, Li-Wei, Tyler K. Lytle, Mithun Radhakrishna, Jason J. Madinya, Jon Vélez, Charles E. Sing, and Sarah L. Perry. 2017. 'Sequence and entropy-based control of complex coacervates', *Nature Communications*, 8: 1273.

Chumachenko, N. V., Y. Novikov, and M. Yarus. 2009. 'Rapid and Simple Ribozymic Aminoacylation Using Three Conserved Nucleotides', *Journal of the American Chemical Society*, 131: 5257-63.

Cleland, Carol E. 2013. 'Pluralism or unity in biology: could microbes hold the secret to life?', *Biology & Philosophy*, 28: 189-204.

Di Giulio, Massimo. 1997. 'On the RNA World: Evidence in Favor of an Early Ribonucleopeptide World', *Journal of Molecular Evolution*, 45: 571-78.

Farias, Sávio T., Thais G. Rêgo, and Marco V. José. 2014. 'Origin and evolution of the Peptidyl Transferase Center from proto-tRNAs', *FEBS Open Bio*, 4: 175-78.

Gao, Ming, and Alexey L. Arkov. 2013. 'Next generation organelles: Structure and role of germ granules in the germline', *Molecular Reproduction and Development*, 80: 610-23.

Gordon-Kim, Christella, Allisandra Rha, George A. Poppitz, Jillian Smith-Carpenter, Regina Luu, Alexis B. Roberson, Russell Conklin, Alexis Blake, and David G. Lynn. 2022. 'Polyanion order controls liquid-to-solid phase transition in peptide/nucleic acid co-assembly', *Frontiers in Molecular Biosciences*, 9.

Gottdenker, Paula. 1979. 'Francesco Redi and the Fly Experiments', *Bulletin of the History of Medicine*, 53: 575-92.

Guo, Qi, Xiangmin Shi, and Xiangting Wang. 2021. 'RNA and liquid-liquid phase separation', *Non-coding RNA Research*, 6: 92-99.

Hsieh, Ming-Chien, David G. Lynn, and Martha A. Grover. 2017. 'Kinetic Model for Two-Step Nucleation of Peptide Assembly', *The Journal of Physical Chemistry. B*, 121: 7401-11.

Hyman, Anthony A., Christoph A. Weber, and Frank Jülicher. 2014. 'Liquid-Liquid Phase Separation in Biology', *Annual Review of Cell and Developmental Biology*, 30: 39-58.

Hyman, Tony, and Cliff Brangwynne. 2012. 'In Retrospect: The Origin of Life', *Nature*, 491: 524-25.

Joyce, Gerald F., and Leslie E. Orgel. 1999. '2 Prospects for Understanding the Origin of the RNA World', *Cold Spring Harbor Monograph Archive*, 37: 49-77.

Kumar Bandela, Anil, Hava Sadihov-Hanoch, Rivka Cohen-Luria, Christella Gordon, Alexis Blake, George Poppitz, David G. Lynn, and Gonen Ashkenasy. 2022. 'The Systems Chemistry of Nucleic-acid-Peptide Networks', *Israel Journal of Chemistry*, 62: e202200030.

Ladouceur, Anne-Marie, Baljyot Singh Parmar, Stefan Biedzinski, James Wall, S. Graydon Tope, David Cohn, Albright Kim, Nicolas Soubry, Rodrigo Reyes-Lamothe, and Stephanie C. Weber. 2020. 'Clusters of bacterial RNA polymerase are biomolecular condensates that assemble through liquid–liquid phase separation', *Proceedings of the National Academy of Sciences*, 117: 18540-49.

Lasker, Keren, Lexy von Diezmann, Xiaofeng Zhou, Daniel G. Ahrens, Thomas H. Mann, W. E. Moerner, and Lucy Shapiro. 2020. 'Selective sequestration of signalling proteins in a membraneless organelle reinforces the spatial regulation of asymmetry in Caulobacter crescentus', *Nature Microbiology*, 5: 418-29.

Lazcano, Antonio. 2012. 'The Biochemical Roots of the RNA World: from Zymonucleic Acid to Ribozymes', *History and Philosophy of the Life Sciences*, 34: 407-23.

———. 2016. 'Alexandr I. Oparin and the Origin of Life: A Historical Reassessment of the Heterotrophic Theory', *Journal of Molecular Evolution*, 83: 214-22.

Le Vay, Kristian, and Hannes Mutschler. 2019. 'The difficult case of an RNA-only origin of life', *Emerging Topics in Life Sciences*, 3: 469-75.

Lincoln, Tracey A., and Gerald F. Joyce. 2009. 'Self-Sustained Replication of an RNA Enzyme', *Science*, 323: 1229-32.

Miller, Stanley L. 1953. 'A Production of Amino Acids Under Possible Primitive Earth Conditions', *Science*, 117: 528-29.

