Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced 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 or dissertation in whole or in part in all forms of media, now or hereafter known, 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 or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____

Li Zhang Date

RE-INVENTING TEMPLATE-DIRECTED REPLICATION IN DYNAMIC

CHEMICAL NETWORKS

By

Li Zhang
Doctor of Philosophy

Chemistry

_____

Dr. David G. Lynn
Advisor

_____

Dr. Khalid Salaita
Committee Member

_____

Dr. Emily Weinert
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____ Date

RE-INVENTNG TEMPLATE-DIRECTED REPLICATION IN DYNAMIC CHEMICAL NETWORKS

By

Li Zhang

B.Eng., Sun Yat-Sen University, China, 2010

Advisor: David G. Lynn, PhD.

An abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Chemistry

2016

Abstract

Re-inventing Template-Directed Replication in Dynamic Chemical Networks

By Li Zhang

Biological replication, either sequence replication defined by the Central Dogma, or prion conformational replication, is driven by a template-directed mechanism. Previous efforts to re-invent templated replication in non-enzymatic systems were unsuccessful partly due to the lack of reversibility. Dynamic chemical networks (DCNs) provide a remarkable platform for selection of replicating species in a dynamic manner, and with adaptive and evolvable potential towards changing environments.

Dynamic chemical networks were constructed using reversible N,O-acetal condensation. Model study establishes dynamic properties, environment impacts and stereochemistry of N,O-acetal condensation reaction, providing insight for our design of networks. DNA templated sequence information transfer in ribose-amine nucleoside polymer (rANP) networks was accomplished, with B-form duplex formed between template and rANP oligomers. Peptide networks based on TTF-CHO building block were generated, with cyclic dimer being selected as major species, and under certain condition self-assembled into novel structures. Prion-like templated conformational replication behavior was achieved in a TTF-CHO + $(TTF)_2$ network, in which both self-templated and external-templated conformation replication were selected and propagated into different assembly structures.

These discoveries achieved selection for replication behavior in dynamic chemical networks by generating diversity on both chemical and physical levels. More importantly, the common reversible linkage that connects both informational biopolymers of biology provides a unique form of nucleic acid/amino acid symbiosis for the emergence of chemical evolution.

RE-INVENTING TEMPLATE-DIRECTED REPLICATION IN DYNAMIC CHEMICAL NETWORKS

By

Li Zhang

B.Eng., Sun Yat-Sen University, China, 2010

Advisor: David G. Lynn, PhD.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Chemistry

2016

Dedicated to My Grandparents Qi and Chunhua, and My Dear Friend Dexin.

Acknowledgements

I would like to express my sincerest and deepest thankfulness to my advisor, Dr. David Lynn, for his mentorship and education over these past years. Dave influences us with his sharp insight into questions, endless curiosity on new subjects, great enthusiasm on teaching science to the public, and dedication to work. He helps me to grow as a chemist with commitment, confidence and creativity, preparing me for my future endeavor as a researcher. His optimistic attitude is a life-long treasure for me to learn.

I would like to thank my committee members, Dr. Khalid Salaita and Dr. Emily Weinert for all the feedback, suggestions and constructive criticisms. I want to thank my collaborators, Dr. Martha Grover and Ming-Chien Hsieh for the kinetics modeling of the networks, which provides essential insights into the system I studied. Thanks to Dr. Shaoxiong Wu and Dr. Bing Wang for their help with NMR experiments and other things.

I want to thank all the Lynn lab members for the help and friendship through my whole graduate study. Every one of you makes the Lynn lab not just a place for scientific research, but a warm space filled with happiness. I especially want to thank the DCN team members, Dr. Jay Goodwin, Dr. Junjun Tan, Dr. Chenrui Chen, Ting Pan, and undergraduate students worked with me, Daniel Bae, Yuka Watanabe and Eric Zhang. It was such a great experience to work with this group of motivated and fun people. I

want to thank Dr. Anil Mehta, Dr. James Simmons and Dr. Dibyendu Das for their help in discussions and experiments. I had great time with my officemates Dr. Yi-Han Lin and Dr. Jill Smith. Many thanks to Dr. Tolulope Omosun, who is always a happy factor and calls me her “brother from another mother”, and to Dr. Sha Li, we had a great Lynn lab journey together. I want to thank Dr. Yue Liu, Dr. Savannah Johnson, Dr. Erin Schuler, Chen Liang, Rolando Rengifo, Noel Li, Allisandra Mowles and Anthony Simentilli, I really enjoyed the time with you and I am looking forward to all of your great discoveries.

I am grateful to Hong Yi and Jeanette Taylor for their help with transmission electron microscopy. Thanks to Dr. John Bacsa for X-ray diffraction, Dr. Fred Strobel, Dr. I-Lin Wu and Dr. Xin Jia for mass spectrometry. Special thanks to Dr. Manshui Zhou for mass analysis on network analysis and Han Zhou for his help with MALDI-TOF.

Thanks to all my other friends, Tao Jiang, Dongmei Xiang, Yang Liu, Hanchao Liu, Naifeng Zhang, Yuxing Wang, Mingda Zhang, Yuling Xie, Le Yan, Claire Yuan, Danny Xu, Cedric Zheng, Jay Kim, Zeding Ding, Han Zhou, and many others. Thank you all for so many unforgettable moments. Special thanks to Dexin Zhou, my best friend and my motivation.

Finally, I feel blessed having my family. Thanks for your unconditional love and support. You are always my rock.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

Aβ	amyloid β
AFM	atomic force microscopy
AIBN	2,2'-azobis(2-methylpropionitrile)
ANP	amine nucleoside polymer
AuNP	gold nanoparticle
Boc	tert-butyloxycarbonyl
CD	circular dichroism
COSY	correlation spectroscopy
DCC	dynamic covalent chemistry
DCL	dynamic combinatorial library
DCM	dichloromethane
DCN	dynamic chemical network
DIEA	N,N-diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
Et_3N	triethylamine
EtOAc	ethyl acetate
FMOC	fluorenylmethyloxycarbonyl
FT-IR	Fourier transform infrared spectroscopy

HBTU	N,N,N′,N′-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HETCOR	heteronuclear correlation
HOBt	hydroxybenzotriazole
HPLC	high performance liquid chromatography
IBX	2-iodoxybenzoic acid
LAH	lithium aluminium hydride
LD	linear dichroism
MALDI-TOF	Matrix-assisted laser desorption/ionization – time of fly
MES	2-(N-morpholino)ethanesulfonic acid
NOE	nuclear Overhauser effect
PNA	peptide nucleic acid
PPh_3	triphenylphosphine
TAPS	N-[tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid
TBAF	tetrabutylammonium fluoride
TBS/TBDMS	tert-butyldimethylsilyl
TEAA	triethylammonium acetate
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
THF	tetrahydrofuran
XRD	X-ray diffraction

Chapter One

Templated Information Replication in Dynamic Chemical Networks

1.1 Template-Directed Replication in Biological Selection and Evolution

Biology is a complex network built upon the dynamic interaction of biopolymers. Nucleic acids and proteins cooperate to continuously function, replicate and mutate to render differentiated species, while environmental fluctuations drive selection and reproduction of specific species. The molecular foundation of this remarkable complex cooperation is best represented as the Central Dogma[1], a process defining how sequential information encoded in DNA template as genotype is transcribed into RNA and translated into proteins as phenotype. The majority of biology relies on this process to accurately pass on genetic information inherent in nucleic acids to offspring, serving as platform for biological replication and Darwinian evolution.

On the other hand, a completely different replication mechanism was recently discovered in infectious protein diseases such as prion infection[2]. For prions, instead of nucleic acid sequences, specific protein conformations are selectively replicated and propagated through a template inherent in the cross-β architecture[3]. Misfolded prion protein can give rise to polymorphic β-sheet aggregates or fibers via a templated nucleation mechanism[4], allowing self-propagation and transfer of specific conformational information. Importantly, this prion-like self-propagation behavior is

not only found in prions or amyloid proteins, but seem to be a generic property of polypeptides. In fact, even very short peptides (3 – 11 residues long) can self-propagate into amyloid fibrils[5].

Both biopolymer replication described by the Central Dogma and prion conformational replication can be viewed as a template-directed replication process from a chemical perspective. In the Central Dogma, linear sequence information is encoded by four nucleotides, and is copied from template molecule following Watson-Crick base-pairing with extremely high precision[6] to give rise to polymerization product. While in prion infectious propagation, cross-β assembly provides the template for self-replication of the β-sheet conformation information and propagate into highly ordered protein assemblies.

However, two templated replication processes are also inherently different in some perspectives, and can be summarized as "digital" and "analog" strategies[7] for replication and evolution, as shown in **Fig 1.1**. For Central Dogma defined replication, highly discrete hydrogen bonding interaction between Watson-Crick base-pairing dominates all other interaction, and resembles "digital" information copying. Followed by nucleobase recognition, nucleotides are polymerized into long sequences by enzymatic phosphodiester condensation. Thus, it is a primarily chemical transformation process, and diversity is generated in the chemical process by forming mismatched sequences, as exemplified by virus replication and evolution in **Fig 1.1A**, exemplified

by virus replication and evolution. While in prion propagation, peptide folding and conformational information is dictated by interaction of complex inter- and intra-molecular forces from amino acid side chains and peptide amide backbones, including electrostatics, aromatic, hydrophobic and hydrogen bonding. Thus cross-β assembly is highly context dependent and subject to environmental changes[8]. In this case, no chemical transformation is involved, diversity and selectively replication is achieved on physical level by self-templated propagation of different protein conformation, similar to an "analog" information response, as shown in **Fig 1.1B**.

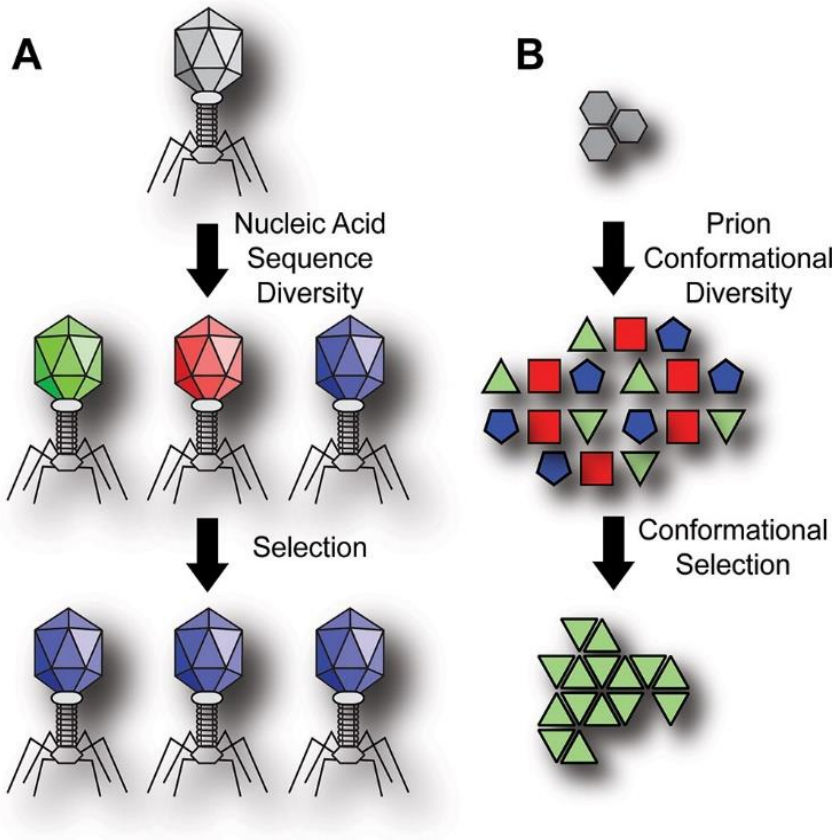


Fig 1.1 Two strategies for replication and evolution: digital strategy (**A**) in nucleic acid sequence dependent replication, exemplified by viral evolution, and analog strategy (**B**) found in prion infectious propagation, which is conformation dependent.

1.2 Re-Inventing Template-Directed Replication

Template-directed biopolymer replication is fundamental to biological replication and evolution. In addition, the broad range diversity of structures and functions in biopolymers clearly demonstrates the power of sequence-controlled polymer synthesis. Inspired by these natural sequence-defined materials, researchers are actively seeking strategies to control the primary structure of synthetic polymers for development of next-generation materials with new structures and functions[9]. However, current efforts to synthesize sequence-controlled polymers fell short in controlling dispersity (for step-growth approaches)[10], sequence precision (for chain-growth strategies)[11] or synthesizing longer sequences (solid-phase iterative synthesis)[12]. In contrast, template-directed replication adopted by biology is highly precise and controlled, with the ability to accurately generate extremely long sequences, superior to any man-invented approaches. Therefore, taking a reductive approach to re-invent template-directed replication process in a simple, enzyme-free context has long been an attractive topic for researchers working in diverse fields from biological evolution[13] to material science.

The well-defined nucleic acids were first used as templates to replicate digital sequential information. Inspired by Orgel's pioneering works[14], the first successful non-enzymatic DNA templated ligation was reported in 1986 by von Kiedrowski[15]. As shown in **Fig 1.2**, ligation of two trinucleotides **A** and **B** is catalyzed by a complementary hexamer DNA template **T**. 3'-phosphate of **A** is activated by 1-

(dimethlyaminopropyl)-3-ethylcarbodiimide (CDI) to overcome the kinetic barrier for phosphodiester condensation without enzymes. Ligation with a fully complementary template gives highest efficiency, strongly suggesting the involvement of template effect[16].

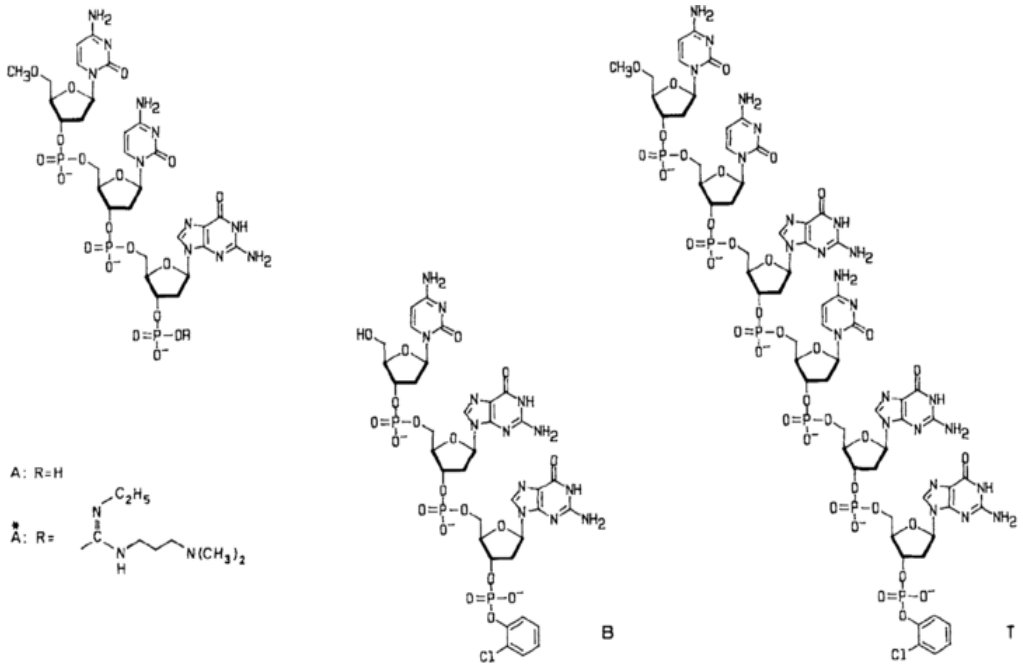


Fig 1.2 Non-enzymatic DNA template directed ligation[15]. Template **T** catalyzes the ligation of **A** and **B**. The product has the same structure as **T**, and serve as new template.

Following Orgel and von Kiedrowski's work, many other DNA templated replication systems have been explored. Evolution and origin of life scientists have developed templated-replication based on triplex recognition[17]; replication coupled with protocell division[19] and replication into alternative nucleic acid backbones such as GNA[20] were also reported. At the same time, material scientists extend Watson-Crick base-pairing into template assisted polymer synthesis [21] as well as synthesis of

structures beyond nucleic acids[22].

On the other hand, rules governing template-directed cross-β peptide conformational replication have been intensively investigated, due to its correlation with many diseases and functional biomaterials. This replication process is primarily a physical transformation, and diversity is rich on conformational level but not on primary sequence level. Recently, researchers started to couple the physical transformation (cross-β assembly) and chemical transformation (peptide ligation reaction) together, potentially provides opportunity to further enrich diversity of the system by introducing another dimension of variation into peptide cross-β assembly. Takahashi and Mihara developed a system of template-directed self-replicating amyloid fibrils[23], where the surface of β-sheet serves as template to direct the ligation of shorter peptide fragments into full-length peptide, and subsequently incorporated into cross-β assembly to become new templates. Ashkenasy[24] reported a similar cross-β replication system that assembles into both fibers and nanotubes, as shown in **Fig 1.3**. In this system, template-assisted formation of ligation products exhibits exponential growth. Both systems require Cys residue at N-terminus of ligation site, and activation of C-terminus with thioester for native ligation reaction to take place.

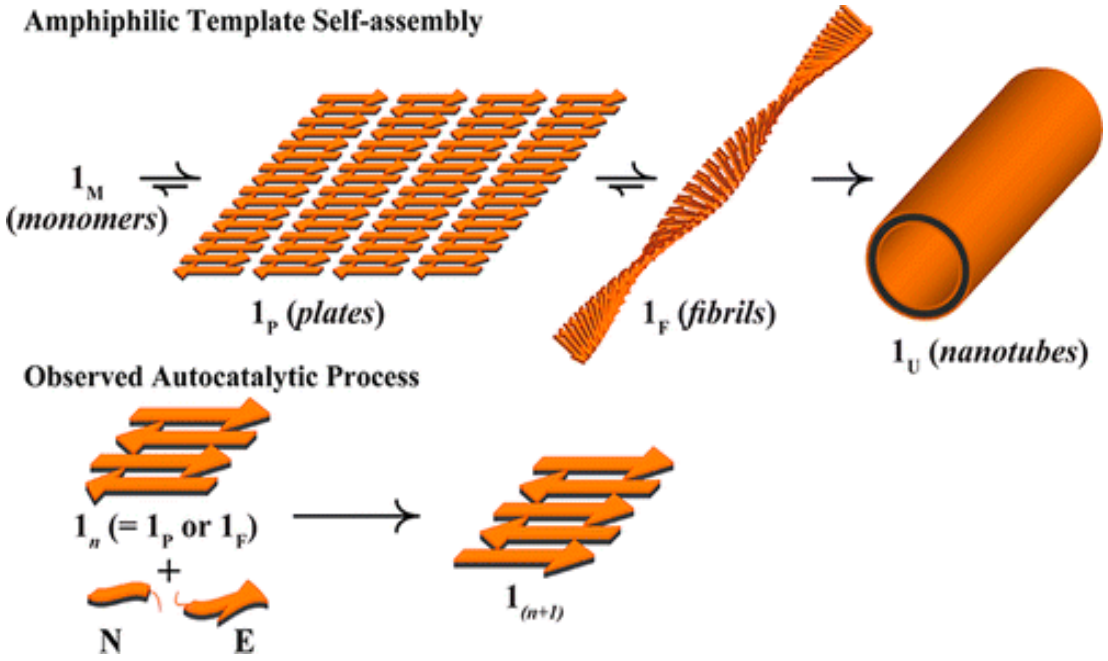


Fig 1.3 Cross-β template directed replication. Cross-β assembly serves as template to direct the ligation of **N** and **E**. The ligation product then incorporated onto the surface of cross-β assembly and serve as new template. The sheets assemble into fibers and nanotubes.

Examples discussed above nicely demonstrate first step success for implementing template-directed replication criteria in chemical systems. However, significant limitations including fidelity and sequence space are present for both systems, impeding an efficient re-creation of replication. For nucleic acids non-enzymatic replication, efficiency of monomer polymerization is considerably lower compared to ligation of oligomers. Enzyme-free templated copying of monomeric nucleotide tends to give high error rate (~20%)[25], low sequence specificity and distribution of truncated products[26], as shown in **Fig 1.4**. Because maximum genome information can be encoded is inversely proportional to the mutation rate[27], a 20% error rate corresponds to a genome of only 5 nucleotides, hardly large enough to carry any sensible function. Some efforts

have been made[28] to improve replication fidelity, but in general DNA templated polymerization of monomers remains to be an error-prone procedure. While for the cross-β templated conformational replication, sequence fidelity is not yet even explored, but the use of native ligation reaction imposes significant limitation on the option of amino acid residue at ligation site.

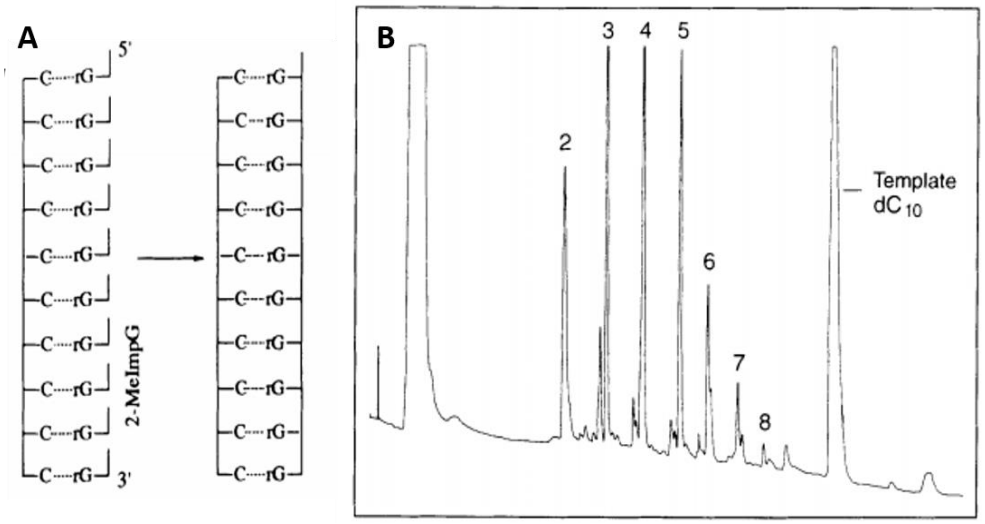


Fig 1.4 DNA template catalyzed RNA monomer polymerization[26]. (**A**) A $(dC)_{10}$ template is used to direct polymerization of activated RNA monomer 2-MeImpG. (**B**) Templated reaction generates distribution of truncated products instead of full-length product

1.3 Dynamic Chemical Networks (DCNs) and Template Effect

With the problems for template-directed replication in chemical systems in mind, next questions become: can template-directed replication be achieved in a different way? Is fidelity only determined by template sequences? How do other factors play role in high-fidelity replication?

Aside from sequence-specific recognition driven by Watson-Crick base-pairing, fidelity

of enzymatic template copying also benefits from reversibility of phosphodiester linkage in the form of polymerases error-correction function. Phosphodiester linkages are extremely stable under physiological condition, but the presence of enzymes can make it temporarily reversible, so that mismatched nucleotides can be removed and replaced. However, in non-enzymatic templated replication, such error correction mechanism is missing. Once the phosphodiester linkage is formed, it remains stable and cannot be hydrolyzed easily. As a result, matched base-pair, even though thermodynamically more preferred, cannot outcompete and replace already-formed mismatched products. In other words, introducing reversibility is one plausible approach to improve templated replication fidelity.

The surging research field of dynamic chemical networks (DCNs) provides opportunity to implement reversibility into templated replication systems. Around two decades after von Kiedrowski's report on first success of DNA-templated ligation, study on dynamic chemical networks (DCNs) became a booming topic in system chemistry. Taking advantage of reversible covalent reactions called dynamic covalent chemistry (DCC)[29], DCNs create complex systems from simple building blocks under thermodynamic control. The members within the system constantly interconvert through formation and breakage of reversible covalent bonds, and the distribution of members is responsive to change in environment conditions and template effect. Such complexity generated by dynamic covalent linkages are previously coined as "dynamic combinatorial library" (DCL), but we instead use the term "dynamic chemical network" (DCN) to better

reflect its interchangeable nature.

With its dynamic property, DCNs offer a unique opportunity to help solve the fidelity problem in templated replication that has haunted researchers for decades. Equilibrium of DCNs can be shifted by template effect and select for species with template-binding structures[29b, 29e], similar to template effect in DNA-templated replication. Actually, DCN was initially envisaged as a potential tool for discovery of new ligands and receptors based upon template-binding. Upon addition of external target molecules as templates, the dynamic nature of DCNs allows equilibrium shifts to amplification of members with highest template-binding affinity. DCN selection based on recognition with external template can generally be categorized into two strategies[29b, 30]: casting (for selection of a guest by a template as host) and molding (for selection of a host by a template as guest), as shown in **Fig 1.5**. External template induced network selection have been achieved for wide range of templates including metal ions[31], small molecules[32] and biomolecules[33].

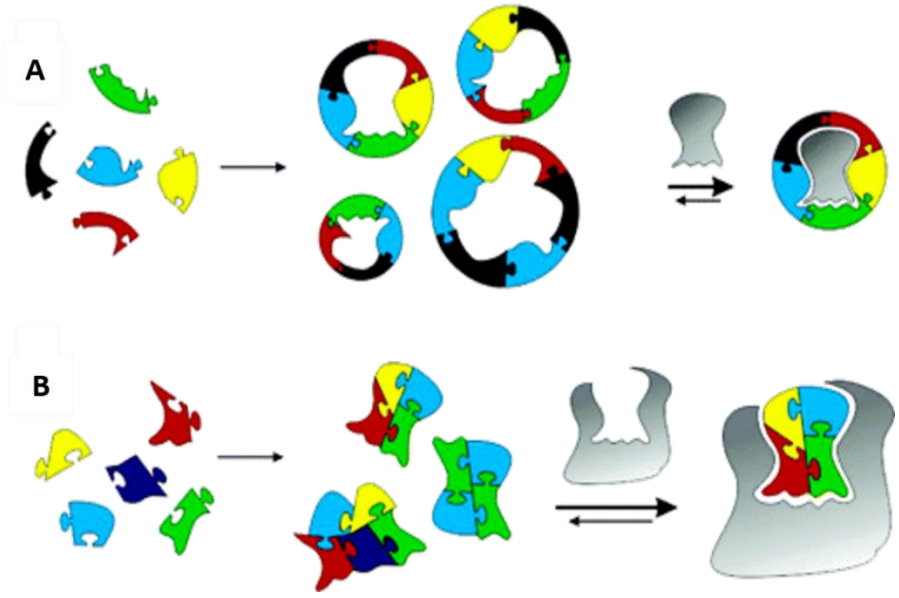


Fig 1.5 Two strategies of DCN member selection based on external template effect: (**A**) Casting, for selection of guests; and (**B**) Molding, for selection of hosts.

Besides external templateing, DCNs can also selectively overexpress members by self-templating, similar to self-templated cross-β propagation. In self-templating DCNs, intermolecular recognition occurs between network members, resulting in self-organization and driving the equilibrium towards self-copying species. **Fig 1.6** illustrates a simple and elegant example of DCN selection based on self-templating effect[34]. This DCN constitutes imine condensation products $\mathbf{Al_{(1\text{-}3)}Am_{(1,2)}}$. Among all possible network members, only $\mathbf{Al_1Am_1}$ is capable stabilizing itself by self-templating effect through hydrogen-bonding induced dimerization to form $\mathbf{[Al_1Am_1]_2}$. This self-duplicating interaction biases DCN equilibrium towards amplification of $\mathbf{Al_1Am_1}$ as the dominant species in the network.

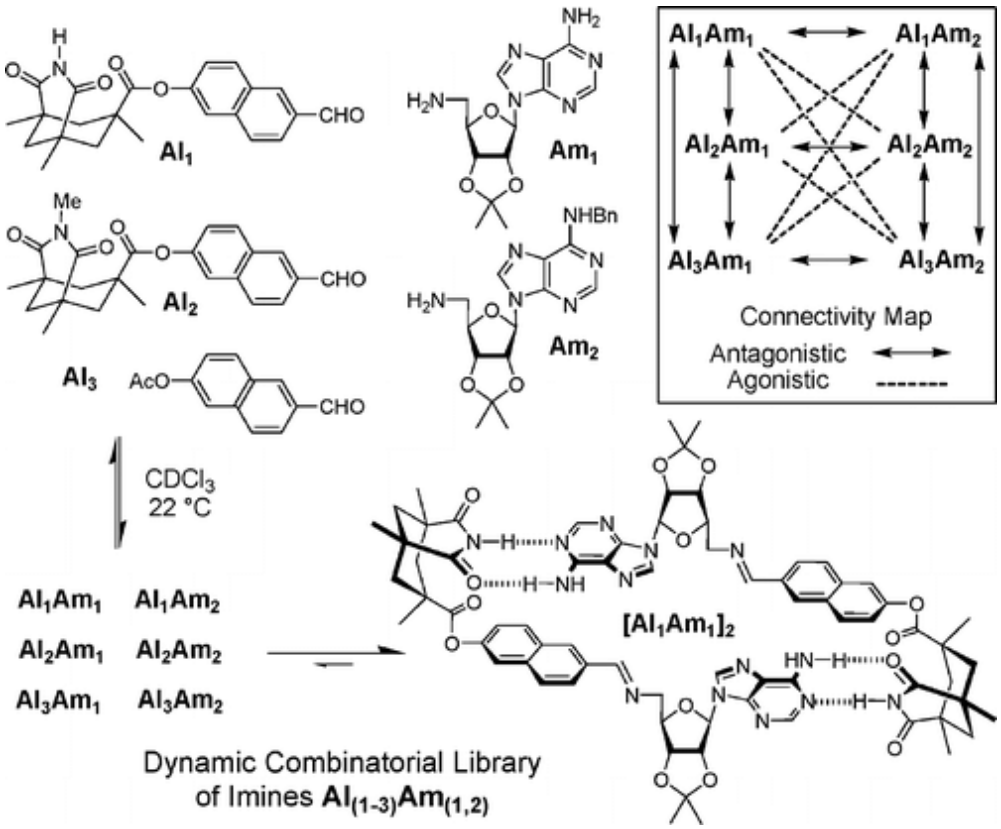


Fig 1.6 Internal Template induced DCN selection[34]. Pairwise combination of aldehydes $\mathbf{Al_{(1-3)}}$ and amines $\mathbf{Am_{(1,2)}}$ generates a DCN containing six imines $\mathbf{Al_{(1-3)}Am_{(1,2)}}$. Only $\mathbf{Al_1Am_1}$ is capable to dimerize, resulting in greater stability and selection of $\mathbf{Al_1Am_1}$.

1.4 Re-Invent Template-Directed Replication in DCNs

In principle, external templating and self-templating selection in DCNs resemble DNA template directed replication and prion self-templated conformational replication respectively. In addition, dynamic property of DCNs allows for selection under thermodynamic control, which provides error-correction in a dynamic fashion, and can potentially enhance fidelity of replication. Therefore, introducing template effect into biopolymer replication systems could help improving fidelity. What's more, the wide scope of reversible linkages to construct DCNs greatly enriches the chemical and sequence space in templated replication.

Our lab reported the first attempt to construct a simple DCN by using DNA template to guide imine condensation of complementary strands[35]. As shown in **Fig 1.7**, in the presence of DNA template dGCAACG, two modified DNA trimers, 5'-H_2N-dTGC and dCGT-3'-CH_2CHO undergoes imine coupling to yield the product with complementary sequence to the template. Dynamic imine condensation allows thermodynamic equilibrium between coupled and uncoupled substrates, allowing for substrate-template recognition to guide product formation. In aqueous environments, products complementary to the template outcompetes non-complementary product by 30-fold at 0^oC (thermodynamic cycle). Since the imine group is highly labile in water and hydrolyzes quickly upon dissociation from template, cyanoborohydride is used to reduce the product into stable secondary amine (catalytic cycle).

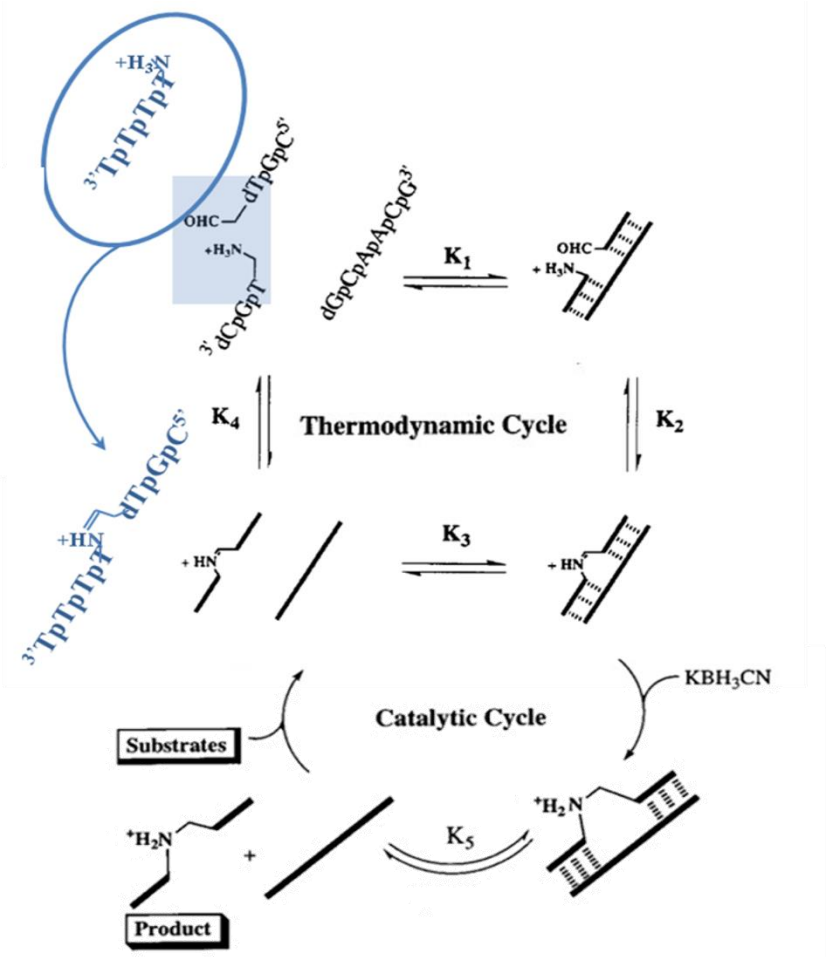


Fig 1.7 Coupled thermodynamic and kinetic cycles of DNA template directed ligation[35].

Later, our lab extended the DNA template guided imine DCN to bis-functionalized substrate 5'-H_2N-dT-3'-CH_2CHO[36], where a $(dA)_8$ template guided the monomeric substrate to polymerize into complementary $(dT)_8$ amine nucleoside polymer (ANP) (**Fig 1.8**). This templated DCN is highly robust, with significantly improved fidelity (< 1%), high sequence specificity, and capability to copy information from longer DNA template such as 32-mer[36b], established that introducing reversibility improves fidelity in templated replication. The reaction is found to proceed with a new step-growth polymerization mechanism, distinct from chain-growth mechanism observed in

enzymatic DNA replication. This templated imine condensation approach was further-extended for templated synthesis of modified PNA[37] and base-filling PNA synthesis[38] by other research groups.

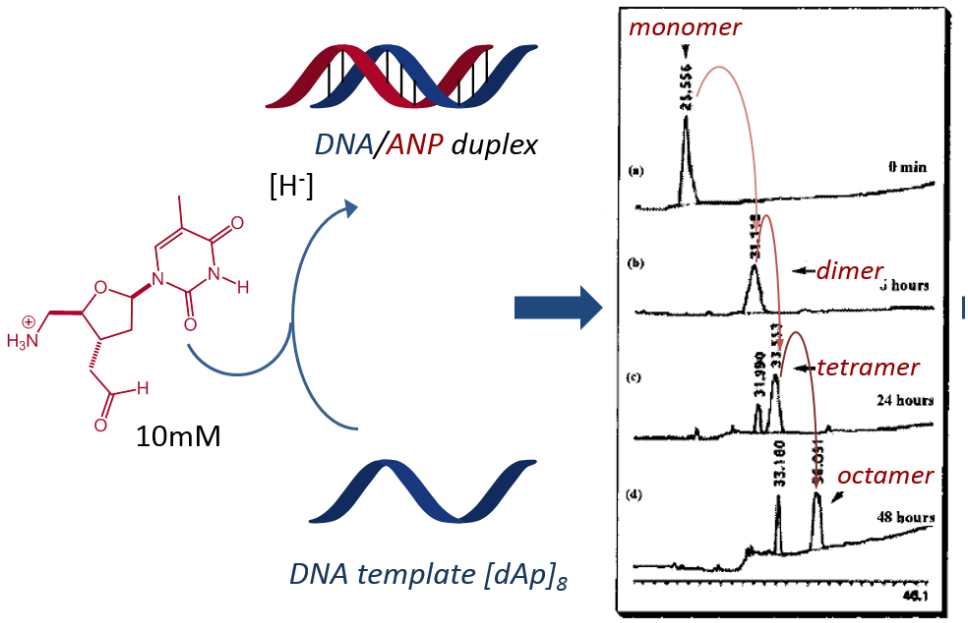


Fig 1.8 Template-directed polymerization of amine nucleoside polymer[36]: ANP monomers association with complementary template gives full-length product, resulting in sequence and chain-length specificity following step-growth mechanism.

Self-templated peptide conformational replication coupled with reversible linkage formation were also reported[39]. Examples include a simple peptide DCNs constructed from disulfide-exchange selecting for β-turn hairpin structure[40], and a dynamic peptide network aggregates into particles[39c]. But in general this topic is still underexplored. Otto and coworkers[41] demonstrate an interesting and unusual system in selection of self-assembly peptide in DCNs. They demonstrate a DCN of dithiol-functionalized

pentapeptide building blocks forming a mixture of macrocyclic oligomers through disulfide exchange, among which hexamer and heptamer can self-propagate into β-sheets and grow as fibers, as shown in **Fig 1.9**. Remarkably, breakage of fibers by different type of mechanically forces select for propagation of different macrocycles: shaking the solution lead to exponential growth of hexamer-based fibers, while stirring favors amplification of heptamers. In such process, product selection is determined kinetically, even though the system is under thermodynamic control. This system can be regarded as a first step towards far-from-equilibrium selection.

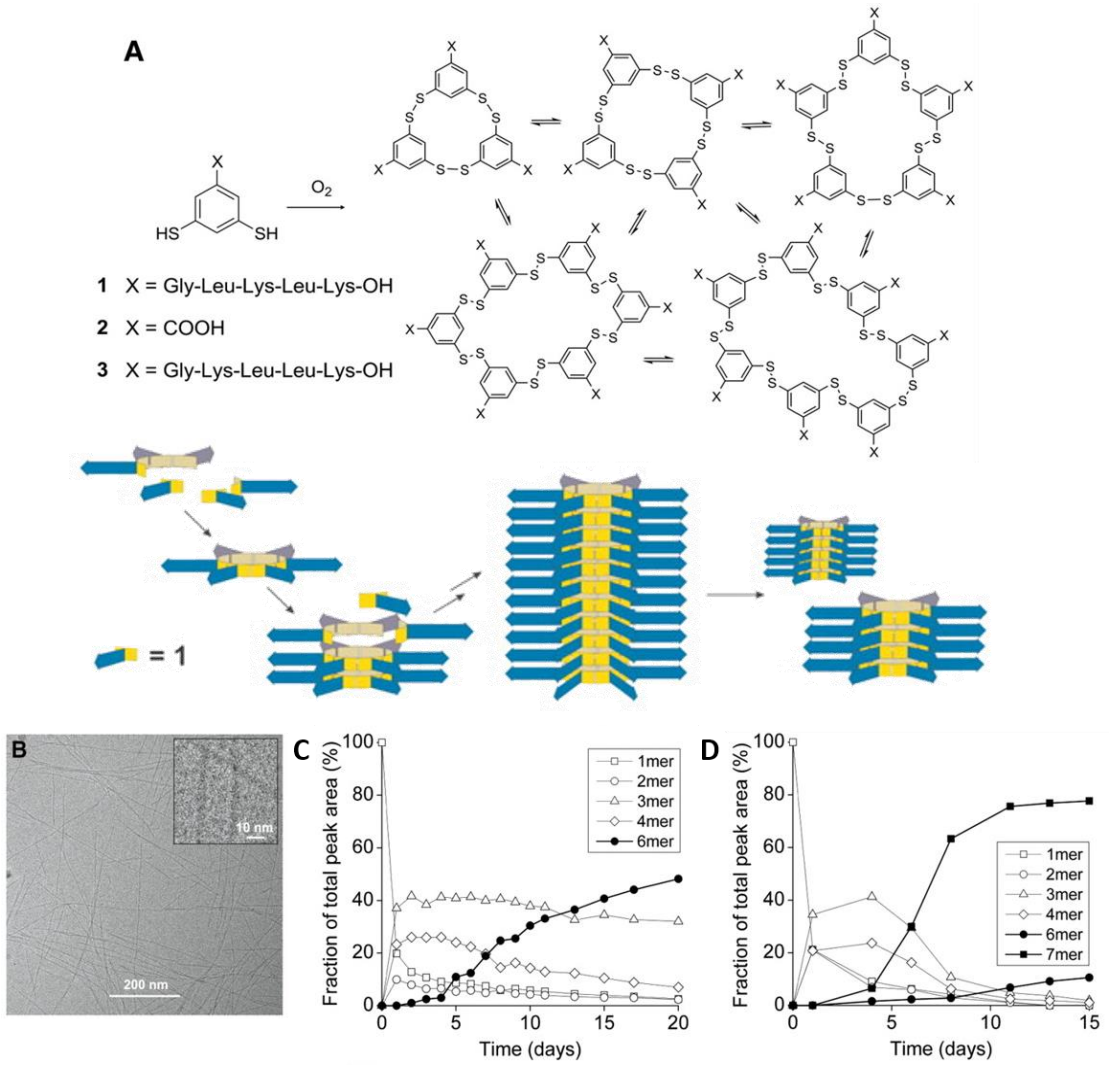


Fig 1.9 A peptide macrocycle DCN selecting for self-assembly species responds differently on different mechanic forces. (**A**) A dithiol-functionalized peptide forms DCN containing different macrocycles through disulfide exchange. (**B**) Hexamers or heptamers can self-assemble into fibers. (**C**) Hexamers are amplified by shaking-induced fiber breakage. (**D**) Heptamers are amplified by stirring.

Benefiting from thermodynamic control in DCNs, DNA template directed ANP replication achieved unprecedented fidelity compared to non-dynamic systems. However, labile imine linkage hydrolyzes rapidly following dissociation from the template, making product separation and analysis infeasible. Reductive amination was

used to overcome this problem, but it falls short in maintaining dynamic nature of the network. On the other hand, peptide replication in DCNs is still rarely studied, and limited reports on such systems are restricted by both sequence diversity and choice of dynamic linkages. For example, in Otto's peptide replication DCN, species diversity is restricted to macrocycles size but not peptide sequence, and choice of dynamic linkage is so far limited to disulfide exchange reaction.

Templated replication in biology occurs in fully dynamic biopolymer networks and generates diversity on both sequence and conformation levels. To re-construct such process in chemical system, it is desirable to create truly dynamic DCNs instead of irreversibly trapped networks, only so that species within the network can be fully responsive towards different template, and select for different sequences.

In this dissertation, I seek to construct DCNs using N,O-acetal linkage, a dynamic linkage with similar condensation mechanism of imine but higher stability. With N,O-acetal linked DCNs, I hope to re-invent both digital and analog template-directed replication in a dynamic fashion. For DNA templated digital replication, N,O-acetal linked dynamic polymer with complementary sequence can be selected. While for peptide cross-β self-templated replication, taking advantage of amino acid side chains to construct dynamic linkages allows for exploring sequence diversity in such replication process, and establishing generality of dynamic linkages in aid of peptide propagation. The establishment of DCNs exhibiting replication of digital and analog

genetic materials sets the first stage to explore chemical evolution and bottom-up synthetic biology.

References

1. Crick, F. H., On protein synthesis. *Symp Soc Exp Biol* **1958,** *12*, 138-163.

2. Li, J.; Browning, S.; Mahal, S. P.; Oelschlegel, A. M.; Weissmann, C., Darwinian Evolution of Prions in Cell Culture. *Science* **2010,** *327* (5967), 869-872.

3. Geddes, A. J.; Parker, K. D.; Atkins, E. D. T.; Beighton, E., "Cross-β" conformation in proteins. *Journal of Molecular Biology* **1968,** *32* (2), 343-358.

4. (a) Chiti, F.; Dobson, C. M., Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry* **2006,** *75* (1), 333-366; (b) Zhang, J.; Muthukumar, M., Simulations of nucleation and elongation of amyloid fibrils. *The Journal of Chemical Physics* **2009,** *130* (3), 035102.

5. (a) Goux, W. J.; Kopplin, L.; Nguyen, A. D.; Leak, K.; Rutkofsky, M.; Shanmuganandam, V. D.; Sharma, D.; Inouye, H.; Kirschner, D. A., The Formation of Straight and Twisted Filaments from Short Tau Peptides. *Journal of Biological Chemistry* **2004,** *279* (26), 26868-26875; (b) Liang, Y.; Pingali, S. V.; Jogalekar, A. S.; Snyder, J. P.; Thiyagarajan, P.; Lynn, D. G., Cross-Strand Pairing and Amyloid Assembly. *Biochemistry* **2008,** *47* (38), 10018-10026; (c) Madine, J.; Copland, A.; Serpell, L. C.; Middleton, D. A., Cross-β Spine Architecture of Fibrils Formed by the

Amyloidogenic Segment NFGSVQFV of Medin from Solid-State NMR and X-ray Fiber Diffraction Measurements. *Biochemistry* **2009,** *48* (14), 3089-3099.

6. (a) Tamm, C.; Strazewski, P., Replication Experiments with Nucleotide Base Analogues. *Angewandte Chemie International Edition in English* **1990,** *29* (1), 36-57; (b) Kuchta, R. D.; Benkovic, P.; Benkovic, S. J., Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. *Biochemistry* **1988,** *27* (18), 6716-6725.

7. Goodwin, J. T.; Mehta, A. K.; Lynn, D. G., Digital and Analog Chemical Evolution. *Accounts of Chemical Research* **2012,** *45* (12), 2189-2199.

8. Childers, W. S.; Anthony, N. R.; Mehta, A. K.; Berland, K. M.; Lynn, D. G., Phase Networks of Cross-β Peptide Assemblies. *Langmuir* **2012,** *28* (15), 6386-6395.

9. (a) Badi, N.; Lutz, J.-F., Sequence control in polymer synthesis. *Chemical Society Reviews* **2009,** *38* (12), 3383-3390; (b) Lutz, J.-F., Sequence-controlled polymerizations: the next Holy Grail in polymer science? *Polymer Chemistry* **2010,** *1* (1), 55-62; (c) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M., Sequence-Controlled Polymers. *Science* **2013,** *341* (6146).

10. (a) Chen, Y.; Guan, Z., Bioinspired Modular Synthesis of Elastin-Mimic Polymers To Probe the Mechanism of Elastin Elasticity. *Journal of the American Chemical Society* **2010,** *132* (13), 4577-4579; (b) Atallah, P.; Wagener, K. B.; Schulz, M. D., ADMET: The Future Revealed. *Macromolecules* **2013,** *46* (12), 4735-4741; (c) Li, Z.-L.; Lv, A.; Du, F.-S.; Li, Z.-C., Intrachain Cyclization via Postmodification of the Internal Alkenes of Periodic ADMET Copolymers: The Sequence Matters.

Macromolecules **2014,** *47* (17), 5942-5951.

11. Gutekunst, W. R.; Hawker, C. J., A General Approach to Sequence-Controlled Polymers Using Macrocyclic Ring Opening Metathesis Polymerization. *Journal of the American Chemical Society* **2015,** *137* (25), 8038-8041.

12. Merrifield, R. B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963,** *85* (14), 2149-2154.

13. Szostak, J. W., The eightfold path to non-enzymatic RNA replication. *Journal of Systems Chemistry* **2012,** *3* (1), 1-14.

14. Inoue, T.; Orgel, L., A nonenzymatic RNA polymerase model. *Science* **1983,** *219* (4586), 859-862.

15. von Kiedrowski, G., A Self-Replicating Hexadeoxynucleotide. *Angewandte Chemie International Edition in English* **1986,** *25* (10), 932-935.

16. von Kiedrowski, G.; Wlotzka, B.; Helbing, J., Sequence Dependence of Template-Directed Syntheses of Hexadeoxynucleotide Derivatives with 3′–5′ Pyrophosphate Linkage. *Angewandte Chemie International Edition in English* **1989,** *28* (9), 1235-1237.

17. Li, T.; Nicolaou, K. C., Chemical self-replication of palindromic duplex DNA. *Nature* **1994,** *369* (6477), 218-221.

18. Luther, A.; Brandsch, R.; von Kiedrowski, G., Surface-promoted replication and exponential amplification of DNA analogues. *Nature* **1998,** *396* (6708), 245-248.

19. Adamala, K.; Szostak, J. W., Nonenzymatic Template-Directed RNA Synthesis Inside Model Protocells. *Science* **2013,** *342* (6162), 1098-1100.

20. Chen, J. J.; Cai, X.; Szostak, J. W., N2′→P3′ Phosphoramidate Glycerol Nucleic

Acid as a Potential Alternative Genetic System. *Journal of the American Chemical Society* **2009,** *131* (6), 2119-2121.

21. (a) Lo, P. K.; Sleiman, H. F., Nucleobase-Templated Polymerization: Copying the Chain Length and Polydispersity of Living Polymers into Conjugated Polymers. *Journal of the American Chemical Society* **2009,** *131* (12), 4182-4183; (b) Liu, Y.; Wang, R.; Ding, L.; Sha, R.; Seeman, N. C.; Canary, J. W., Templated synthesis of nylon nucleic acids and characterization by nuclease digestion. *Chemical Science* **2012,** *3* (6), 1930-1937.