Müller, Felix, Luis Escobar, Felix Xu, Ewa Węgrzyn, Milda Nainytė, Tynchtyk Amatov, Chun-Yin Chan, Alexander Pichler, and Thomas Carell. 2022. 'A prebiotically plausible scenario of an RNA–peptide world', *Nature*, 605: 279-84.

Mushegian, A. R., and E. V. Koonin. 1996. 'A minimal gene set for cellular life derived by comparison of complete bacterial genomes', *Proceedings of the National Academy of Sciences of the United States of America*, 93: 10268-73.

Neveu, Marc, Hyo-Joong Kim, and Steven A. Benner. 2013. 'The "strong" RNA world hypothesis: fifty years old', *Astrobiology*, 13: 391-403.

Paul, Natasha, and Gerald F. Joyce. 2002. 'A self-replicating ligase ribozyme', *Proceedings of the National Academy of Sciences*, 99: 12733-40.

Preiner, Martina, Silke Asche, Sidney Becker, Holly C. Betts, Adrien Boniface, Eloi Camprubi, Kuhan Chandru, Valentina Erastova, Sriram G. Garg, Nozair Khawaja, Gladys Kostyrka, Rainer Machné, Giacomo Moggioli, Kamila B. Muchowska, Sinje Neukirchen, Benedikt Peter, Edith Pichlhöfer, Ádám Radványi, Daniele Rossetto, Annalena Salditt, Nicolas M. Schmelling, Filipa L. Sousa, Fernando D. K. Tria, Dániel Vörös, and Joana C. Xavier. 2020. 'The Future of Origin of Life Research: Bridging Decades-Old Divisions', *Life*, 10: 20.

Prosdocimi, Francisco, Marco V. José, and Sávio Torres de Farias. 2021. 'The Theory of Chemical Symbiosis: A Margulian View for the Emergence of Biological Systems (Origin of Life)', *Acta Biotheoretica*, 69: 67-78.

Rha, Allisandra K., Dibyendu Das, Olga Taran, Yonggang Ke, Anil K. Mehta, and David G. Lynn. 2020. 'Electrostatic Complementarity Drives Amyloid/Nucleic Acid Co-Assembly', *Angewandte Chemie (International Ed. in English)*, 59: 358-63.

Rich, A. 1962. 'On the problems of evolution and biochemical information transfer', *Horizons in Biochemistry*: 103-26.

Ruiz-Mirazo, Kepa, Carlos Briones, and Andrés de la Escosura. 2014. 'Prebiotic Systems Chemistry: New Perspectives for the Origins of Life', *Chemical Reviews*, 114: 285-366.

Saurabh, Saumya, Trisha N. Chong, Camille Bayas, Peter D. Dahlberg, Heather N. Cartwright, W. E. Moerner, and Lucy Shapiro. 2022. 'ATP-responsive biomolecular condensates tune bacterial kinase signaling', *Science Advances*, 8: eabm6570.

Sheehan, Fahmeed, Deborah Sementa, Ankit Jain, Mohit Kumar, Mona Tayarani-Najjaran, Daniela Kroiss, and Rein V. Ulijn. 2021. 'Peptide-Based Supramolecular Systems Chemistry', *Chemical Reviews*, 121: 13869-914.

Sheldon, Robert B. 2005. "Historical development of the distinction between bio- and abiogenesis." In *Astrobiology and Planetary Missions*, 444-56. SPIE.

Szostak, Jack W. 2012. 'The eightfold path to non-enzymatic RNA replication', *Journal of Systems Chemistry*, 3: 2.

Tan, Wei, Sihua Cheng, Yingying Li, Xiao-Yang Li, Ning Lu, Jingxian Sun, Guiyue Tang, Yujiao Yang, Kezhu Cai, Xuefei Li, Xijun Ou, Xiang Gao, Guo-Ping Zhao, W. Seth Childers, and Wei Zhao. 2022. 'Phase separation modulates the assembly and dynamics of a polarity-related scaffold-signaling hub', *Nature Communications*, 13: 7181.

Tomares, Dylan T., Sara Whitlock, Matthew Mann, Emma DiBernardo, and W. Seth Childers. 2022. 'Repurposing Peptide Nanomaterials as Synthetic Biomolecular Condensates in Bacteria', *ACS Synthetic Biology*, 11: 2154-62.

van de Waterbeemd, Michiel, Sem Tamara, Kyle L. Fort, Eugen Damoc, Vojtech Franc, Philipp Bieri, Martin Itten, Alexander Makarov, Nenad Ban, and Albert J. R. Heck. 2018. 'Dissecting ribosomal particles throughout the kingdoms of life using advanced hybrid mass spectrometry methods', *Nature Communications*, 9: 2493.

Wang, Bin, Lei Zhang, Tong Dai, Ziran Qin, Huasong Lu, Long Zhang, and Fangfang Zhou. 2021. 'Liquid–liquid phase separation in human health and diseases', *Signal Transduction and Targeted Therapy*, 6: 1-16.