22. (a) Datta, B.; Schuster, G. B.; McCook, A.; Harvey, S. C.; Zakrzewska, K., DNA-Directed Assembly of Polyanilines: Modified Cytosine Nucleotides Transfer Sequence Programmability to a Conjoined Polymer. *Journal of the American Chemical Society* **2006,** *128* (45), 14428-14429; (b) Datta, B.; Schuster, G. B., DNA-Directed Synthesis of Aniline and 4-Aminobiphenyl Oligomers: Programmed Transfer of Sequence Information to a Conjoined Polymer Nanowire. *Journal of the American Chemical Society* **2008,** *130* (10), 2965-2973; (c) Chen, W.; Schuster, G. B., Precise Sequence Control in Linear and Cyclic Copolymers of 2,5-Bis(2-thienyl)pyrrole and Aniline by DNA-Programmed Assembly. *Journal of the American Chemical Society* **2013,** *135* (11), 4438-4449; (d) Niu, J.; Hili, R.; Liu, D. R., Enzyme-free translation of DNA into sequence-defined synthetic polymers structurally unrelated to nucleic acids. *Nat Chem* **2013,** *5* (4), 282-292.

23. Takahashi, Y.; Mihara, H., Construction of a chemically and conformationally self-replicating system of amyloid-like fibrils. *Bioorganic & Medicinal Chemistry* **2004,** *12*

(4), 693-699.

24. (a) Rubinov, B.; Wagner, N.; Rapaport, H.; Ashkenasy, G., Self-Replicating Amphiphilic β-Sheet Peptides. *Angewandte Chemie International Edition* **2009,** *48* (36), 6683-6686; (b) Rubinov, B.; Wagner, N.; Matmor, M.; Regev, O.; Ashkenasy, N.; Ashkenasy, G., Transient Fibril Structures Facilitating Nonenzymatic Self-Replication. *ACS Nano* **2012,** *6* (9), 7893-7901.

25. Hagenbuch, P.; Kervio, E.; Hochgesand, A.; Plutowski, U.; Richert, C., Chemical Primer Extension: Efficiently Determining Single Nucleotides in DNA. *Angewandte Chemie International Edition* **2005,** *44* (40), 6588-6592.

26. Bohler, C.; Nielsen, P. E.; Orgel, L. E., Template switching between PNA and RNA oligonucleotides. *Nature* **1995,** *376* (6541), 578-581.

27. Eigen, M., Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften 58* (10), 465-523.

28. (a) Deck, C.; Jauker, M.; Richert, C., Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nat Chem* **2011,** *3* (8), 603-608; (b) Kaiser, A.; Spies, S.; Lommel, T.; Richert, C., Template-Directed Synthesis in 3′- and 5′-Direction with Reversible Termination. *Angewandte Chemie International Edition* **2012,** *51* (33), 8299-8303.

29. (a) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M., Dynamic combinatorial chemistry. *Drug Discovery Today* **2002,** *7* (2), 117-125; (b) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S., Dynamic Combinatorial Chemistry. *Chemical Reviews* **2006,** *106* (9), 3652-3711; (c) Cougnon, F. B. L.; Sanders,

J. K. M., Evolution of Dynamic Combinatorial Chemistry. *Accounts of Chemical Research* **2012,** *45* (12), 2211-2221; (d) Jin, Y.; Yu, C.; Denman, R. J.; Zhang, W., Recent advances in dynamic covalent chemistry. *Chemical Society Reviews* **2013,** *42* (16), 6634-6654; (e) Li, J.; Nowak, P.; Otto, S., Dynamic Combinatorial Libraries: From Exploring Molecular Recognition to Systems Chemistry. *Journal of the American Chemical Society* **2013,** *135* (25), 9222-9239.

30. Herrmann, A., Dynamic mixtures and combinatorial libraries: imines as probes for molecular evolution at the interface between chemistry and biology. *Organic & Biomolecular Chemistry* **2009,** *7* (16), 3195-3204.

31. Furlan, R. L. E.; Ng, Y.-F.; Otto, S.; Sanders, J. K. M., A New Cyclic Pseudopeptide Receptor for Li+ from a Dynamic Combinatorial Library. *Journal of the American Chemical Society* **2001,** *123* (36), 8876-8877.

32. Hamieh, S.; Ludlow, R. F.; Perraud, O.; West, K. R.; Mattia, E.; Otto, S., A Synthetic Receptor for Nicotine from a Dynamic Combinatorial Library. *Organic Letters* **2012,** *14* (21), 5404-5407.

33. (a) Huc, I.; Lehn, J.-M., Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly. *Proceedings of the National Academy of Sciences* **1997,** *94* (6), 2106-2110; (b) Klekota, B.; Hammond, M. H.; Miller, B. L., Generation of novel DNA-binding compounds by selection and amplification from self-assembled combinatorial libraries. *Tetrahedron Letters* **1997,** *38* (50), 8639-8642.

34. Xu, S.; Giuseppone, N., Self-Duplicating Amplification in a Dynamic

Combinatorial Library. *Journal of the American Chemical Society* **2008,** *130* (6), 1826-1827.

35. (a) Goodwin, J. T.; Lynn, D. G., Template-directed synthesis: use of a reversible reaction. *Journal of the American Chemical Society* **1992,** *114* (23), 9197-9198; (b) Zhan, Z.-Y. J.; Lynn, D. G., Chemical Amplification through Template-Directed Synthesis. *Journal of the American Chemical Society* **1997,** *119* (50), 12420-12421.

36. (a) Li, X.; Zhan, Z.-Y. J.; Knipe, R.; Lynn, D. G., DNA-Catalyzed Polymerization. *Journal of the American Chemical Society* **2002,** *124* (5), 746-747; (b) Li, X.; Hernandez, A. F.; Grover, M. A.; Hud, N. V.; Lynn, D. G., Step-growth control in template-directed polymerization. *Heterocycles* **2011,** *82* (2), 1477-1488.

37. (a) Rosenbaum, D. M.; Liu, D. R., Efficient and Sequence-Specific DNA-Templated Polymerization of Peptide Nucleic Acid Aldehydes. *Journal of the American Chemical Society* **2003,** *125* (46), 13924-13925; (b) Kleiner, R. E.; Brudno, Y.; Birnbaum, M. E.; Liu, D. R., DNA-Templated Polymerization of Side-Chain-Functionalized Peptide Nucleic Acid Aldehydes. *Journal of the American Chemical Society* **2008,** *130* (14), 4646-4659.

38. Heemstra, J. M.; Liu, D. R., Templated Synthesis of Peptide Nucleic Acids via Sequence-Selective Base-Filling Reactions. *Journal of the American Chemical Society* **2009,** *131* (32), 11347-11349.

39. (a) Williams, R. J.; Smith, A. M.; Collins, R.; Hodson, N.; Das, A. K.; Ulijn, R. V., Enzyme-assisted self-assembly under thermodynamic control. *Nat Nano* **2009,** *4* (1), 19-24; (b) Sadownik, J. W.; Ulijn, R. V., Dynamic covalent chemistry in aid of peptide

self-assembly. *Current Opinion in Biotechnology* **2010,** *21* (4), 401-411; (c) Hirsch, A. K. H.; Buhler, E.; Lehn, J.-M., Biodynamers: Self-Organization-Driven Formation of Doubly Dynamic Proteoids. *Journal of the American Chemical Society* **2012,** *134* (9), 4177-4183; (d) Pappas, C. G.; Sasselli, I. R.; Ulijn, R. V., Biocatalytic Pathway Selection in Transient Tripeptide Nanostructures. *Angewandte Chemie International Edition* **2015,** *54* (28), 8119-8123.

40. Krishnan-Ghosh, Y.; Balasubramanian, S., Dynamic Covalent Chemistry on Self-Templating Peptides: Formation of a Disulfide-linked β-Hairpin Mimic. *Angewandte Chemie* **2003,** *115* (19), 2221-2223.

41. Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S., Mechanosensitive Self-Replication Driven by Self-Organization. *Science* **2010,** *327* (5972), 1502-1506.

Chapter Two

Investigating Dynamic N,O-acetal Condensation on Nucleic Acid and Peptide Backbones

2.1 Introduction

2.1.1 From dANP to rANP: Use of N,O-acetal Linkage

By replacing irreversible phosphodiester linkage with a reversible imine linkage to generate a DCN, our group successfully addressed the hurdle of low fidelity in non-enzymatic template-directed polymerization[1]. With < 1% error rate, we were able to accurately copy information on a DNA template with long sequences such as 32 bps into Amine Nucleoside Polymers (ANPs)[2]. However, $NaCNBH_3$ was required to drive polymerization (**Fig 2.1**). While this reductive amination approach yielded stable polymers, the dynamic property of DCN was lost, as amine formation is irreversible. Maintaining the dynamic nature of linkages is critical to build a network that responds to changing environmental stimuli and templates.

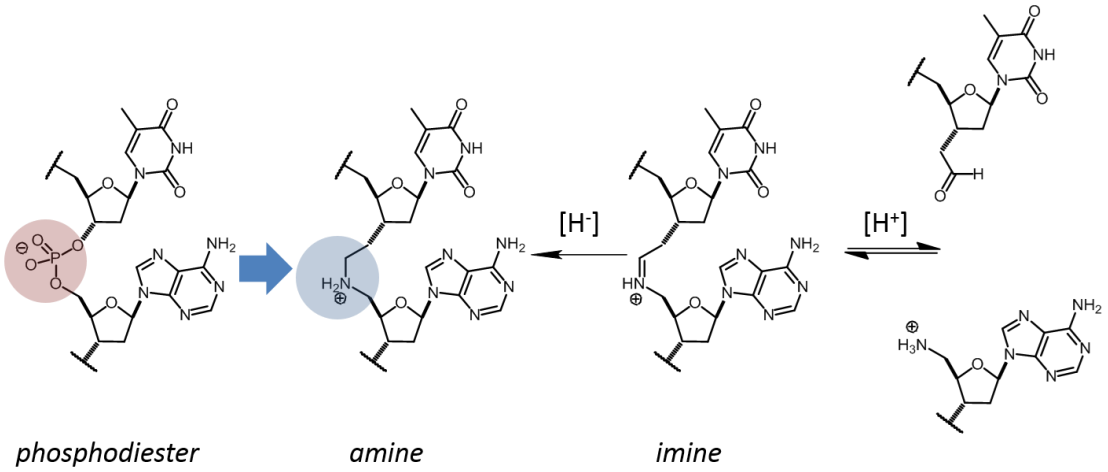


Fig 2.1 Replacing phosphodiester with imine to construct DCN. Imine coupling occurs between an aldehyde moiety and an amine group. Imine is trapped by reducing into secondary amine.

If the imine can be stabilized by another reversible reaction instead of irreversible reductive amination, then a conditionally dependent (such as pH) DCN can be preserved. On nucleic acid backbone, one simple approach is to switch to RNA (ribose Amine Nucleoside Polymer, rANP) as the DCN building block. As shown in **Fig 2.2**, the 2'-OH of rANP monomer can intramolecularly trap the imine intermediate to generate dynamic cyclic N,O-acetal linkage[3] (**Fig 2.2**), whose stability should be pH-dependent[4], providing a convenient tool to fine-tune the dynamics of DCNs. To study the dynamics of N,O-acetal formation on rANP backbone, the synthesis to generate two mono-functionalized uridine monomers were designed and achieved, one bearing an aldehyde moiety, another with amine, to investigate the N,O-acetal condensation reaction by NMR.

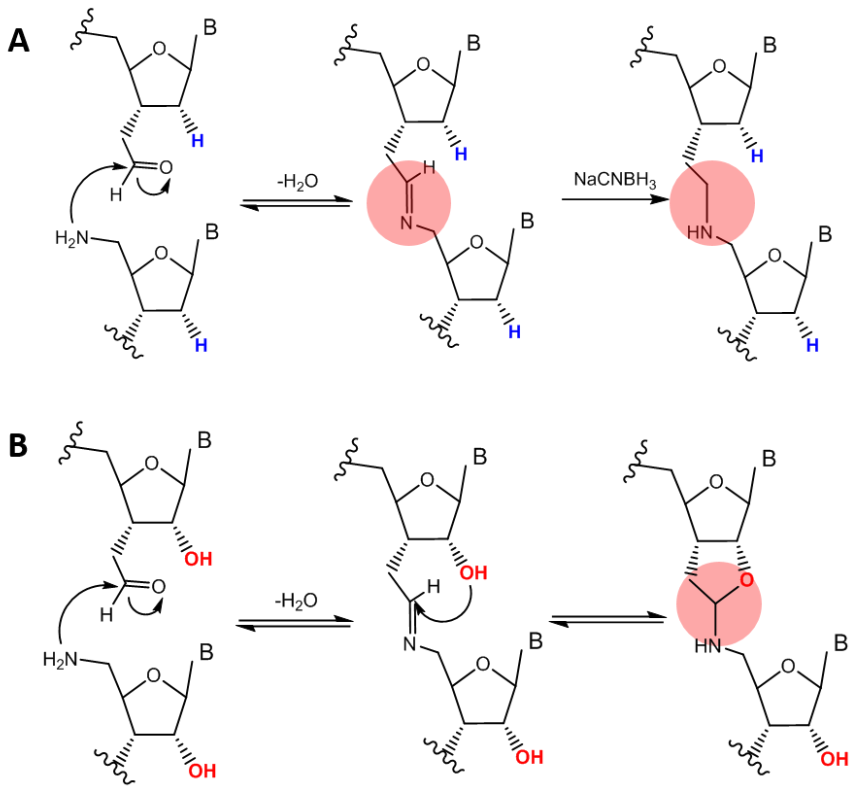


Fig 2.2 From dANP to rANP: trappling the imine linkage as N,O-acetal. (**A**) reductive amination of imine on deoxyribose-Amine Nucleoside Polymer (dANP). (**B**) 2'-Hydroxyl group reversibly trap the imine as N,O-acetal on ribose-Amine Nucleoside Polymer (rANP)

2.1.2 Extending N,O-acetal Linkages on Peptide Backbones

Reversible N,O-acetal linkage can be extended to peptide backbones for investigation of templated conformational replication in peptide dynamic chemical networks. Simple (-1) change of oxidation state of C-terminus into an aldehyde, allowing coupling with N'-terminus of another peptide to generate an imine. The inventory of diverse side chain functional groups on amino acids provides various options to further stabilize the imine as different acetals, as shown in **Fig 2.3**. Among them, threonine and serine side chains are positioned to trap the imine as a five member ring N,O-acetal. The amino acid chosen to be reduced into aldehyde is phenylalanine, due to the propensity of the

phenyl ring for promoting high-order assembly structures in peptides[5].

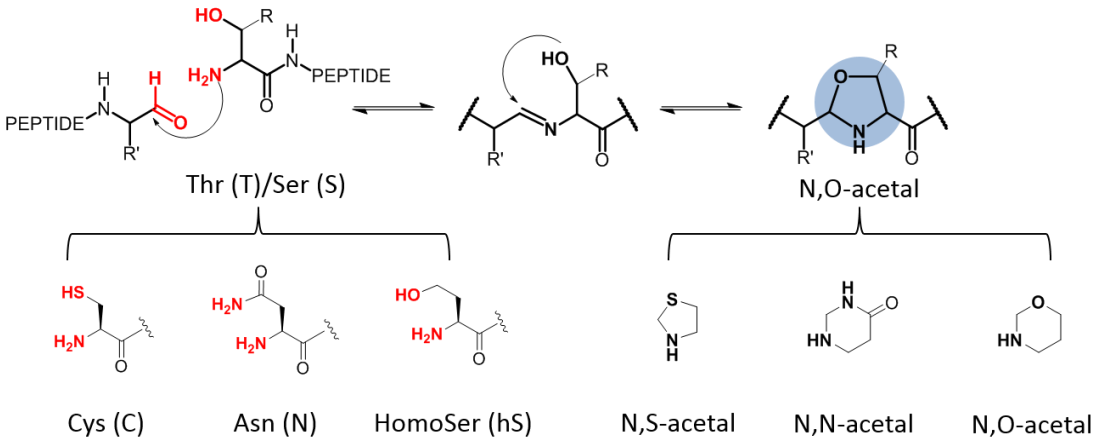


Fig 2.3 Illustration of reversible acetal linkages on peptide from side-chain trapping of imine. Inventory of amino acid side chains allows for trapping of imine as acetals with different heteroatoms and ring sizes.

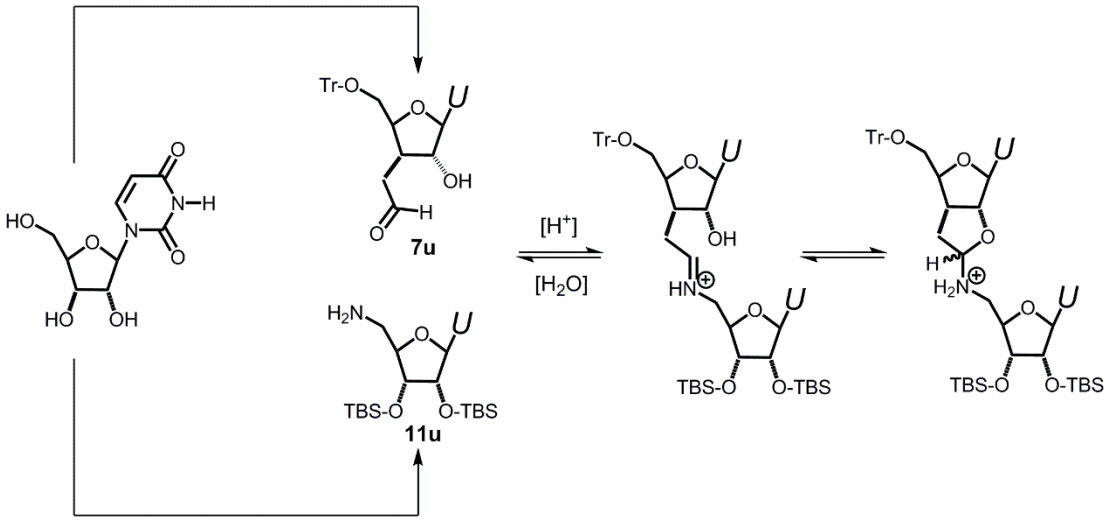
2.1.3 Stereo-Selectivity in N,O-acetal Condensation

N,O-acetal condensation gives rise to a new chiral center at acetal carbon with two possible configuration: *trans* and *cis*. Previous investigations[6] suggest energy difference between the two configurations in five-member ring is not significant, and acid-catalyzed acetalization of serine, threonine or cysteine esters yields 1:1 mixture of both isomers. But the ratio of two configurations is subject to change depending on substitution, N-acylation, ring-fusion and other factors. Therefore, knowledge of N,O-acetal condensation stereo-selectivity is a critical piece of information to define DCN behavior.

2.2 Results and Discussion

2.2.1 N,O-acetal on rANP Backbone

Mono-functionalized substrate **7u** and **11u** were synthesized from uridine, with **7u** as aldehyde monomer and **11u** as amine monomer, as shown in **Scheme 2.1**. After mixing two substrates 7.5 mM each in THF-d_8, but the aldehyde monomer **7u** is in slight excess to prevent possible nucleophilic attack on imine from primary amine of **11u**. The course of N,O-acetal condensation reaction was monitored by NMR over time under room temperature, as shown in **Fig 2.4**.



Scheme 2.1 N,O-acetal condensation on rANP monomer

After 24h, two NMR resonances assigned to N,O-acetal proton emerged at 5.25 ppm and 4.84 ppm respectively, with a 2:1 ratio. The formation of N,O-acetal is confirmed by 2D COSY (**Fig 2.5 A**), which shows coupling between N,O-acetal proton and adjacent methylene protons as well as one of the 5' protons. The N,O-acetal dimer is further confirmed by mass spectrometry (**Fig 2.5 B**).

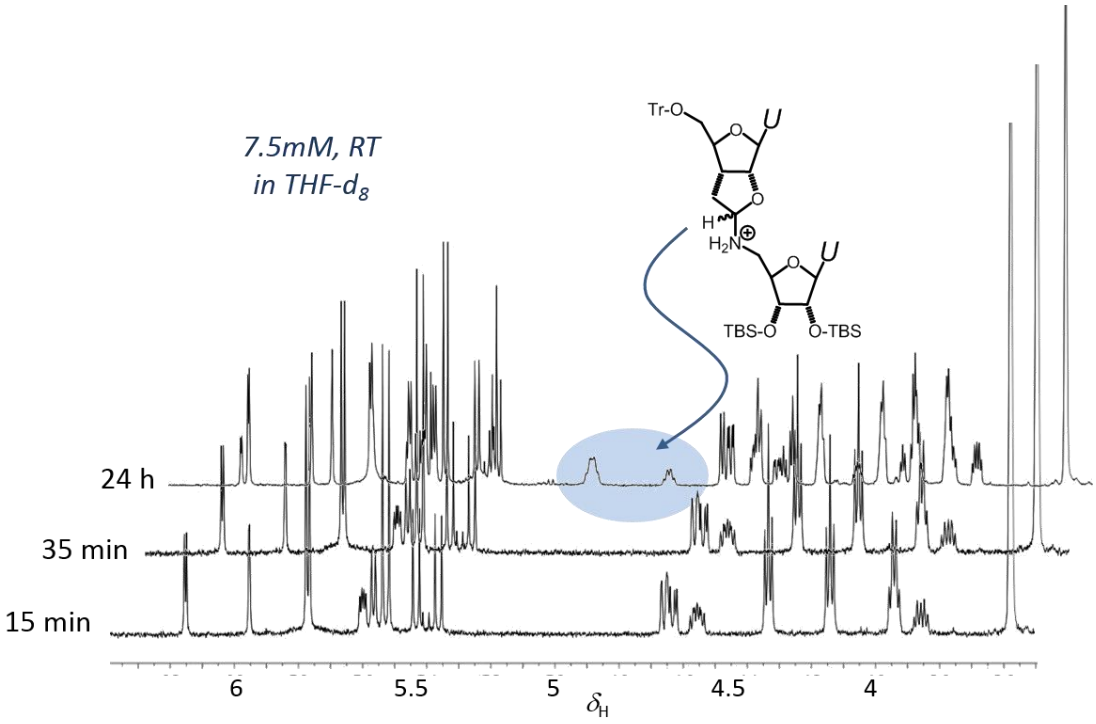


Fig 2.4 NMR time-course monitoring of N,O-acetal condensation on rANP backbone. Two resonances at 5.25 ppm and 4.84 ppm emerges as the reaction progresses.

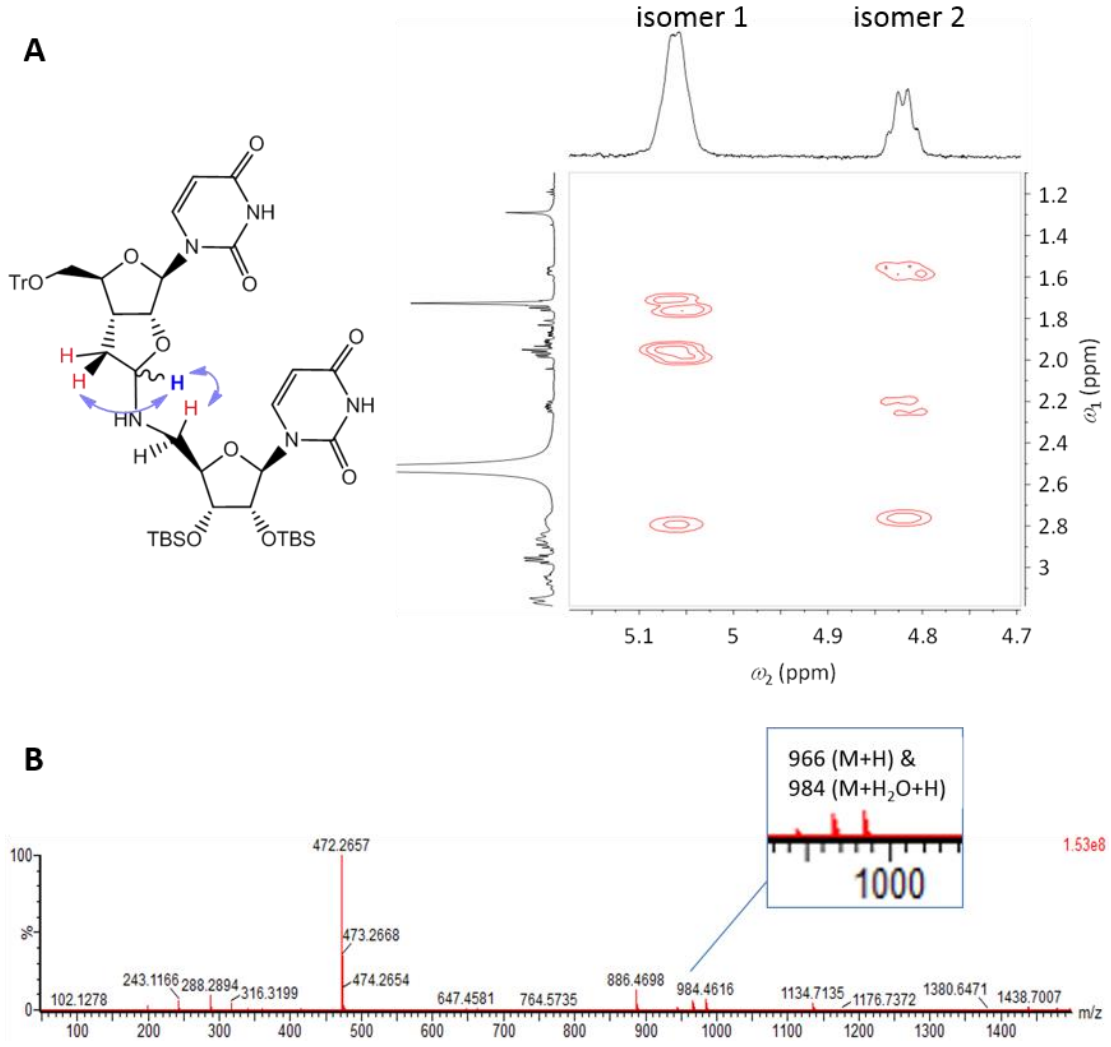
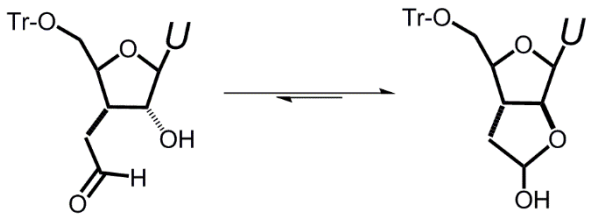


Fig 2.5 Characterization of N,O-acetal linked dimer. (**A**) 2D COSY spectrum indicates coupling between N,O-acetal proton and adjacent protons on methylene carbon. (**B**) Mass spectrometry confirms dimer formation

2.2.2 Kinetics of N.O-acetal Condensation on rANP Dimer

We noticed that N,O-acetal condensation between **7u** and **11u** is substantially slower than literature reports[4]. NMR analysis of starting material reveals that compound **7u**, the aldehyde substrate, does not exist in the free aldehyde form. Instead, the 2'-OH intramolecularly traps the methyl aldehyde to yield ring-closure product lactol, as shown in **Scheme 2.2**. As a result, ring-opening step is required to release the free

aldehyde for N, O-acetal condensation to take place. The five member ring lactol tautomer is energetically favored over and ring-opening step becomes rate-limiting for N,O-acetal condensation reaction.



Scheme 2.2 Ring-closure lactol is favored over free aldehyde in **7u**

To improve the reaction rate, catalytic (< 0.1eq) TFA-d was added upon mixing of **7u** and **11u**, as shown in **Fig 2.6**, addition of acid significantly facilitates N,O-acetal condensation. The reaction reaches equilibrium with ~90% conversion within 24h. The ratio of the two isomers remains to be 2:1.

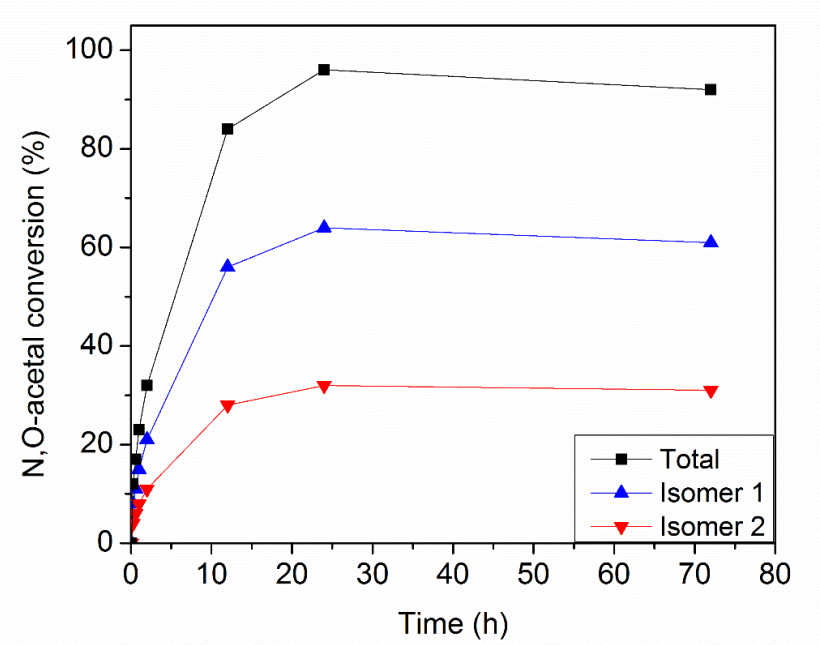


Fig 2.6 TFA catalysis of N,O-acetal condensation between **7u** and **11u**. Catalytic amount of TFA greatly accelerates the condensation reaction, with the equilibrium being reached around 24 hrs. The ratio of two isomers does not change. The N,O-acetal conversion rate is measured by integration of N,O-acetal resonances vs TMS internal standard.

Raising the reaction temperature can also increase the condensation reaction rate. As shown in **Fig 2.7**, increasing temperature to 45°C gives even more pronounced rate-acceleration compared to acid catalysis, but at the cost of fidelity in templated system. At 45°C, the reaction reaches equilibrium within 2 hours. However, compared to acid catalysis, thermal catalysis gives lower equilibrium constant (K = 16 at 45°C without TFA catalysis vs K = 360 at 25°C with TFA catalysis).

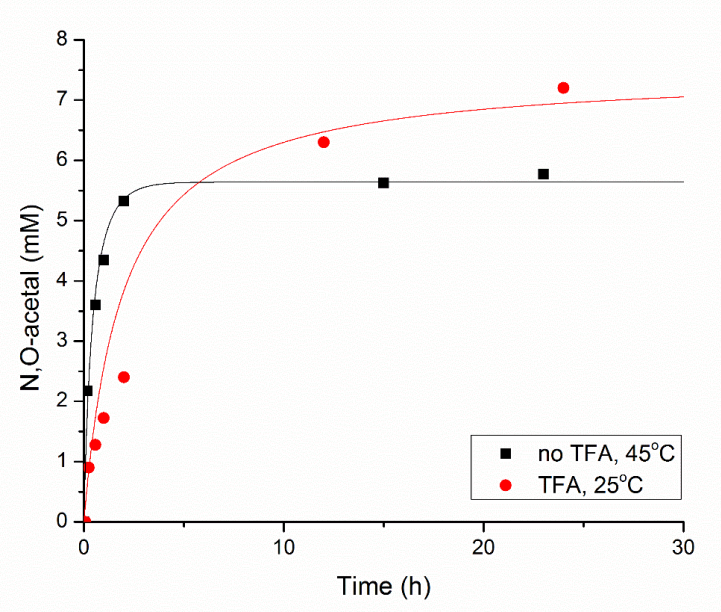


Fig 2.7 Comparison of thermal catalysis and TFA catalysis. At 45°C, condensation reaction reaches equilibrium at ~2 hours, faster than TFA catalyzed reaction at 25°C, but has lower equilibrium constant. Concentration is measured by integration of N,O-acetal resonance on NMR spectra and calculated with internal standard TMS and initial concentration of substrates. Data is fitted to a simple second-order reaction mechanism.

2.2.3 Kinetics of N, O-acetal Hydrolysis on rANP Dimer

Mechanism of N, O-acetal condensation involves releasing one molecule of water, thus addition of water to the system should be able to drive the reaction backwards. On addition 20 μL of D_2O into 750 μL THF-d_8 solution (~2.5% D_2O (v/v)) at 25°C, no change in N,O-acetal ^{1}H-NMR intensity was observed. After increasing temperature (**Fig 2.8**, after 30 min) to 45°C, however, hydrolysis began to take place and starting materials were re-generated.

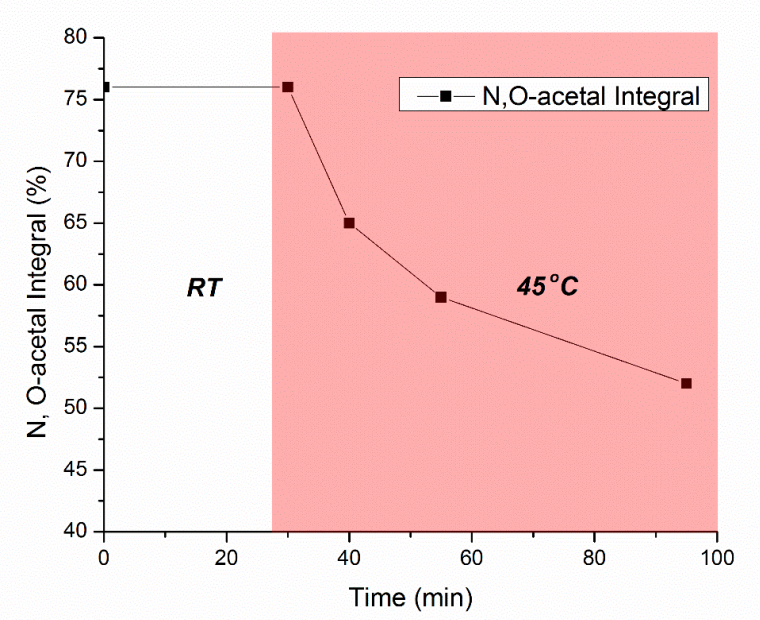


Fig 2.8 N,O-acetal hydrolysis in 2.5% D_2O in THF-d_8 (v/v) at 25°C and after increasing temperature to 45°C. At 25°C, no change in N,O-acetal ^{1}H-NMR intensity is observed, suggesting N,O-acetal is stable under such condition. After heating to 45°C, hydrolysis starts to occur and starting materials are regenerated. No TFA added in solution.

pH dependence of N,O-acetal hydrolysis was also investigated. The equilibrated N,O-acetal was mixed in 1:1 with D_2O buffer of different pHs. Acidic buffer pD = 4.5 acetate-d_4 and pD = 5.5 pyridine-d_5 buffer hydrolyzes the N,O-acetals within 5 min. In near-neutral buffer (pD = 7.0 KD_2PO_4 buffer, pD = 7.5 KD_2PO_4), the hydrolysis is much slower, with the half-life of N, O-acetal more than 1 hr (**Fig 2.9**), consistent with literature reported N,O-acetal pH-dependence profile[4c]. These experiments clearly demonstrate the temperature and pH dependence of N,O-acetals, opening up opportunity for environmental and template regulation.

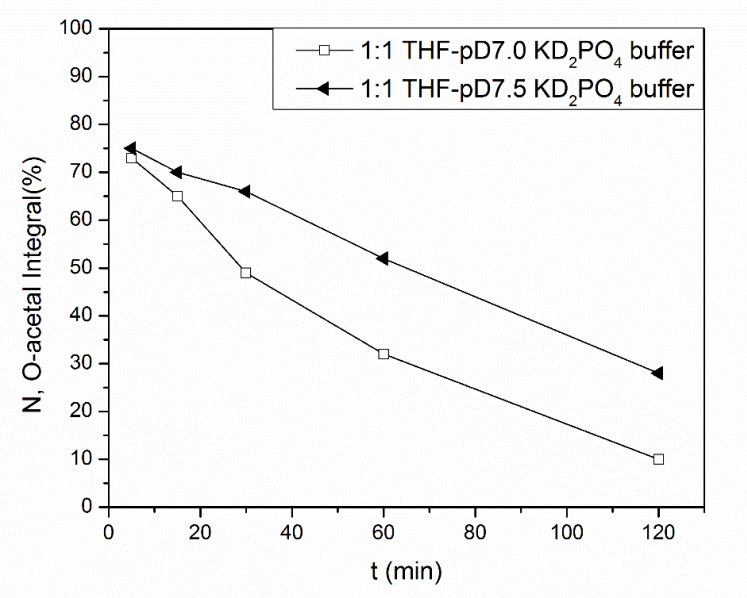


Fig 2.9 Hydrolysis of N,O-acetal in 1:1 THF-d_8-D_2O mix solvent is pH-dependent. In acidic buffer (pD = 4.5 and 5.5), N,O-acetals hydrolyze within 5 min, while under neutral pH (pD = 7 and 7.5), N,O-acetal is relatively stable against hydrolysis, with half-life more than 1 hr.

2.2.4 Stereochemistry Assignment of Diastereomers

Two diastereomers in a 2:1 ratio are generated in N,O-acetal condensation reaction on ANP backbone. Nuclear Overhauser Effect (NOE) experiments were designed to elucidate the configuration of the two diastereomers. By irradiating each N,O-acetal proton resonance individually, and combining these assignments with COSY spectra, all resonances with significant (> 0.5%) enhancement upon irradiation were assigned. As shown in **Fig 2.10**, irradiation on isomer 1 N,O-acetal proton resonance at 5.25 ppm causes enhancement of 1'-H proton and 4'-H on aldehyde segment of the dimer, suggesting on isomer 1, N,O-acetal linkage is at the same side of ring with these two protons. While for the isomer 2, upon irradiation of N,O-acetal proton resonance at 4.48 ppm, positive enhancement of 2'-H was observed, indicating this N,O-acetal shares

same face with 2’-H. (3’-H resonance overlaps with several other proton signals and is difficult to assign).

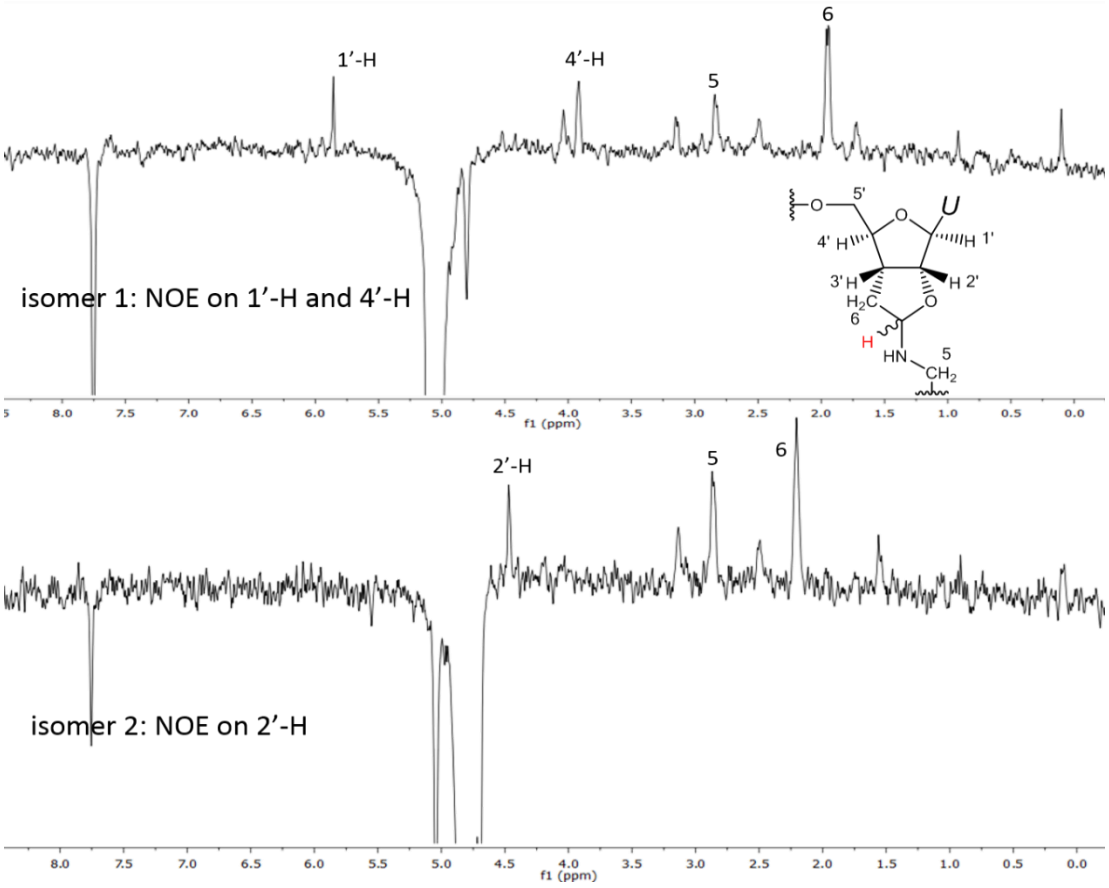


Fig 2.10 NOE for N,O-acetal condensation product. Isomer 1 N,O-acetal proton is on the same face of 1’-H and 4’-H, while isomer 2 N,O-acetal proton is on the same face of 2’-H.

Based on NOE results, we proposed models for two isomers as shown in **Fig 2.11**. Note that we flip the ring pucker of isomer 2 to place the bulky group equatorial, instead of keeping the ring conformation and pointing bulky group towards axial position because we observed similar percentage enhancements for 1’-H & 4’-H on isomer 1, and 2’-H on isomer 2, suggesting the distances between protons are similar for both isomers. If the bulky group is axial for isomer 2, the distance between 2’-H and acetal proton will be further away and weaker enhancements are expected.

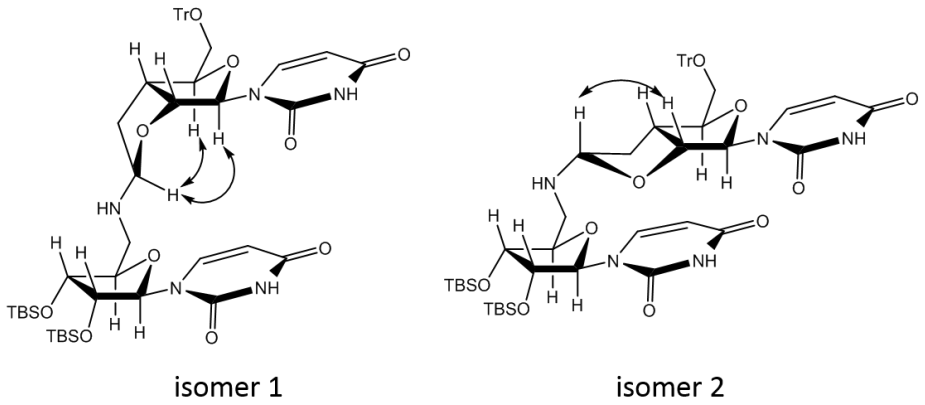
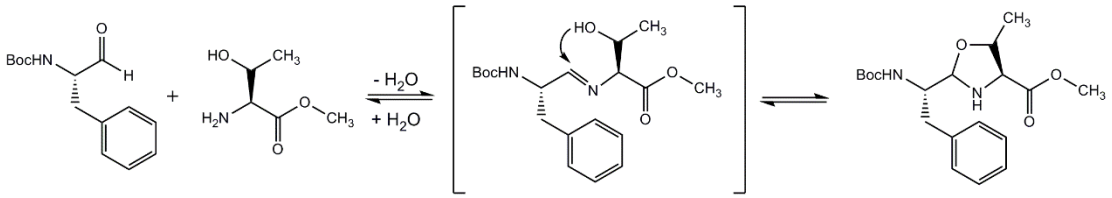


Fig 2.11 Proposed structure models for 2 isomers

2.2.5 N,O-acetal Condensation and Kinetics on Peptide Backbones

Two mono-functionalized substrates, *N*-Boc-L-Phe-CHO and L-Thr methyl ester are selected for similar analysis of N,O-acetals on peptide backbone. As shown in **Scheme 2.3**, the hydroxyl group on Thr side chain can trap the intermediate imine and generate the cyclic N,O-acetal.



Scheme 2.3 N,O-acetal condensation mechanism from *N*-Boc-L-Phe-CHO and L-Thr methyl ester substrates

In benzene-d_6, N,O-acetal proton NMR resonance appears at 4.43 ppm, concomitant with the loss of aldehyde proton at 9.04 ppm. The yield and kinetics of N,O-acetal condensation were determined by integration of resonances relative to the TMS internal

standard over time as shown in **Fig 2.12**. Within two hours, the reaction reaches equilibrium (**Fig 2.13**). Detailed assignment of each resonance on acetal product was achieved by COSY analysis and is shown in **Fig 2.14**.

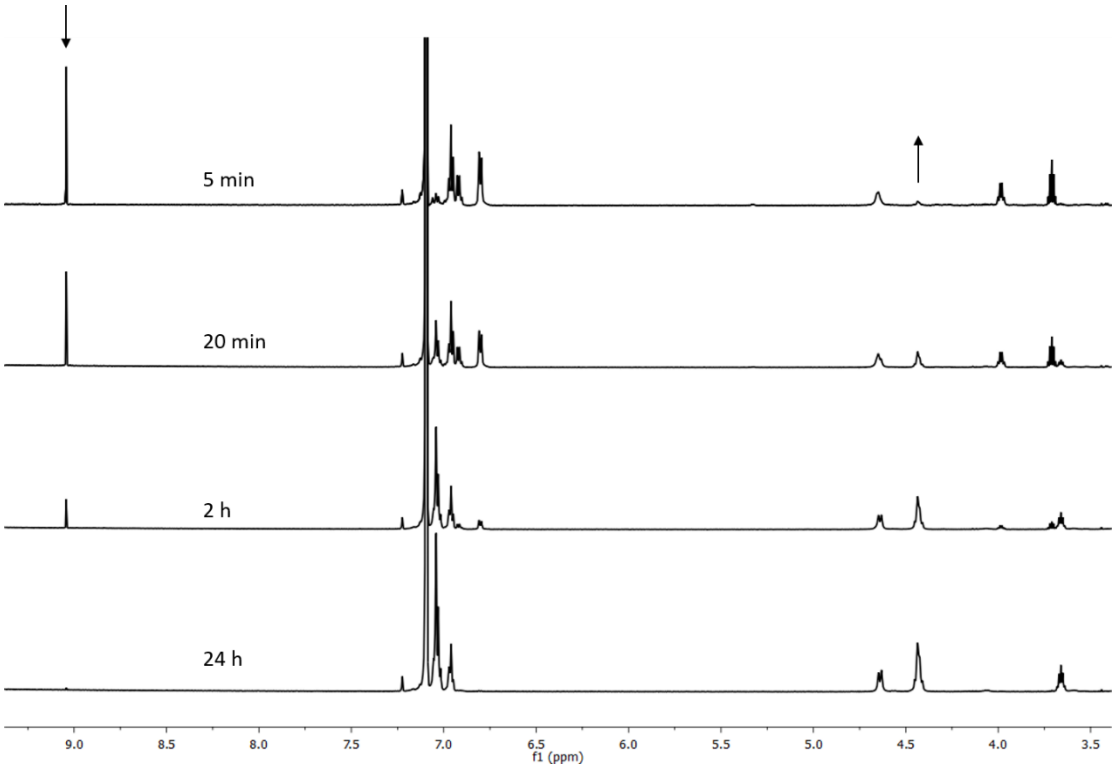


Fig 2.12 ^{1}H-NMR monitor of N,O-acetal condensation on F-T peptide backbone. The loss of aldehyde proton resonance at 9.04 ppm is concomitant with emergence of acetal proton resonance at 4.43 ppm.

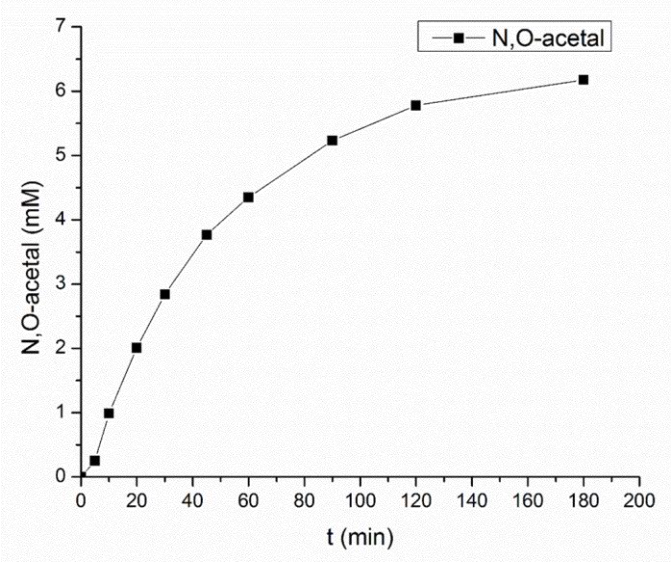


Fig 2.13 Kinetics of N,O-acetal condensation on F-T peptide backbone. Condensation reaction reaches equilibrium within 3 hrs.

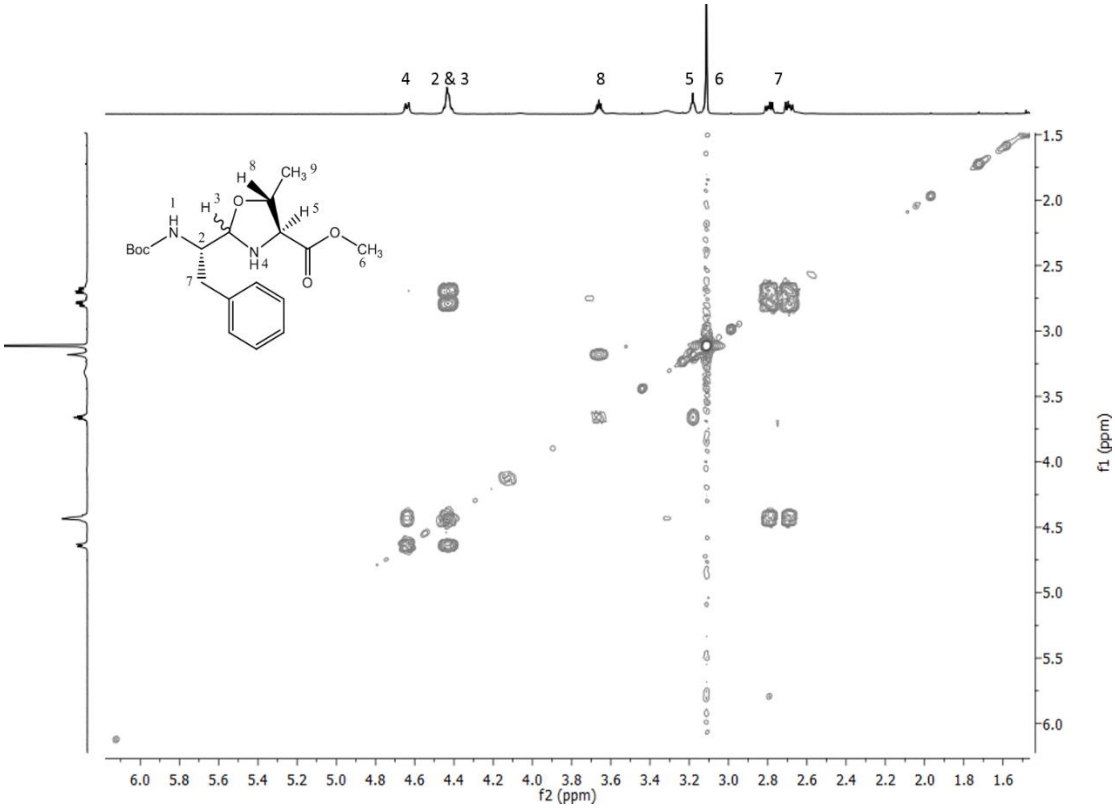


Fig 2.14 2D COSY spectrum of N,O-acetal condensation product on F-T peptide backbone. All resonances between 1.5 ppm to 6 ppm are assigned and labeled accordingly. Resonance on 4.43 ppm contains 2 protons, on coupled with proton 4, assigned as acetal proton 3, another coupled with phenylalanine methyl proton 7, assigned as proton 2.

2.2.6 Diastereoselectivity of N,O-acetal Condensation on Peptide Backbone

Two diastereomers are possible for cyclic N,O-acetal: (S,S)-*trans* and (R,S)-*cis*, as shown in **Fig 2.15**, however, only a single N,O-acetal proton resonance was observed. The (R,S)-*cis* diastereomer, with the bulky group positioned quasi-equatorial, is less sterically hindered and energetically more favorable, as shown in **Fig 2.15**, top. As shown in **Fig 2.16**, in NOE experiment, irradiating the acetal proton enhances proton 5 resonance relative to proton 8. Due to the overlap of the acetal proton and the Phe α proton (proton 2), proton 7 resonance is also enhanced. This result supports that N,O-acetal adopts *cis* configuration, as shown in **Fig 2.15**.

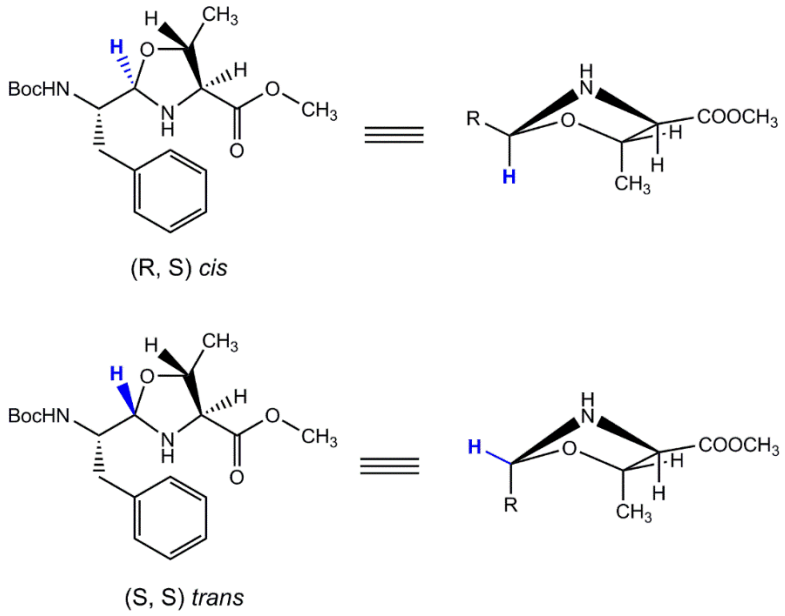


Fig 2.15 Two possible configurations of N,O-acetal on T-F peptide backbone. (R,S)-*cis* configuration is predicted to be energetically more favored, for the bulky R group is at quasi-equatorial position and sterically less hindered, compared to (S,S)-*trans* configuration, which the R group is at axial position.

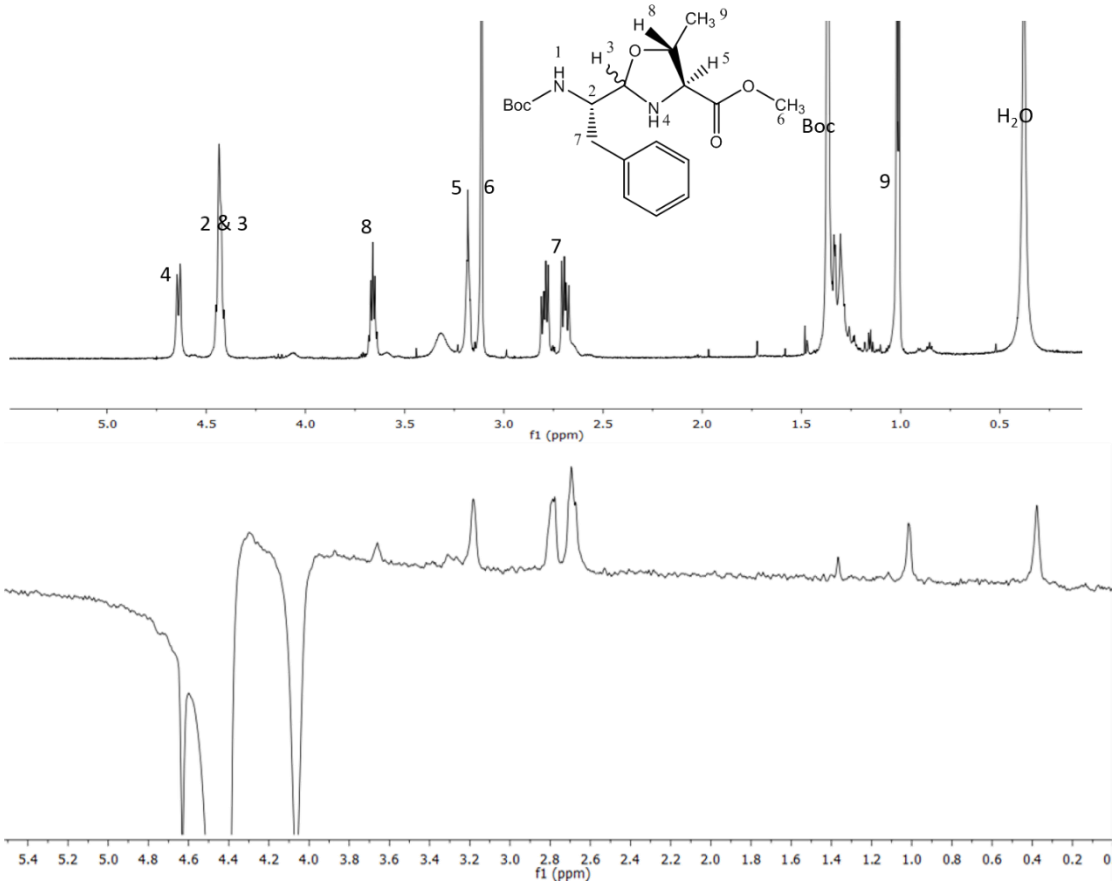


Fig 2.16 NOE experiment of N,O-acetal on F-T peptide backbone. Irradiation of acetal proton 3 resonance gives positive enhancement of proton 5 but not 8, indicates protons 3 and 5 are at the same face of the ring, supporting the acetal adopts (R, S)-*cis* configuration.

2.3 Conclusion

In this chapter, we proposed the use of the N,O-acetal linkage to reversibly trap intermediate imines to construct dynamic DCNs responsive to environmental and template effects. Using two monofunctionalized monomers, one with aldehyde, another with amine, we are able to create model systems suitable for detailed study of the kinetics and stereoselectivity of the reversible N, O-acetal linkage by NMR.

For condensation on two rANP monomers, the kinetics are substantially slower than

the N,O-acetal on peptide backbones, likely because of the ring closure of aldehyde monomer generates energy-favored fused bicyclic lactol structure. However, by controlling pH or temperature, the reaction rate can be accelerated. Likewise, hydrolysis of the N, O-acetal is also temperature and pH dependent. The N,O-acetal being generated on rANP backbone is a mixture of two diastereomers with 2:1 ratio, and the configurations of two isomers are tentatively assigned by NOE and COSY experiments. While for N, O-acetal condensation between two peptide monomers, N-Boc-L-Phe-CHO and L-Thr methyl ester, the condensation reaction in benzene occurs fast and reaches equilibrium within hours. This reaction selects for the (R, S)-*cis* configuration, a configuration that is sterically more preferred.

Understanding how different environmental factors impact the rate and stability of the N,O-acetal is critical for our design and development of dynamic chemical networks. Both condensation and hydrolysis are pH dependent, suggesting we should be able to drive the reaction to proceed or "freeze" the network from exchange by tuning the pH. The reaction proceeds more readily in organic solvent, meaning that hydrophobic environment is preferred to generate and stabilize the product. Hydrophobic environments can also be achieved by other approaches, such as template-induced peptide self-assembly and DNA template surfaces. Higher temperature can destabilize association between template and substrates, but can be used to drive to different region of the reaction surface[7].

In summary, the NMR study in this chapter provides a solid foundation for our understanding of the nature of the reversible linkage, and strengths and limitations for establishing our fully-functionalized DCNs described later in this thesis.

2.4 Experimental

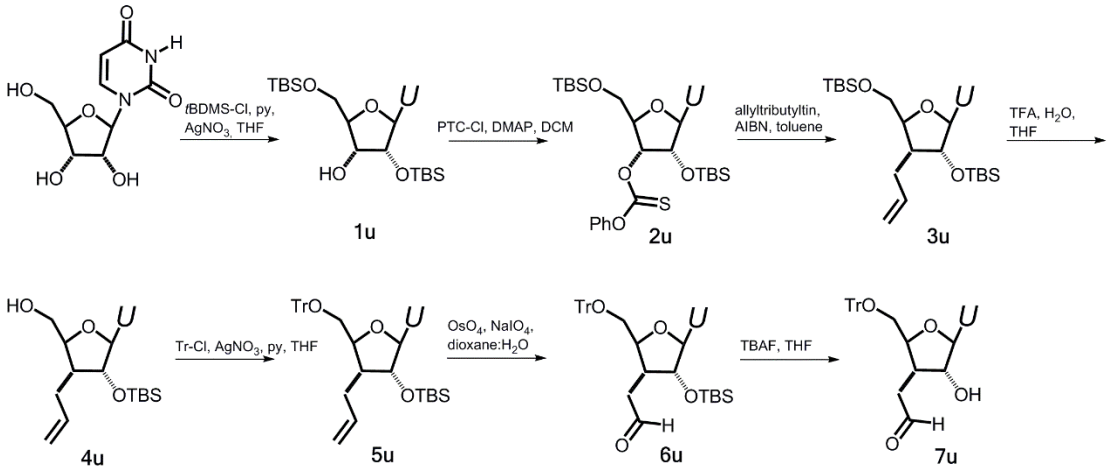
2.4.1 General Methods

All chemical reagents and solvents are purchased from Sigma-Aldrich, Fisher, AnalSpec, Alfa-Aesar and Cambridge Isotope Laboratories without further purification unless otherwise noted. All reactions are performed under N_2 protection unless otherwise noted. All 1D and 2D NMR was performed on an INOVA 400 MHz or an INOVA 600 MHz NMR spectrometer. FT-IR spectra were acquired using a Jasco FT-IR 4100 at room temperature and averaging 500 to 1000 scans with 2 cm^{-1} resolution, using an MCT detector with a 5mm aperture and a scanning speed of 4mm/sec. Aliquots (10μL) of substrate ethyl acetate solution were dried as thin films on a Pike Galdi ATR diamond crystal.

2.4.2 Synthesis of 3'-CHO-5'-Tr-Uridine (7u)

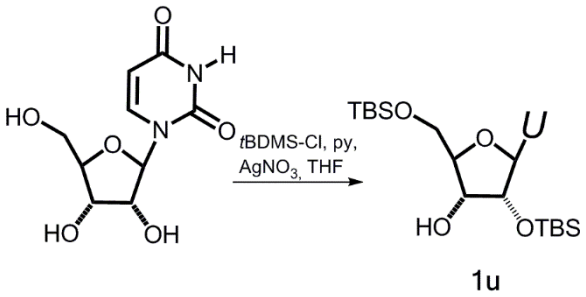
Synthetic route of **7u** is shown in **Scheme 2.4**. The 2'- and 5'- hydroxyl groups are protected with TBS using $AgNO_3$ as catalyst. The free 3'-hydroxyl group is activated by a phenylthiolcarbonyl group. Then the carbon skeleton of 3'-aldehyde moiety is introduced by a stereoselective radical allylation reaction initiated by AIBN, and masked as an olefin. The bulky 5'-TBDMS protecting group prevents the radical

coupling to attack the β face of ribose ring, giving a single diastereomer with selectivity > 9:1. The 5'-TBDMS group can be selectively removed using TFA: H_2O : THF = 1:1:4 (v/v/v) mixed solvent, and protected with trityl group. The aldehyde is later unmasked by oxidation of olefin, and TBS group on 2'- position is deprotected by TBAF.



Scheme 2.4 Synthetic route of 3'-CHO-5'-Tr-Uridine (**7u**)

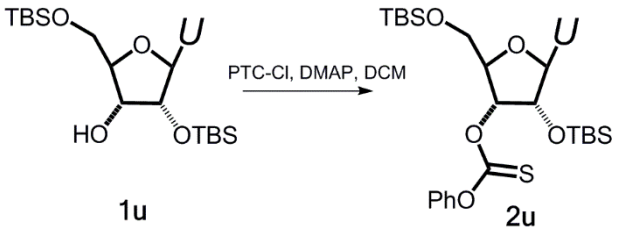
*1. 2', 5'-O-bis(tert-butyldimethylsilyl)uridine (**1u**)*[8]



To a suspension of 2.44 g uridine (10mmol) in 100 mL THF were added 3.89 mL pyridine (5eq), 3.74 g $AgNO_3$ (2.2eq) then 3.32 g TBDMS-Cl were added. Stir overnight at room temperature under N_2 protection. The precipitant was filtered off and solvent was removed *in vacuuo.* Residue was diluted by EtOAc, washed with H_2O, saturated $NaHCO_3$ solution then brine, dried and concentrated. Purified by column with Ether:Hexane = 4:1 to give the product **1u**. Yield: 83%.^{1}H NMR (600 MHz, cdcl$_3$) δ

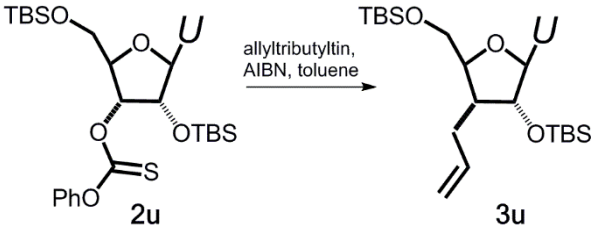
8.78 – 8.59 (s, 1H), 8.01 – 7.96 (m, 1H), 5.96 (dd, *J* = 10.1, 4.4 Hz, 1H), 5.69 (dd, *J* = 8.1, 1.1 Hz, 1H), 4.19 (t, *J* = 4.4 Hz, 1H), 4.10 (dt, *J* = 4.3, 2.6 Hz, 2H), 3.98 (d, *J* = 11.6 Hz, 1H), 3.81 (d, *J* = 11.6 Hz, 1H), 2.60 (s, 1H), 1.04 – 0.73 (m, 18H), 0.29 – -0.03 (m, 12H). HRMS: $M+H^+$ ($C_{21}H_{41}N_2O_6Si_2$): calculated: 473.2503, found: 473.2514

*2. 2', 5'-O-bis(tert-butyldimethylsilyl)-3'-O-phenylthiocarbonyl-uridine (**2u**)*[9]



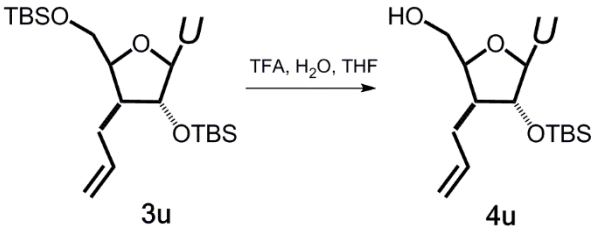
472 mg **1u** (1mmol) and 488 mg DMAP (4eq) were dissolved in DCM. Under N_2 protection, PTC-Cl was added dropwise at 0°C. The reaction mixture then was warmed up to RT and stir for 18h. After removal of solvent under reduced pressure, column chromatography (DCM then DCM : MeOH = 10:1) was used to purify the residue and give the product **2u**. Yield: 81%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8.34 (s, 1H), 7.98 (t, *J* = 7.5 Hz, 1H), 7.45 – 7.40 (m, 2H), 7.31 (dt, *J* = 13.2, 3.3 Hz, 1H), 7.07 (dd, *J* = 7.6, 1.0 Hz, 2H), 6.05 (t, *J* = 5.8 Hz, 1H), 5.71 (d, *J* = 8.2 Hz, 1H), 5.59 (t, *J* = 4.7 Hz, 1H), 4.50 (t, *J* = 4.7 Hz, 1H), 4.48 – 4.43 (m, 1H), 4.05 (d, *J* = 11.8 Hz, 2H), 3.91 (dd, *J* = 8.2, 3.5 Hz, 2H), 1.00 – 0.88 (m, 18H), 0.17 – 0.06 (m, 12H). HRMS: $M+H^+$ ($C_{28}H_{45}N_2O_7SSi_2$): calculated: 609.24861, found: 609.24847.

*3. 3'-C-Allyl-2', 5'-O-bis(tert-butyldimethylsilyl)-uridine (**3u**)*[10]



To a solution of 3.66g **2u** (7.75mmol) in 50 mL benzene was added 104 mg AIBN (0.1eq) and 14 mL allyltributylstanne (7.5eq). The solution was degassed for 15min, then reflux at 80°C under N_2 protection overnight. After cooling down to room temperature, the solvent then was removed *in vacuuo* and residue was purified by silica gel column chromatography with Hexane : EtOAc = 5:1 through 3:1 to give the product **3u**. Yield: 68%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8.25 (d, *J* = 8.2 Hz, 1H), 8.15 (s, 1H), 5.78 – 5.68 (m, 1H), 5.65 (s, 1H), 5.61 (t, *J* = 6.6 Hz, 1H), 5.10 – 5.01 (m, 2H), 4.24 (d, *J* = 4.0 Hz, 1H), 4.15 (d, *J* = 12.1 Hz, 1H), 4.05 (d, *J* = 10.1 Hz, 1H), 3.72 (dt, *J* = 10.3, 5.1 Hz, 1H), 2.39 – 2.31 (m, 1H), 2.14 (ddd, *J* = 14.1, 8.4, 3.5 Hz, 1H), 2.04 – 1.97 (m, 1H), 0.96 – 0.85 (m, 18H), 0.27 – 0.05 (m, 12H).M+H$^+$ ($C_{24}H_{45}N_2O_5Si_2$): calculated: 497.28670, found: 497.28647.

*4. 3'-C-Allyl-5'-O- tert-butyldimethylsilyl-uridine (**4u**)*[11]



400 mg **3u** (0.81 mmol) was dissolved in 8 mL THF and cooled down to 0°C, 4 mL TFA : H_2O = 1:1 mixture was added dropwise. Stirred for 4h at 0°C, then neutralized

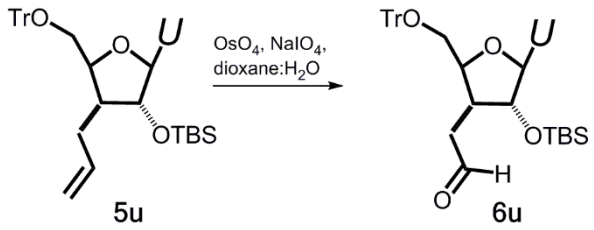
with saturated $NaHCO_3$ solution. Extract with 50 mL EtOAc 3 times, combined the organic layer, washed with 20 mL H_2O followed by 20 mL brine, dried with $MgSO_4$ then concentrated. The residue was purified by column with DCM:MeOH = 20:1 to give **4u**. Yield: 87%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 9.14 (s, 1H), 8.13 (d, *J* = 8.1 Hz, 1H), 5.80 – 5.69 (m, 1H), 5.66 (t, *J* = 6.6 Hz, 1H), 5.61 (s, 1H), 5.12 – 5.01 (m, 2H), 4.31 (d, *J* = 4.3 Hz, 1H), 4.14 – 4.05 (m, 2H), 3.80 – 3.74 (m, 1H), 2.34 (dt, *J* = 14.1, 7.0 Hz, 1H), 2.16 – 2.03 (m, 2H), 0.96 – 0.77 (m, 9H), 0.25 – 0.07 (m, 6H). HRMS: M+H$^+$ ($C_{18}H_{31}N_2O_5Si$): calculated: 383.20023, found: 383.20007.

*5. 3'-C-Allyl-5'-O-tert-butyldimethylsilyl-2'-O-trityl-uridine (**5u**)*[8]

4u 5u

1.07 g **4u** (2.8 mmol) was dissolved in 80 mL THF. 0.856 g Tr-Cl (1.1 eq), 0.521g $AgNO_3$ (1.1 eq) and 1.08mL pyridine (5 eq) was added. The reaction mixture was stirred at room temperature overnight, filtered, concentrated and purified by flash column with Hexane : EtOAc = 3:1 to give **5u**. Yield: 79%.^{1}H NMR (400 MHz, cdcl$_3$) δ 9.74 (s, 1H), 8.25 (d, *J* = 8.1 Hz, 1H), 7.46 – 7.36 (m, 6H), 7.36 – 7.20 (m, 9H), 5.75 – 5.60 (m, 2H), 5.24 (dd, *J* = 8.1, 2.0 Hz, 1H), 5.02 (dd, *J* = 17.0, 1.5 Hz, 1H), 4.92 (dd, *J* = 10.1, 1.4 Hz, 1H), 4.28 (t, *J* = 7.0 Hz, 1H), 3.70 (dd, *J* = 11.3, 1.7 Hz, 1H), 3.34 (dd, *J* = 11.4, 2.7 Hz, 1H), 2.36 – 2.21 (m, 2H), 1.88 (dd, *J* = 17.5, 8.9 Hz, 2H), 0.89 (s, 9H), 0.19 (d, *J* = 56.3 Hz, 6H).

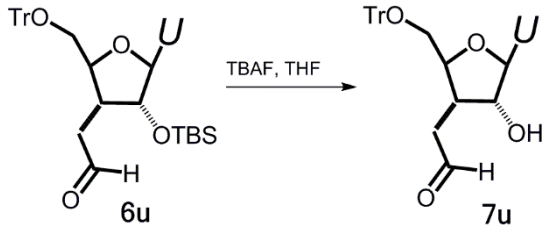
*6. 3'-C-acetaldehyde-5'-O-tert-butyldimethylsilyl-2'-O-trityl-uridine (**6u**)[10]*



5u 6u

100 mg **5u** (0.16 mmol) was dissolved into 4 mL dioxane and 2 mL H_2O. 0.1 mL OsO_4 (4 wt% in H_2O) was added dropwise. Solution turned brown in 5min. Stir for another 15 min, then $NaIO_4$ was dissolved into 30 mL dioxane and 20 mL H_2O and added dropwise over 1h. Stirred for another 1h, then diluted the reaction mixture with 50 mL H_2O plus 10 mL brine, extract with 50 mL DCM for 3 times, dried and concentrated. Residue was purified by silica gel column with Hexane : EtOAc = 2:1 to give product **6u**. Yield: 67%. ^{1}H NMR (400 MHz, cdcl$_3$) δ 9.66 (s, 1H), 9.31 (s, 1H), 8.19 (d, *J* = 8.2 Hz, 1H), 7.39 (dt, *J* = 10.8, 5.6 Hz, 6H), 7.35 – 7.20 (m, 9H), 5.73 (s, 1H), 5.34 (dd, *J* = 8.1, 2.2 Hz, 1H), 4.45 (d, *J* = 3.6 Hz, 1H), 4.02 (d, *J* = 10.3 Hz, 1H), 3.79 (dd, *J* = 11.7, 2.1 Hz, 1H), 3.28 (dd, *J* = 11.7, 2.6 Hz, 1H), 2.69 – 2.54 (m, 2H), 1.95 (t, *J* = 11.8 Hz, 1H), 0.85 (s, 9H), 0.12 (d, *J* = 82.8 Hz, 6H). HRMS: M+H$^+$ ($C_{36}H_{43}N_2O_6Si$): calculated: 627.28904, found: 627.28747

*7. 3'-C-Acetaldehyde-5'-O-Trityl-uridine (**7u**)*

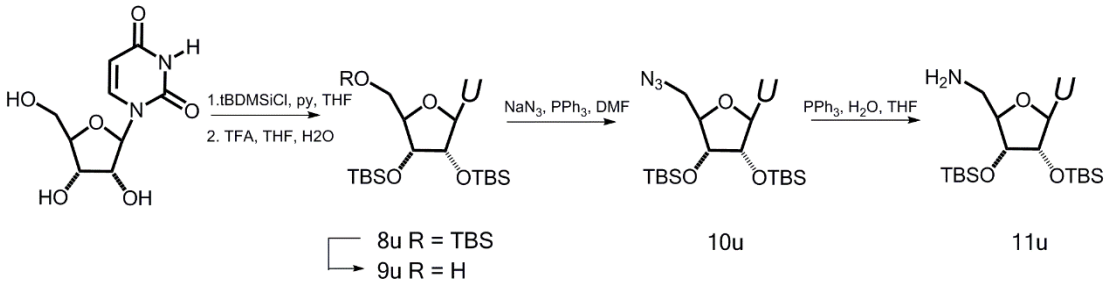


67 mg **6u** (0.11 mmol) was dissolved in 5 mL THF at 0°C. 0.1 mL TBAF (1.0 M in THF) was added at 0°C then stir for 2h at room temperature. Solvent was removed *in vacuuo* and product **7u** was obtained by column chromatography with Hexane : EtOAc = 5:1. NMR indicates product is a mixture of two diastereomers of ring-closure lactol structure with 1:1 ratio. Yield: 92%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 9.44 (s, 0.5H) & 9.34 (s, 0.5H), 7.72 (d, *J* = 8.1 Hz, 0.5H) & 7.60 (d, *J* = 8.1 Hz, 0.5H), 7.46 – 7.36 (m, 6H), 7.35 – 7.20 (m, 9H), 6.17 (d, *J* = 2.6 Hz, 0.5H) & 5.97 (s, 0.5H), 5.75 (dt, *J* = 8.8, 4.4 Hz, 0.5H) & 5.73 (d, *J* = 5.0 Hz, 0.5H), 5.51 (dd, *J* = 8.1, 1.8 Hz, 0.5H) & 5.45 (dd, *J* = 8.1, 1.7 Hz, 0.5H), 4.80 (d, *J* = 6.4 Hz, 0.5H) & 4.66 (dd, *J* = 8.2, 2.5 Hz, 0.5H), 4.58 (m, 0.5H) & 3.93 – 3.88 (m, 1H), 3.47 (dd, *J* = 10.8, 3.1 Hz, 0.5H) & 3.41 (dt, *J* = 14.1, 7.1 Hz, 0.5H), 3.36 (ddd, *J* = 10.4, 9.1, 4.3 Hz, 1H), 3.11 (qd, *J* = 8.5, 2.9 Hz, 0.5H) & 3.00 (q, *J* = 8.2 Hz, 0.5H), 2.04 – 1.99 (m, 2H), 1.93 (m, 2H). HRMS: M+Na$^+$ ($C_{30}H_{28}N_2O_6Na$): calculated: 535.18451, found: 535.18341.

2.4.3 Synthesis of 2'-, 3'- di-TBS - 5'-Amino Uridine (11u)

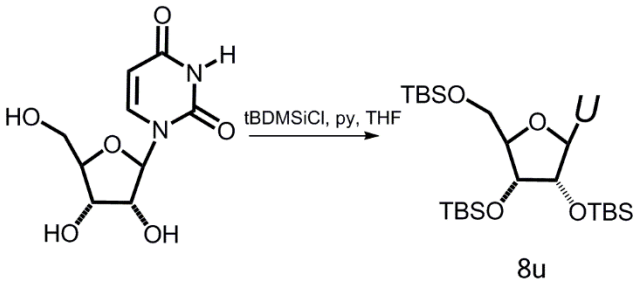
Synthesis of **11u** is outlined in **Scheme 2.5**. After protecting all three hydroxyl groups with TBS, 5'-TBS can be selected deprotected. Incorporation of 5'-amine group is achieved by azidolation instead of previous phthalimide approach, due to easiness of

working-up and higher yield. Azido group is later converted to primary amine.



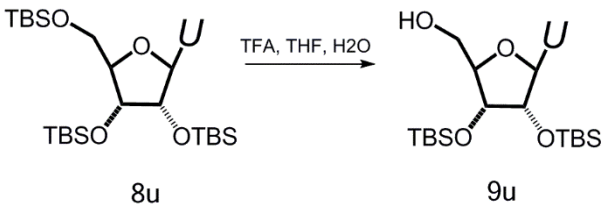
Scheme 2.5 Synthetic route of 2'-, 3'- di-TBS - 5'-Amino Uridine (**11u**)

*1. 2', 3', 5'-O-tris(tert-butyldimethylsilyl)uridine (**8u**)*[12]



488 mg (2 mmol) uridine, 2.11 g TBDMS-Cl (7 eq) and 1.91g imidazole (14 eq) are dissolved in 50 mL anhydrous DMF and stir overnight at room temperature under N_2 protection. Dilute reaction mixture with EtOAc and wash with H_2O three times to get rid of most DMF. Organic layer was combined, dried and concentrated. Residue was purified by silica gel column chromatography with Hexane : EtOAc = 3:1 to obtain the product **8u**. Yield: 93%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8.39 (s, 1H), 8.02 (d, *J* = 8.2 Hz, 1H), 5.87 (d, *J* = 3.8 Hz, 1H), 5.67 (d, *J* = 8.1 Hz, 1H), 4.12 – 4.03 (m, 3H), 3.98 (d, *J* = 11.5 Hz, 1H), 3.75 (d, *J* = 11.6 Hz, 1H), 1.03 – 0.81 (m, 27H), 0.20 – 0.02 (m, 18H). HRMS: M+H$^+$ ($C_{27}H_{55}N_2O_6Si_3$): calculated: 587.33680, found: 587.33840.

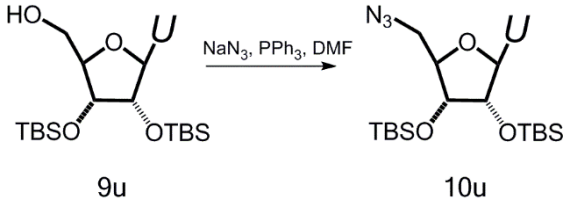
*2. 2', 3'-O-bis(tert-butyldimethylsilyl)uridine (**9u**)[11]*



8u 9u

2.74g of **8u** (4.68 mmol) was dissolved in 40 mL THF at 0°C. H_2O and TFA (10 mL each) were slowly added to make THF : TFA : H_2O = 4:1:1. The solution was stirred for 6h at 0°C under N_2 then neutralized with saturated $NaHCO_3$ solution and dilute with 250mL EtOAc. The organic layer was washed with water then brine, dried and subject to column chromatography with Hexane : EtOAc = 1:1 to give the product **9u**. Yield: 57%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8.16 (s, 1H), 7.59 (d, *J* = 8.1 Hz, 1H), 5.72 (d, *J* = 8.1 Hz, 1H), 5.45 (d, *J* = 5.5 Hz, 1H), 4.56 (t, *J* = 5.0 Hz, 1H), 4.19 – 4.14 (m, 1H), 4.10 – 4.07 (m, 1H), 3.94 (dd, *J* = 12.3, 2.1 Hz, 1H), 3.72 (dd, *J* = 12.3, 2.0 Hz, 1H), 0.97 – 0.84 (m, 18H), 0.12 – 0.00 (m, 12H). HRMS: M+H$^+$ ($C_{21}H_{41}N_2O_6Si_2$) calculated: 473.25032, found: 473.24917.

*3. 5'-Azido-2', 3'-O-bis(tert-butyldimethylsilyl)uridine (**10u**)[13]*

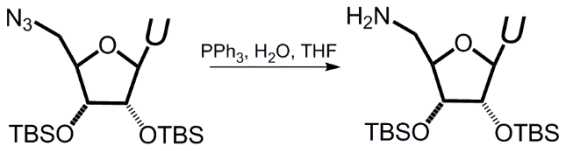


9u 10u

472 mg of **9u** (1 mmol) was dissolved into 10 mL anhydrous DMF. 288 mg of PPh_3 (1.1 eg) and 487 mg of NaN_3 (7.5 eq) were subsequently added into solution. 365 mg CBr_4 (1.1 eq) was then added in portions. Stirred overnight under N_2 protection at room

temperature. 2 mL MeOH was added to quench the reaction. The reaction mixture was filtered, diluted with EtOAc, washed with brine 3 times, dried and concentrated. The residue was purified by column with Hexane : EtOAc = 2:1 to give the product **10u**. Yield: 53%. ^{1}H NMR (400 MHz, cdcl$_3$) δ 8.71 (s, 1H), 7.71 (d, *J* = 8.2 Hz, 1H), 5.78 (d, *J* = 8.2 Hz, 1H), 5.69 (d, *J* = 3.0 Hz, 1H), 4.24 – 4.18 (m, 1H), 4.18 – 4.10 (m, 1H), 3.98 (dd, *J* = 6.1, 4.4 Hz, 1H), 3.84 (dd, *J* = 13.5, 3.1 Hz, 1H), 3.63 (dd, *J* = 13.5, 3.2 Hz, 1H), 0.98 – 0.82 (m, 18H), 0.18 – 0.04 (m, 12H). HRMS: M+H$^+$ ($C_{21}H_{40}N_5O_5Si_2$) calculated: 498.25680, found: 498.25678. IR of ν_{as}(N=N=N): 2105cm^{-1}.

*4. 5'-Amino-2', 3'-O-bis(tert-butyldimethylsilyl)uridine (**11u**)*[14]



10u 11u

248 mg **10u** (0.5 mmol) and 262 mg PPh$_3$ (2 eq) was dissolved into 20 mL THF + 2 mL H$_2$O and reflux overnight. The reaction mixture was concentrated under reduced pressure and purified by column chromatography with DCM : MeOH = 10:1 to give the product **11u**. Yield: 80%.^{1}H NMR (400 MHz, THF-d8) δ 8.00 (d, *J* = 8.1 Hz, 1H), 5.76 (d, *J* = 4.3 Hz, 1H), 5.51 (d, *J* = 8.1 Hz, 1H), 4.33 (t, *J* = 4.4 Hz, 1H), 4.14 (t, *J* = 4.7 Hz, 1H), 3.94 (dd, *J* = 8.6, 4.0 Hz, 1H), 3.04 – 2.82 (m, 2H), 1.03 – 0.81 (m, 18H), 0.11 (dd, *J* = 9.0, 3.5 Hz, 12H). HRMS: M+H$^+$ ($C_{21}H_{42}N_3O_5Si_2$): calculated: 472.26630, found: 472.26553.

2.4.4 Mixing of 7u with 11u for N, O-acetal Condensation

Free amine substrates of **11u** and the aldehyde substrate **11u** are dissolved individually in THF-d_8. Then the amine substrate is added into the aldehyde solution closely to 0.99: 1 with 7.5 mM concentration for each substrate. The reaction progress is followed by 1H NMR spectroscopy and recorded on an INOVA 600 NMR spectrometer.

For hydrolysis study, pD = 4.5 acetate-d_4 buffer, pD = 5.5 pyridine-d_5 buffer, pD = 7.0 KD_2PO_4 buffer, pD = 7.5 KD_2PO_4 buffer were prepared with D_2O. Deuterium chloride solution (35% wt in D_2O) and Sodium Deuterioxide (40% in D_2O) were used to adjust the pD on a Fisher Scientific Accumet Basic AB15 pH meter using the relationship[15]: pD = pH(meter) + 0.5. Ionic strength is controlled to be 100 mM over all buffers using NaCl.

2.4.5 Preparation of Free Amine Substrates by Neutralizing Thr Ester HCl Salts

Before mixing aldehyde substrate with the amine substrate for acetal condensation, the Thr ester hydrochloride salts (50 mg) are neutralized by 30% NH_4OH solution (1 mL) and the resulting water and excess NH_4OH is removed in vacuum at room temperature. The resulting residue is then dissolved in acetonitrile (1 mL) in a microcentrifuge tube and centrifuged at 9,000 ×g for 3 minutes to separate the precipitated NH_4Cl salt at the bottom, and then the supernatant is transferred and dried *in vacuuo* to yield the free amine of Thr ester residues.

2.4.6 Mixing of Thr Ester with Phe Aldehyde for N, O -acetal Product

The ^{1}H NMR spectra of the free amine and the aldehyde substrates are recorded on an INOVA 600 NMR spectrometer as follows:

L-Thr-methyl ester: ^{1}H NMR (600 MHz, cd_3cn) δ 3.91 (qd, *J* = 6.4, 4.4 Hz, 1H), 3.69 (m, 3H), 3.33 (d, *J* = 4.4 Hz, 1H), 1.17 (d, *J* = 6.5 Hz, 1H).

N-Boc-L-Thr-CHO: ^{1}H NMR (600 MHz, cd_3cn) δ 9.58 (s, 1H), 7.34 – 7.27 (m, 2H), 7.23 (dd, *J* = 15.4, 7.5 Hz, 3H), 5.59 (s, 1H), 4.19 (dd, *J* = 13.4, 8.5 Hz, 1H), 3.17 (dd, *J* = 14.1, 5.0 Hz, 1H), 2.83 (dd, *J* = 14.0, 9.5 Hz, 1H), 1.35 (s, 9H).

Free amine substrates of Thr residues and the aldehyde substrate N-Boc-L-Phe-CHO are dissolved individually in solvents. Then the amine substrate is added into the aldehyde solution closely to 0.99: 1 with 10 mM concentration for each substrate. The reaction progress is followed by ^{1}H NMR spectroscopy and recorded on an INOVA 600 NMR spectrometer or INOVA 400 NMR spectrometer.

References

1. (a) Goodwin, J. T.; Lynn, D. G., Template-directed synthesis: use of a reversible reaction. *Journal of the American Chemical Society* **1992,** *114* (23), 9197-9198; (b) Li, X.; Zhan, Z.-Y. J.; Knipe, R.; Lynn, D. G., DNA-Catalyzed Polymerization. *Journal of the American Chemical Society* **2002,** *124* (5), 746-747.

2. Li, X.; Hernandez, A. F.; Grover, M. A.; Hud, N. V.; Lynn, D. G., Step-growth control in template-directed polymerization. *Heterocycles* **2011,** *82* (2), 1477-1488.

3. Bergmann, E. D., The Oxazolidines. *Chemical Reviews* **1953,** *53* (2), 309-352.

4. (a) Isbell, H. S.; Frush, H. L., Effect of pH in the mutarotation and hydrolysis of clycosylamines. *Journal of the American Chemical Society* **1950,** *72* (2), 1043-1044; (b) Fife, T. H.; Hagopian, L., Oxazolidine hydrolysis. Participation of solvent and buffer in ring opening. *Journal of the American Chemical Society* **1968,** *90* (4), 1007-1014; (c) Fife, T. H.; Hutchins, J. E. C., General-acid-catalyzed ring opening of oxazolidines. Hydrolysis of 2-[4-(dimethylamino)styryl]-N-phenyl-1,3-oxazolidine. *The Journal of Organic Chemistry* **1980,** *45* (11), 2099-2104.

5. (a) Adler-Abramovich, L.; Reches, M.; Sedman, V. L.; Allen, S.; Tendler, S. J. B.; Gazit, E., Thermal and Chemical Stability of Diphenylalanine Peptide Nanotubes: Implications for Nanotechnological Applications. *Langmuir* **2006,** *22* (3), 1313-1320; (b) Meital, R.; Ehud, G., Designed aromatic homo-dipeptides: formation of ordered nanostructures and potential nanotechnological applications. *Physical Biology* **2006,** *3* (1), S10; (c) Scanlon, S.; Aggeli, A., Self-assembling peptide nanotubes. *Nano Today* **2008,** *3* (3–4), 22-30.

6. Seebach, D.; Amatsch, B.; Amstutz, R.; Beck, A. K.; Doler, M.; Egli, M.; Fitzi, R.; Gautschi, M.; Herradón, B.; Hidber, P. C.; Irwin, J. J.; Locher, R.; Maestro, M.; Maetzke, T.; Mouriño, A.; Pfammatter, E.; Plattner, D. A.; Schickli, C.; Schweizer, W. B.; Seiler, P.; Stucky, G.; Petter, W.; Escalante, J.; Juaristi, E.; Quintana, D.; Miravitlles, C.; Molins, E., Structure and Reactivity of Five- and Six-Ring N, N-, N, O-, and O, O-

acetals: A lesson in allylic 1, 3-strain (A1, 3 strain). *Helvetica Chimica Acta* **1992,** *75* (3), 913-934.

7. Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S., Mechanosensitive Self-Replication Driven by Self-Organization. *Science* **2010,** *327* (5972), 1502-1506.

8. Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K., New catalysts and procedures for the dimethoxytritylation and selective silylation of ribonucleosides. *Canadian Journal of Chemistry* **1982,** *60* (9), 1106-1113.

9. Meier, C.; Huynh-Dinh, T., Improved Conversion of Adenosine to 3′-Deoxyadenosine. *Synlett* **1991,** *1991* (04), 227-228.

10. Fiandor, J.; Y. Tam, S., Synthesis of 3′-deoxy-3′-(2-propynyl)thymidine and 3′-cyanomethyl-3′-deoxythymidine, analogs of AZT. *Tetrahedron Letters* **1990,** *31* (5), 597-600.

11. Zhu, X.-F.; Williams, H. J.; Scott, A. I., Facile and highly selective 5[prime or minute]-desilylation of multisilylated nucleosides. *Journal of the Chemical Society, Perkin Transactions 1* **2000,** (15), 2305-2306.

12. Ogilvie, K. K.; Beaucage, S. L.; Schifman, A. L.; Theriault, N. Y.; Sadana, K. L., The synthesis of oligoribonucleotides. II. The use of silyl protecting groups in nucleoside and nucleotide chemistry. VII. *Canadian Journal of Chemistry* **1978,** *56* (21), 2768-2780.

13. (a) Hata, T.; Yamamoto, I.; Sekine, M., A simple method for the synthesis of 5'-azido-5'-deoxyribonucleosides. *Chemistry Letters* **1975,** *4* (9), 977-980; (b) Yamamoto,

I.; Sekine, M.; Hata, T., One-step synthesis of 5'-azido-nucleosides. *Journal of the Chemical Society, Perkin Transactions 1* **1980,** (0), 306-310.

14. Mungall, W. S.; Greene, G. L.; Heavner, G. A.; Letsinger, R. L., Nucleotide chemistry. XX. Use of the azido group in the synthesis of 5'-terminal aminodeoxythymidine oligonucleotides. *The Journal of Organic Chemistry* **1975,** *40* (11), 1659-1662.

15. Glasoe, P. K.; Long, F. A., Use of glass electrodes to measure acidities in Deuterium Oxide. *The Journal of Physical Chemistry* **1960,** *64* (1), 188-190.

Chapter Three

DNA Templated rANP Networks: Formation of DNA/rANP Duplexes

3.1 Introduction

In biology, the highly accurate replication of sequence information encoded on DNA is accomplished by template-directed polymerization with sophisticated polymerases. The power of evolution in this process is greatly appreciated by researchers, and decades of efforts have been put into creating simple chemical systems to mimic the process, but most previous efforts suffer from low fidelity, possibly due to lack of reversibility as discussed in Chapter 1. In our laboratory[1], by replacing phosphodiester with reversible imine to create a DCN, we successfully achieved a template-directed replication under thermodynamic control, and greatly improved information copying fidelity. However, the use of reductive amination traps the DCN in an irreversible fashion and eliminates the dynamic property of the network.

To construct DCNs with full potential to responds to different environmental inputs and templates, it is desirable to fully maintain the dynamic property. It has also been postulated that the prebiotic selection of biopolymers might have taken advantage of reversible backbones to provide a means to recycle, repair and transform in response to environment[2].

Ghadiri et al[3] reported the first fully dynamic system that can copy sequence information from DNA template to new polymers (**Fig 3.1**). In this system, nucleobase thioester forms reversible disulfide linkages with a poly-cysteine backbone to generate thioester peptide nucleic acids (tPNA) (**Fig 3.1A**). In the presence of a DNA template $(dT)_{20}$, adenine thioester is selectively anchored on the backbone, over-competes other nucleobase thioesters, and form tPNA with complementary sequences (**Fig 3.1B**). Apart from this DNA templated dynamic base-filling, DNA template directed reversible ligation using disulfide exchange[4] and boronic ester condensation[5] have also been reported. However, no attempt of using monomeric network building block to achieve DNA template directed sequence-specific replication have been successful[6].

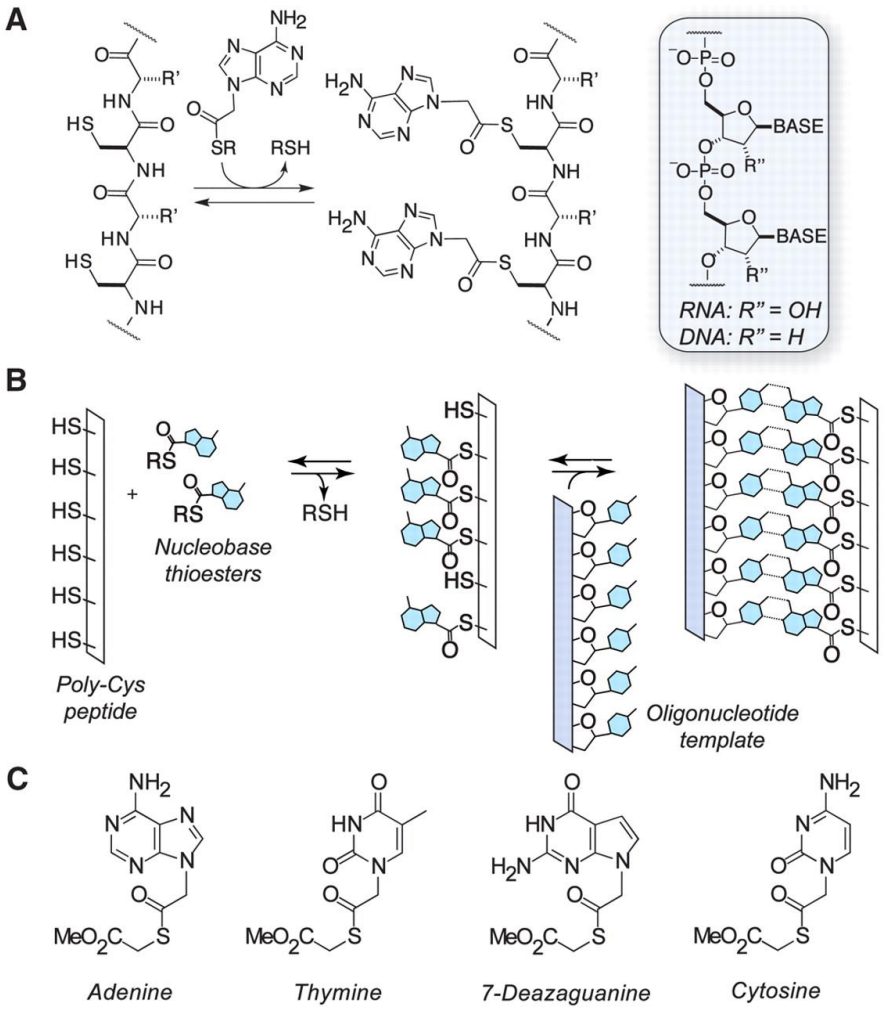


Fig 3.1 DNA template directed dynamic tPNA oligomerization via reversible base-filling[3]. (**A**) Mechanism for reversible covalent oligomer assembly via anchoring of thioester-derived building blocks onto an oligopeptide backbone. (**B**) Illustration of oligomer assembly and binding to a complementary oligonucleotide template. (**C**) Structures of nucleobase thioester monomers for construction of network.

Here, to create DCNs with the potential to achieve DNA template directed replication from monomers, we propose to use ribose-ANP monomer (rANP) instead of deoxyribose-ANP (dANP) to construct DCNs and copy information from a DNA template. As discussed in Chapter 2, the 2'-OH group on ribose is well-positioned to trap the imine in a reversible fashion and generate an N,O-acetal, a dynamic linkage

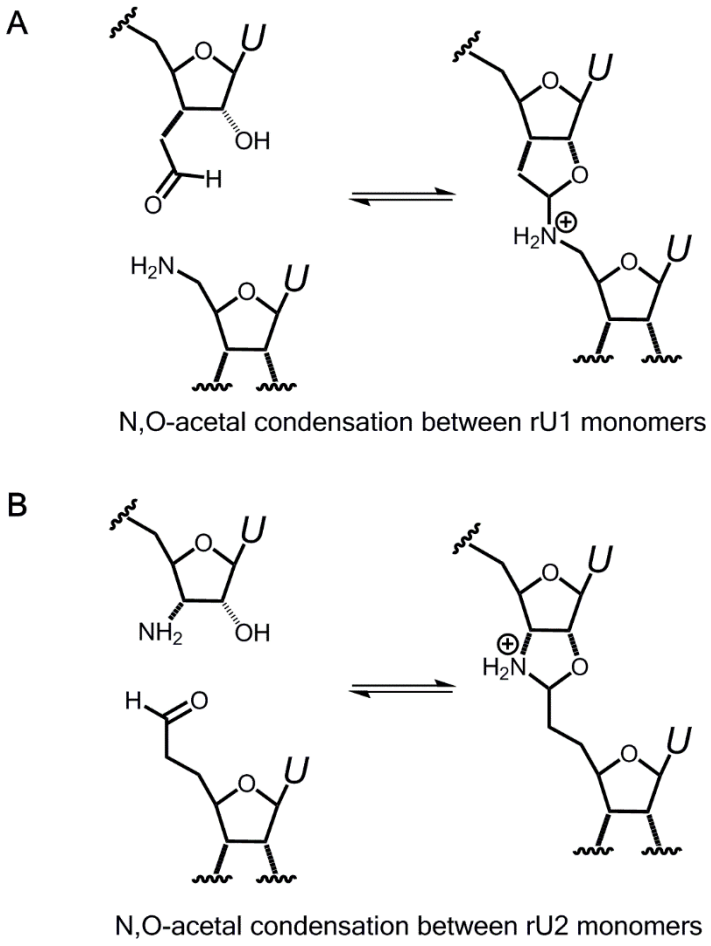
fused as a five member ring structure. Not only can N,O-acetal linkage remains dynamic and responsive towards environment changes, but its fused bicyclic structure can further increase the organization on DNA template[7], due to less degree of freedom from the restricted conformation flexibility observed in locked nucleic acids (LNA)[8].

3.2 Results and Discussion

3.2.1 Synthesis of rANP Monomer Building Blocks

Uridine-derived monomers are selected as the building blocks for construction of dynamic chemical networks on poly-dA oligonucleotide templates for the selection of rANP networks. Uridine-based building block benefits from a relative easy synthesis. Another benefit of A-U base-pairing relative to G-C is to avoid the complications of G-quadruplex structure formation in the presence of salt.

As summarized in **Scheme 3.1**, two geometries of rANP monomers, 5'-amino-3'-aldehyde-uridine (rU1) and 3'-amino-5'-aldehyde-uridine (rU2) are both prepared. In Chapter 2 we found that for rU1, the proximity of 2'-hydroxyl group and the 3'-allylaldehyde results in formation of lactol structure, and complicates the reaction. Positioning the aldehyde at 5'- position and the amine at 3'- position avoids this complication. The design of both monomers presents a six atom distance linkage, same as the native phosphodiester, to maximize the likelihood of effective template-substrate binding.



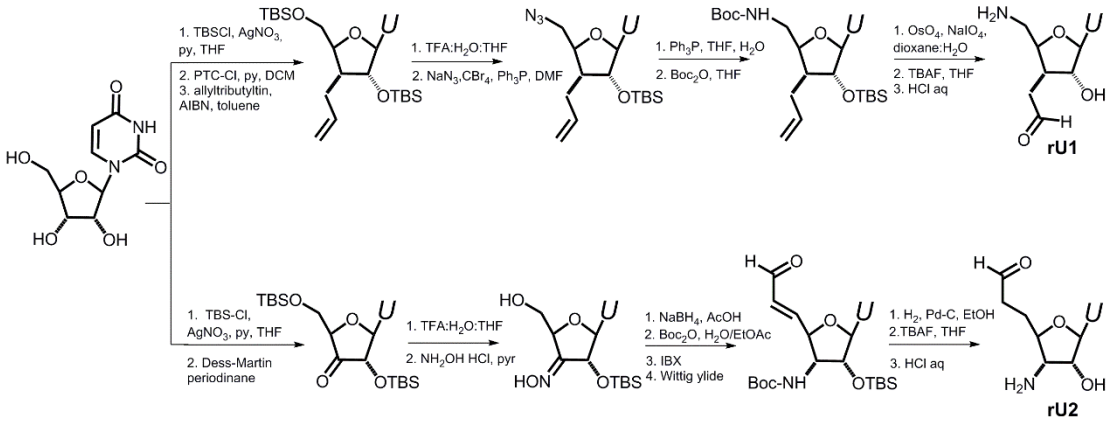
Scheme 3.1 N,O-acetal condensation of rANP monomers rU1 and rU2. (**A**) rU1 with 5'-amine and 3'-aldehyde generates reversible N,O-acetal linkage with O in cyclic structure and N out of cyclic structure. (**B**) rU2 with 3'-amine and 5'-aldehyde generates N,O-acetal linkage with N and O in cyclic structure.

Synthetic routes to create rU1 and rU2 monomers are summarized in **Scheme 3.2**. Both monomer synthetic routes start from commercially available uridine. For rU1, uridine's 2'- and 5'-OH groups are selectively protected by TBDMS using $AgNO_3$ as catalyst[9]. The free 3'-OH group is activated by phenylthiocarbonyl[10], then the carbon skeleton of 3'-aldehyde moiety is introduced by a stereoselective radical allylation reaction initiated by AIBN[11], and masked as an olefin. The bulky 5'-TBDMS protecting group prevents the radical coupling to attack the β face of ribose ring, giving a single

diastereomer with selectivity > 9:1. The 5'-TBDMS group can be selectively removed using TFA: H_2O : THF = 1:1:2 (v/v/v) mixed solvent[12]. Incorporation of 5'-amine group is achieved by azidolation[13] instead of the previous phthalimide approach, due to the easiness of working-up and higher yield. The azido group is later converted to primary amine[14] and protected with Boc group. The 3'-olefin is oxidized to the aldehyde by OsO_4 in the presence of $NaIO_4$[11a]. The final rU1 is generated by deprotecting Boc group *in situ* using 1 M HCl aqueous solution at room temperature.

For the rU2 monomer, after selectively protecting 2'- and 5'- OH with TBDMS, the 3'-position is oxidized to the ketone by Dess-Martin oxidation[15], converted into the oxime[16], and the C=N double bond of oxime product is reduced by $NaBH_4$ acetate solution to generate the hydroxylamine[17]. This reduction is stereoselective if reaction temperature is carefully controlled under 5ºC. Even though 5ºC is below the freezing point of solvent acetate, high concentrations of $NaBH_4$ effectively lower the freezing point. In addition, adding a small volume of THF as co-solvent helps avoid freezing without compromising stereoselectivity. The hydroxylamine is further converted to the amine using Pd/C catalyzed hydrogenation, and protected with Boc group. The 5'-aldehyde is introduced by oxidation of 5'-hydroxyl group using IBX oxidation[18], followed by Wittig reaction[19]. Previous efforts to use Dess-Martin reagent or Swern oxidation resulted in low yield of the subsequent Wittig reaction. We found that it was due to byproducts formed from those reactions hampering the following Wittig reaction. The 5'-aldehyde uridine is not stable and cannot be purified using silica gel

chromatography. The IBX reaction greatly simplifies the reaction and improves yield because the byproduct is not soluble and can be easily removed by filtration. Selective reduction of the C=C double bond on α, β-unsaturated aldehyde is accomplished by Pd/C reduction[20]. The TBDMS is deprotected by classical TBAF approach, and as with rU1, the final rU2 is generated by deprotecting the Boc group *in situ* using 1M HCl aqueous solution at room temperature.



Scheme 3.2 Summary of synthetic routes for rU1 and rU2 monomers

3.2.2 $(dA)_{16}$ Template in rANP Networks Induces DNA/rANP Hybrid Duplexes

The DNA templated rANP network reaction is illustrated in **Fig 3.2**. For the rU1 network, rU1 monomer and DNA template $(dA)_{16}$ were incubated together with the ratio of template $(dA)_{16}$ = 1 mM and monomer rU1 = 16 mM. Acetate buffer with pH = 4.5 (concentration:50 mM, ionic strength:100 mM) was used and the template was slowly heated to 75^{o}C, then cooled down to room temperature to eliminate any possible folding structures, and the rANP substrate was added. The final network solution was

incubated at 4°C for 2 weeks.

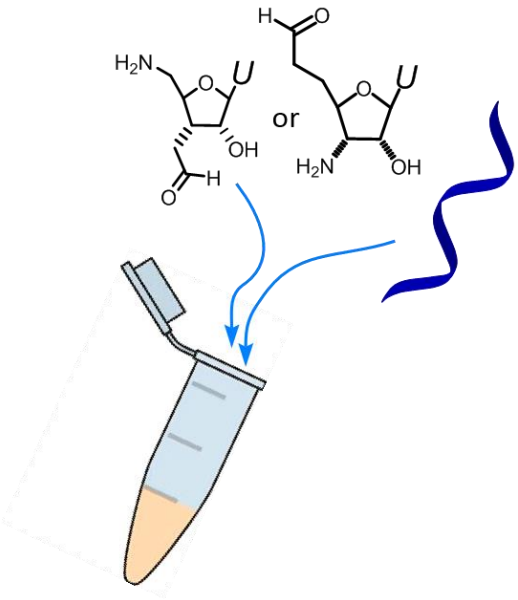


Fig 3.2 Illustration of DNA templated rANP networks. DNA template and rANP monomers are incubated in aqueous buffer solution in stoichiometric ratio (template $(dA)_{16}$=1 mM, rU1 monomer = 16mM) at 4°C.

As shown in **Fig 3.3**, $(dA)_{16}$ template exhibits a negative band at 248 nm and a strong positive band at 217 nm, characteristic of a lack of secondary structure[5a, 21], and rU1 monomer alone gives no obvious CD signals. However, for the $(dA)_{16}$ templated rU1 network, after 30 days of incubation, a drastic CD signal change was observed after 14 days of incubation. CD displays a strong positive band at 265 nm and a negative band at 245 nm, diagnostic of B-form duplex conformation[22]. The formation of B-form duplex is consistent with oligomerization of rANP and hybridization with template.

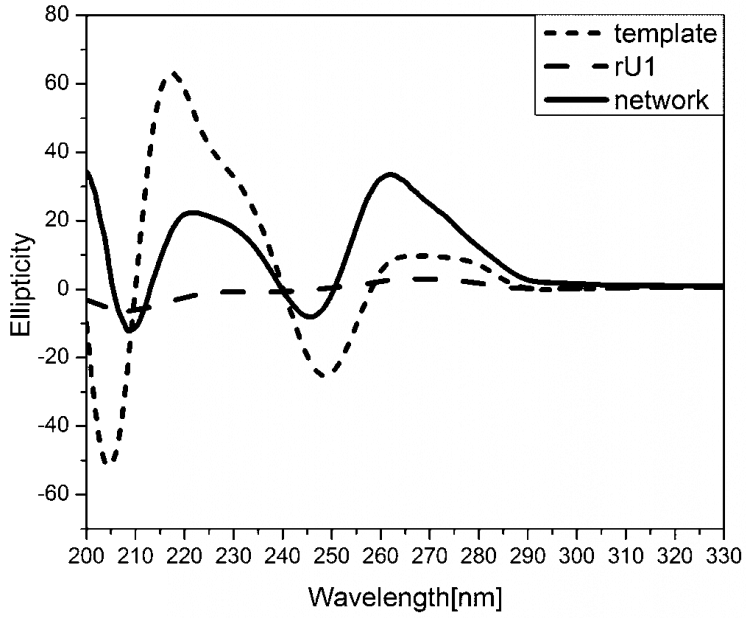


Fig 3.3 Circular Dichroism of $(dA)_{16}$ templated rU1 network and B-form poly[d(AT)] duplex. Circular Dichroism of $(dA)_{16}$ templated rU1 network. Short dashed line: CD of $(dA)_{16}$ template; dashed line: CD of rU1 monomer; solid line: CD of $(dA)_{16}$ templated rU1 network after 30 days of incubation. Condition: $(dA)_{16}$ = 1 mM, rU1 = 16 mM, buffer: pH = 4.5 acetate, c = 50 mM, I = 100 mM. Incubation temperature: 4°C.

The templated rU2 network requires lower substrate concentration and higher ionic strength for the putative B-form duplex to form. On mixing of template and monomer under the same conditions (($dA)_{16}$ = 1 mM, rU2 = 16 mM with pH = 4.5 acetate buffer (c = 50 mM, I = 100 mM), precipitant formed immediately, likely due to non-specific electrostatic interaction between negatively charged template and positively charged rANP monomers. To minimize the electrostatic interaction, the concentration of substrates was reduced to to 10%, with $(dA)_{16}$ = 0.1 mM, rU2 = 1.6 mM; and the ionic strength of the buffer was raised to 500 mM instead of 100 mM using NaCl. Higher ionic strength solutions shield non-specific electrostatic interactions, favoring Watson-

Crick base-pair recognition, and as shown in **Fig 3.4**, under such condition, a strong CD ellipticity similar to templated rU1 network with a positive band at 265 nm and a negative band at 245 nm was observed.

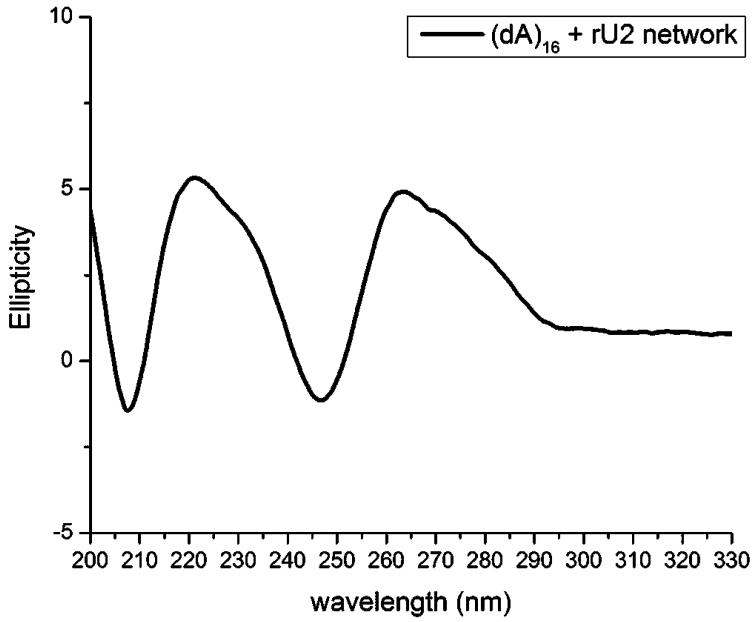


Fig 3.4 Circular Dichroism of $(dA)_{16}$ templated rU2 network. Templated network exhibits B-form duplex type of signal, with a strong positive band at 265 nm and a negative band at 245 nm, similar to templated rU1 network. Condition: $(dA)_{16}$ = 0.1 mM, rU2 = 1.6 mM, buffer: pH = 4.5 acetate, c = 50 mM, I = 500 mM. Incubation temperature: 4°C.

To verify that the duplex formation we observed in the DNA templated rANP network is driven by specific Watson-Crick base pairs, rather than other non-specific interactions, we examined whether $(dT)_{16}$ template can impact the behavior of rANP networks. In both $(dT)_{16}$ templated rU1 and rU2 networks, after 30 days of incubation, no change of ellipticity was observed, confirming the recognition between rANP and DNA template is driven by specific Watson-Crick base pairs, establishing digital information transfer within templated rANP networks.

HPLC separation of rANP oligomers in templated networks were unsuccessful, possibly due to product hydrolysis. Adjusting HPLC condition to pH = 4.5, the pH for network incubation, resulted in extremely short retention time and poor resolution. Therefore, the templated networks were directly applied for MALDI-TOF analysis. As shown in **Fig 3.5A,** with DHB matrix, an m/z = 3887 peak was observed for templated rU2 network, consistent with a 15-mer ANP oligomer ($M + AcOH + H_2O + Na^+$). This 15-mer oligomer was further confirmed by a bivalent ionization signal observed with CHCA matrix (m/z = 1940, $M + AcOH + 2Na^+$, data not shown). Similarly, as shown in **Fig 3.5B**, for templated rU1 network a 15-mer ANP oligomer peak was also observed in CHCA matrix, with m/z = 3986 ($M + 3AcOH + Na^+$). It is unclear why 15-mer instead of 16-mer were detected. Possible explanation includes decomposition in gas phase, or incomplete oligomerization in templated reaction.

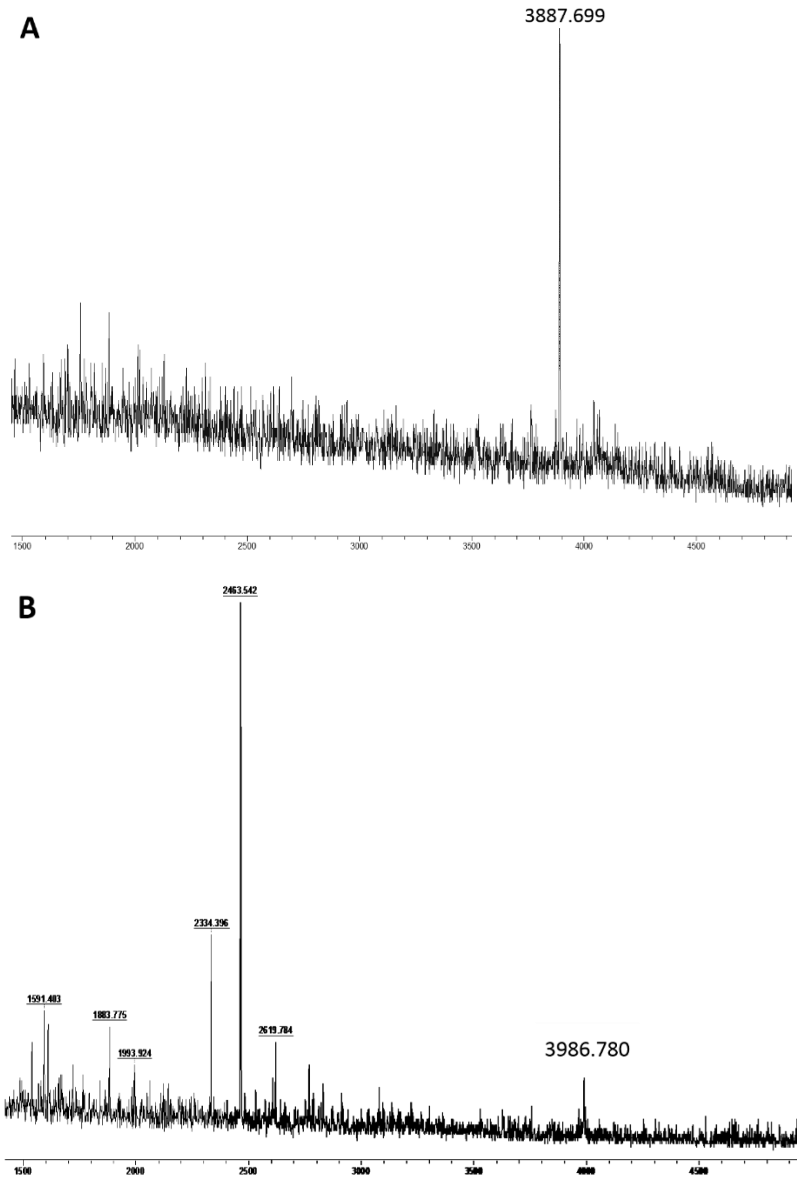


Fig 3.5 MALDI-TOF analysis of templated rU1 and rU2 networks detects 15-mer oligomers. (**A**) MALDI-TOF of templated rU1 network, with DHB matrix. (**B**) MALDI-TOF of templated rU2 network, with CHCA matrix.

3.2.3 Kinetics of Duplex Formation in Templated rANP Networks

To better resolve the transformation, the kinetics were monitored by CD over time. As shown in **Fig 3.6A**, for templated rU1 network, the signature for B-form duplex formation reached completion after 5 days, while for templated rU2 network, even

though the concentration is 10 times lower than rU1 network, duplex formation was completed within 48 hours (**Fig 3.6B**). Since the structure of rU2 monomer, with 5'-aldehyde away from 2'-hydroxyl, avoids the lactol formation complicating the rU1 reaction, the higher effective concentration of reactive aldehyde likely contributes to the faster kinetics of duplex formation in the templated rU2 network.

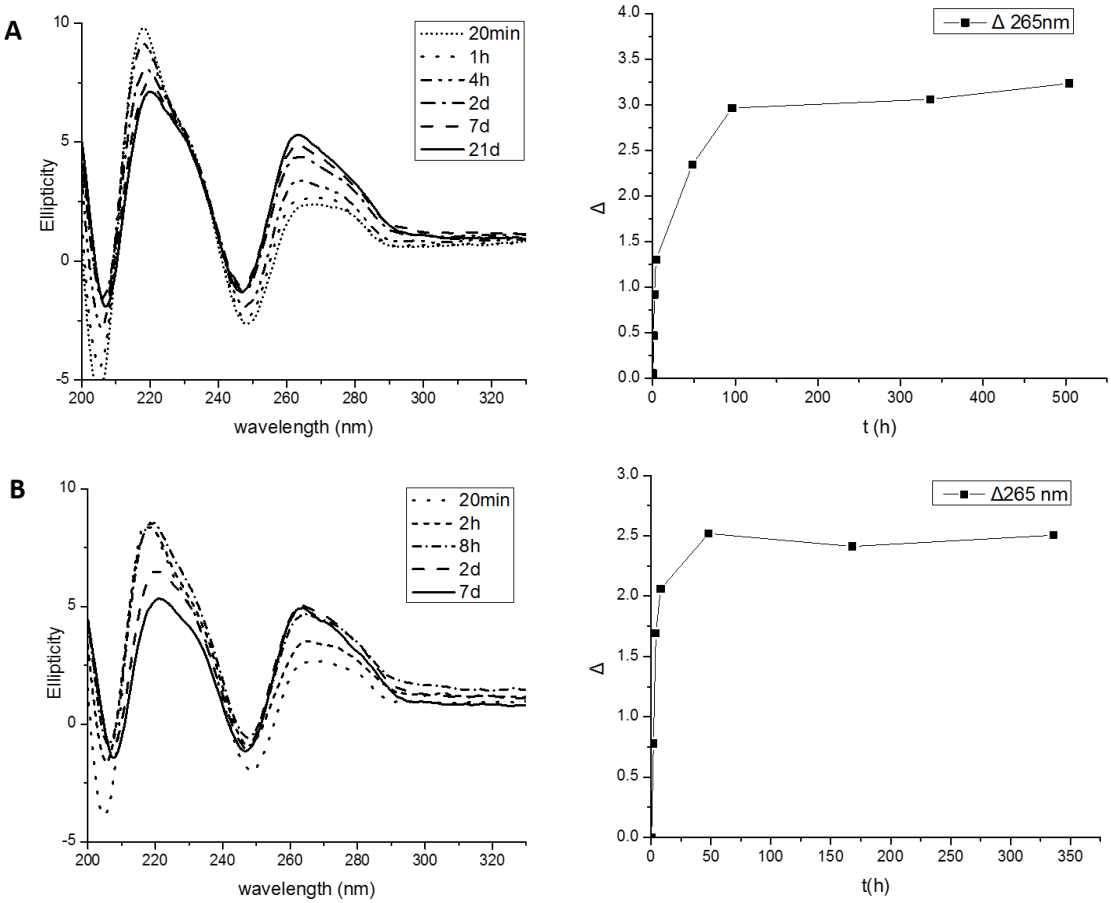


Fig 3.6 Kinetics of $(dA)_{16}$ templated rU1 and rU2 networks. (**A**) Kinetics of $(dA)_{16}$ templated rU1 network monitored by the increase of 265 nm ellipticity. The transition is completed after 5 days. Condition: $(dA)_{16}$ = 1 mM, rU1 = 16 mM, buffer: pH = 4.5 acetate, c = 50 mM, I = 100 mM. 4°C. (**B**) Kinetics of (dA)16 templated rU2 network monitored by increase of 265 nm positive band on CD. Assembly is completed after 2 days. Condition: $(dA)_{16}$ = 0.1 mM, rU2 = 1.6 mM, buffer: pH = 4.5 acetate, c = 50 mM, I = 500 mM. 4°C.

3.2.4 Thermal Melting Experiments Suggest High Duplex Binding Affinity

The CD signal intensity at 265 nm was recorded with incrementally increased temperature to determine the thermodynamic stability of the presumptuous duplex structure. For the $(dA)_{16}$: rU1 = 1 mM : 16 mM network in 100 mM NaCl, as shown in **Fig 3.7A**, the thermal melting curve gives a highly cooperative transition at T_m = 67^{o}C, 12^{o}C higher than control $(dA)_{16}/(dT)_{16}$ duplex. While for $(dA)_{16}$: rU2 = 0.1 mM : 1.6 mM network in 500 mM NaCl, the T_m is so high a T_m is not determined (**Fig 3.7B**).

Such high T_m and cooperativity indicates strong binding between template and rANP oligomers. The high thermal stability is likely due to the backbone charge attraction instead of repulsion in native DNA duplexes, as well as the restricted conformation flexibility from the N,O-acetal five-member ring.

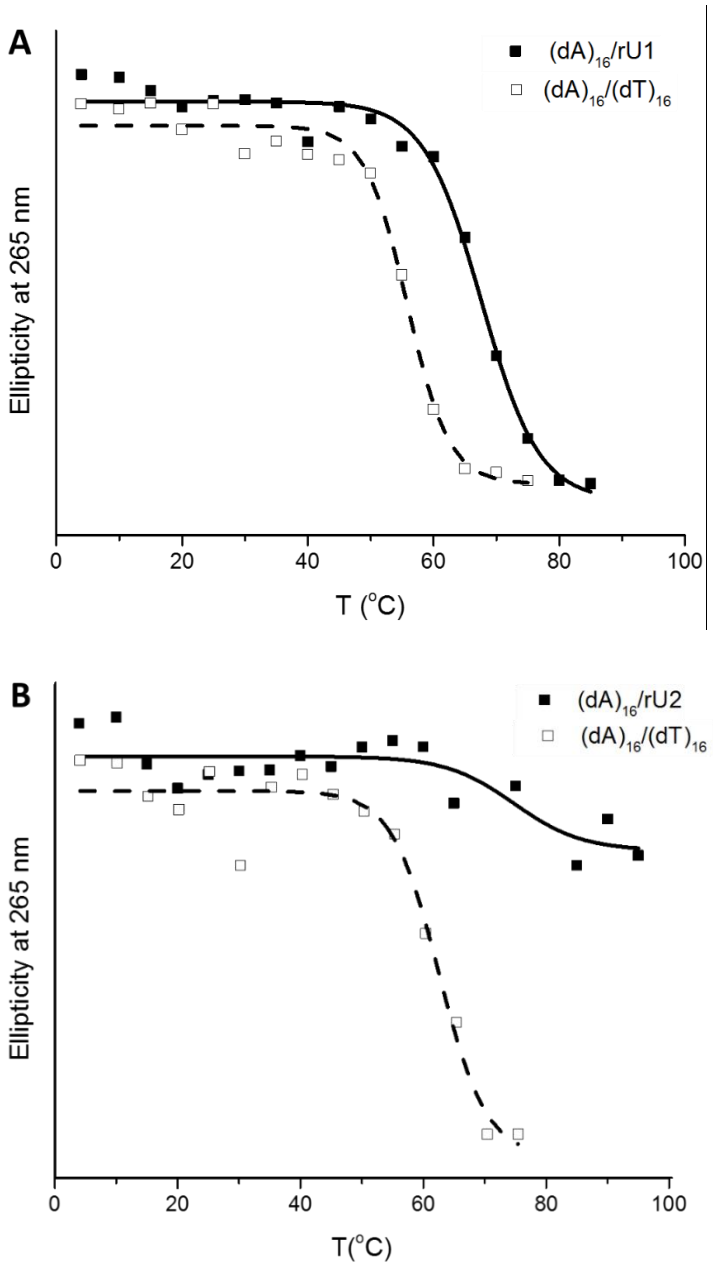


Fig 3.7 Thermal melting temperature experiment of $(dA)_{16}$ templated rU1 and rU2 networks. Melting temperature is monitored by 265 nm band on CD over increasing temperature. (**A**) DNA/ANP duplex generated in rU1 network gives a high T_m at 67°C with high cooperativity. Control $(dA)_{16}/(dT)_{16}$ duplex: T_m = 55°C. (**B**) DNA/ANP duplex in rU2 network has a very high T_m that high temperature plateau cannot be reached. Control $(dA)_{16}/(dT)_{16}$ duplex: T_m = 62°C

3.2.5 Impact of Template Length on Duplex Formation

Template-directed replication rates are correlated with not only substrate/template binding affinity, but also the number of template binding sites. For a homopolymer $(dA)_n$ template, it provides n binding sites for rU monomer that contribute to the first N,O-acetal ligation event. The more binding sites available, the more the binding equilibrium is biased towards duplex formation. To use this procedure to evaluate duplex formation, g a shorter template $(dA)_8$ was used to template the rANP networks. The condition for templated rU1 network is $(dA)_8$ = 1 mM, rU1 = 8 mM, in pH = 4.5 acetate buffer (c=50 mM, I=100 mM) and incubated at 4ºC. Similarly, for templated rU2 network, $(dA)_8$ = 0.1 mM, rU1 = 0.8 mM, in pH = 4.5 acetate buffer (c=50 mM, I=500 mM) is incubated at 4ºC for 30 days.

After 30 days of incubation, neither of two $(dA)_8$ templated networks indicates any formation of B-form duplex structure. The CD signals show unassembled free $(dA)_8$ template. This result suggests that for shorter templates, $(dA)_8$ cannot provide sufficient amount of binding site to reach the threshold for catalyzing on-template N,O-acetal condensation. This result is in agreement with literature reports[23], which shows that short DNA template was found unable to provide effect binding to monomeric substrate, while longer template is highly efficient.

3.2.6 Impact of pH on DNA Templated ANP Oligomerization

As was discussed in Chapter 2, reactivity of N,O-acetal condensation and stability of

N,O-acetal are pH-dependent. In general, N,O-acetal condensation shows a bell-shape pH-dependence curve[24]: slightly acidic pH facilitates condensation by activating carbonyl carbon, and this is the rationale to choose pH = 4.5 in DNA templated networks. Basic pH is also favorable, because deprotonated primary amine is more electrophilic. However, pK_a of adenosine and uridine (3.5 for N6 of adenosine and 9.2 for N3 of uridine) limits the pH range for effective DNA/ANP duplex formation in templated networks. Besides, rANP monomers might possess different pK_a than native uridine, giving their different backbone structures.

Therefore, we tested the DNA templated rANP polymerization in pH = 6.5 MES buffer and pH = 8.5 TAPS buffer, with all other conditions maintained the same. For both rU1 and rU2 templated networks, under both pH = 6.5 and pH = 8.5, no duplex formation can be detected by CD, suggesting no oligomerization or hybridization occurred. For pH = 6.5, the N,O-condensation reaction rate is expected to be low, consistent with the observation that no duplex formation. While pH 8.5 is very close to pK_a of Uridine (9.2), therefore it is likely that pH 8.5 is out of the pH range for efficient Watscon-Crick base-pairing between $(dA)_{16}$ and rU monomers, hampering efficient oligomerization to take place. Another possibility is that for pH above 6.5, N,O-acetal is at a different protonation state (pK_a range of N,O-acetal[25] is 5.2 – 6.9), and results in unfavorable configuration for formation of duplexes.

3.2.7 Impact of Ionic Strength on DNA Templated ANP Oligomerization

Solvent ionic strength impacts the stability of DNA duplexes as well as DNA/ANP hybrid duplexes. In general, higher ionic strength can shield the electrostatic interaction between backbone charges to reduce non-specific interaction. Higher ionic strength also increases solvent polarity, in return enhancing hydrophobic stacking between base-pairs, a critical factor to stabilize duplex structures and favors the templating reaction. On the other hand, increased solvent polarity also weakens hydrogen bonding interaction between base-pairing. Therefore, the balance between these factors dictates the ideal ionic strength to facilitate templated duplex formation in rANP networks.

As we discovered previously, for templated rU2 network, higher ionic strength (I = 500 mM) is critical to reduce the non-specific interaction. Lower ionic strength (I = 100 mM) results in non-specific interaction between template and substrate to form precipitants. Here, we also explored the impact of ionic strength on rU1 networks.

For templated rU1 network, higher ionic strength facilitates kinetics of duplex formation. As shown in **Fig 3.8**, in I = 500 mM acetate pH4.5 buffer, the kinetics of duplex formation in network took place faster compared to network in I = 100 mM, with transition completed within 2 days for I = 500 mM network and 5 days for I = 100 mM network. Higher salt concentration is also found to decrease critical concentration for duplex formation. As shown in **Fig 3.9**, in I = 100mM buffer, if the concentration for template and substrates was lowered to $(dA)_{16}$ = 0.1 mM and rU1 = 1.6 mM

respectively, no assembly formation can be detected after 35 days of incubation, likely due to lack of efficient binding under lower concentration. However, for I = 500 mM buffer, duplex structure can be detected at the same low concentrations, though at a substantially lower rate compare to templated rU2 network. This result suggests that in DNA templated rU1 replication, higher ionic strength is favorable.

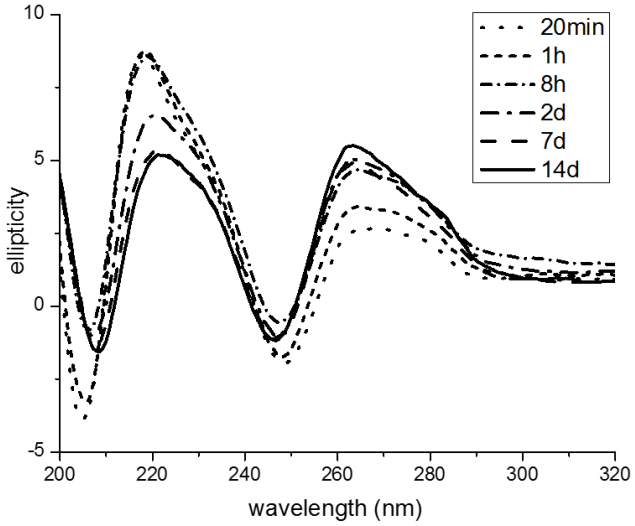


Fig 3.8 Higher ionic strength facilitates kinetics of duplex formation in templated rU1 network. In I = 100 mM networks, duplex formation completes after 5 days, while in I = 500 mM, duplex formation reaches completion after 2 days. Conditions: $(dA)_{16}$ = 1 mM, rU1 = 16 mM, buffer: pH = 4.5 acetate, c = 50 mM, I = 500 mM. Incubation temperature: 4^{o}C.

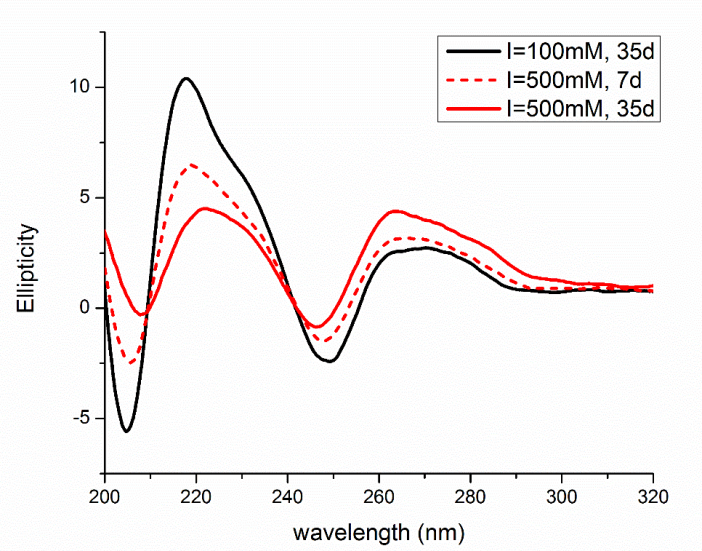


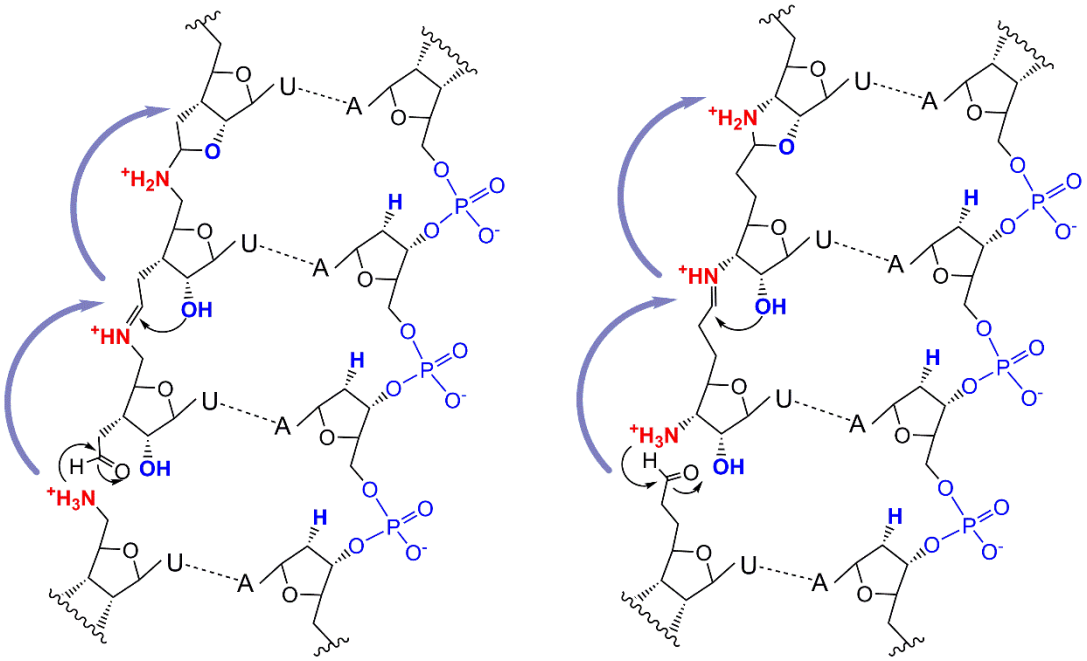
Fig 3.9 Higher ionic strength lowers critical duplex assembly concentration for templated rU1 network. For $(dA)_{16}$ = 0.1 mM and rU1 = 1.6 mM, under I = 100 mM (black solid line), no duplex formation can be detected after 35 days of incubation. While same substrate concentration under I = 500 mM, conformation change can be observed (red dashed line) and completes into B-form duplex after 35 days (red solid line).

3.3 Conclusion

By creating a dynamic chemical network based on imine condensation, our laboratory successfully overcame the low fidelity problem of non-enzymatic DNA templated replication and achieved template-directed deoxyribose amine nucleoside polymer (dANP) synthesis. However, the use of reductive amination results in loss of dynamic property. Efforts of creating fully dynamic networks to achieve DNA templated replication have been reported, but replication of monomers have not been successful. In this Chapter, I sought to use rANP building blocks to construct a truly dynamic network that selects for DNA-templated replication behavior.

As shown in **Scheme 3.3**, with the 2'-hydroxyl group, rANP can form reversible N,O-

acetal linkage, a reversible linkage with higher stability compared to imine. We synthesized two uridine derived rANP monomer, rU1 and rU2, and introduced DNA template into rANP networks for template-directed information copying.



Scheme 3.3 poly(dA) template directed rU1 and rU2 oligomerization.

DNA templated reactions were investigated in poly(dA) templated rU networks. Circular dichroism is consistent with DNA/ANP duplex structure upon incubation of rANP with DNA template, and the formation of ANP oligomers is confirmed by MALDI-TOF. The duplexes are found to have very high melting transition temperatures. Stabilization likely comes from charge attraction and restricted backbone conformation. DNA/ANP hybridized B-from duplex structure formation is driven by Watson-Crick base-pairing, with poly(dT) template unable to drive formation of duplexes.

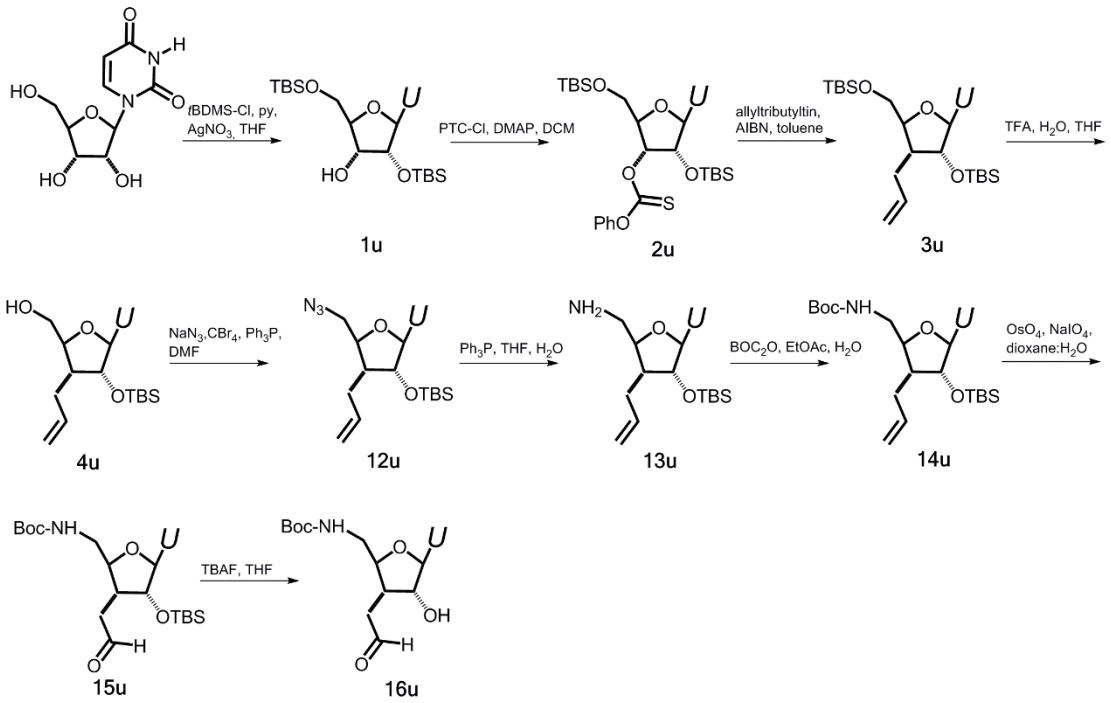
We studied how environmental factors including pH and ionic strength impact DNA

template reaction. The duplexes formation were shown to be highly pH dependent, with slightly acidic pHs being optimal. This result is consistent with the pH profile of N,O-acetal condensation. Neutral or slightly basic pH not favored likely due to slow reaction rate, or unfavorable configuration from deprotonation of N,O-acetal. Ionic strength also impacts assembly. High ionic strength is required for rU2 network to avoid precipitation upon mixing with DNA template, and is also favored (though not required) for templated rU1 network. It is likely because higher ionic strength stabilizes hydrophobic stacking of duplexes and screens charges, herein facilitates the reaction.

3.4 Experimental

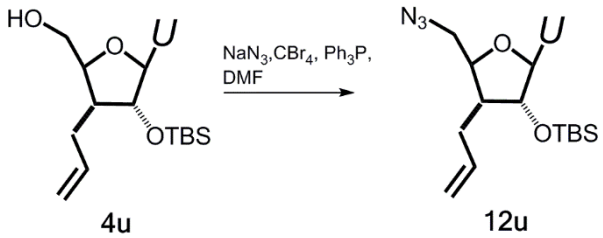
3.4.1 Synthesis of rANP Monomer rU1

Synthetic route of rU1 is shown in **Scheme 3.4**. The 2'- and 5'- hydroxyl groups are protected with TBS using $AgNO_3$ as catalyst. The free 3'-hydroxyl group is activated by a phenylthiolcarbonyl group. Then the carbon skeleton of 3'-aldehyde moiety is introduced by a stereoselective radical allylation reaction initiated by AIBN, and masked as an olefin. The 5'-TBDMS group can be selectively removed using TFA: H_2O : THF = 1:1:4 (v/v/v) mixed solvent, and 5'-hydroxyl is converted to azido group and reduced to primary amine, followed by Boc protection. The aldehyde is unmasked by OsO_4 oxidation of olefin, and deprotection of 2'-TBS is achieved by TBAF protocol.



Scheme 3.4 Synthetic scheme of rU1

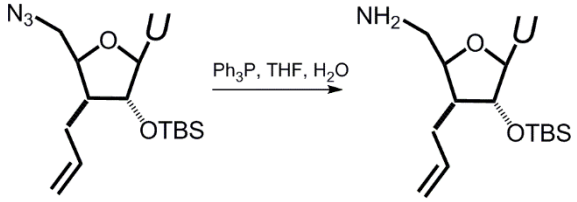
1. 3'-C-Allyl-5'-deoxy-5-azido-2'-O-tert-butyldimethylsilyluridine ***12u***[13a, 14]*:*



Synthesis of **4u** is described in Chapter 2. To 20 mL DMF solution containing 270 mg of **4u** (0.71 mmol) was added 231 mg NaN_3 (5 eq) and 196 mg PPh_3 (1.05 eq). CBr_4 247 mg (1.05 eq) was added in portions. The reaction mixture was stirred overnight under N_2 protection at room temperature. After 24 h, 1.5 mL MeOH was added to quench the reaction. After filtering off precipitants, the filtrate was diluted with EtOAc, washed with H_2O once then brine once, dried by $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash column chromatography with

Hexane : EtOAc = 2:1 to give product **12u**. Yield: 87%. ^{1}H NMR (400 MHz, cdcl$_3$) δ 10.06 – 9.90 (s, 1H), 7.83 (d, *J* = 7.4 Hz, 1H), 5.82 – 5.67 (m, 2H), 5.62 (s, 1H), 5.16 – 5.04 (m, *2*H), 4.29 (d, *J* = 4.4 Hz, 1H), 4.14 (m, 1H), 3.88 (dd, *J* = 12.4, 2.5 Hz, 1H), 3.59 (dd, *J* = 13.8, 4.0 Hz, 1H), 2.34 (m, 1H), 2.10 – 1.93 (m, 2H), 0.92 (m, 9H), 0.29 – 0.03 (m, 6H). HRMS: M-N_2+H^+ ($C_{18}H_{30}N_3O_4Si$): calculated: 380.20056, found: 380.20013. FT-IR of v_{as}(N=N=N): 2105cm^{-1}.

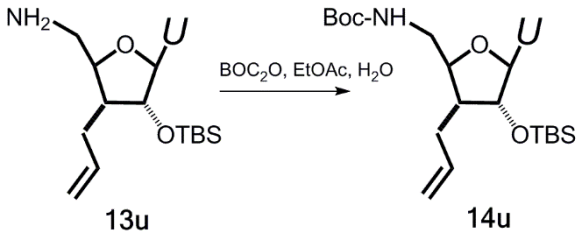
2. 3'-C-Allyl-5'-deoxy-5-amino-2'-O-tert-butyldimethylsilyluridine ***13u***[14]*:*



12u **13u**

1.205g **12u** (0.515 mmol) was dissolved in 30 mL THF + 3 mL H_2O. 1.55 g (2 eq) PPh_3 was added. Reflux overnight at 80^{o}C under protection of N_2. Removed solvent *in vacuuo*, and the residue was purified by silica gel column chromatography with DCM : MeOH = 10:1 to give the product **13u**. Yield: 84%. ^{1}H NMR (400 MHz, cdcl$_3$) δ 8.13 – 8.02 (m, 1H), 5.84 – 5.67 (m, 2H), 5.64 (s, 1H), 5.08 (dd, *J* = 18.6, 13.6 Hz, 2H), 4.31 (d, *J* = 4.3 Hz, 1H), 4.11 – 4.02 (m, 1H), 3.18 (dd, *J* = 14.0, 2.5 Hz, 1H), 2.94 (dd, *J* = 14.0, 5.9 Hz, 1H), 2.43 – 2.30 (m, 1H), 2.11 – 1.99 (m, 1H), 1.91 (m, 1H), 0.91 (m, 9H), 0.28 – 0.06 (m, 6H). HRMS: M+H^+ ($C_{18}H_{32}N_3O_4Si$): calculated: 382.21621, found: 382.21546

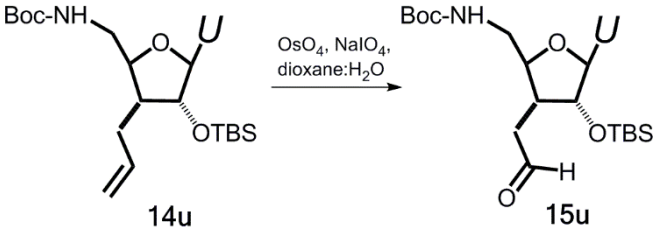
3. *3'-C-Allyl-2'-O-tert-butyldimethylsilyl-5'-deoxy-5-(tert-butyloxycarbonylamino)-uridine **14u**:*



13u 14u

After 120 mg **13u** (0.315 mmol) was dissolved in 7 mL EtOAc and 7 mL H_2O, 79 mg $(Boc)_2O$ (1.15 eq) was added. The reaction mixture was stirred overnight at 25°C. The organic layer was collected, washed with saturated $NaHCO_3$ solution then H_2O, dried through anhydrous $MgSO_4$ and concentrated. Silica gel column purification with Hexane : EtOAc = 1:1 gives product **14u.** Yield: 76%. ^1H NMR (600 MHz, cdcl$_3$) δ 10.05 (s, 1H), 7.58 (t, *J* = 25.4 Hz, 1H), 5.77 – 5.62 (m, 2H), 5.57 (s, 1H), 5.12 – 4.95 (m, 3H), 4.33 – 4.28 (m, 1H), 4.08 – 4.00 (m, 1H), 3.59 (dd, *J* = 13.2, 5.5 Hz, 1H), 3.31 – 3.18 (m, 1H), 2.37 – 2.27 (m, 1H), 2.08 (ddd, *J* = 25.2, 15.4, 7.9 Hz, 1H), 1.69 (d, *J* = 4.0 Hz, 1H), 1.49 – 1.31 (m, 9H), 0.91 – 0.81 (m, 9H), 0.24 – 0.04 (m, 6H). HRMS: $M+H^+$ ($C_{23}H_{40}N_3O_6Si$): calculated: 482.26864, found: 482.26819.

4. *3'-C-acetaldehyde-2'-O-tert-butyldimethylsilyl-5'-deoxy-5-(tert-butyloxycarbonylamino)-uridine **15u**[11a]:*

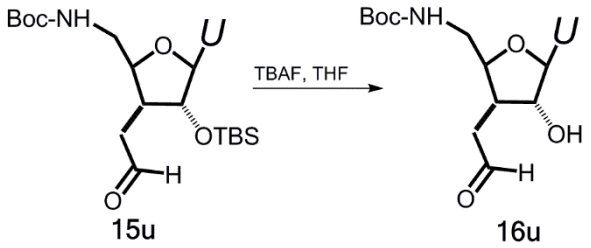


14u 15u

0.967 g **14u** (2 mmol) was dissolved into 40 mL dioxane and 20 mL H_2O. 1 mL OsO_4

(4wt% in H_2O, 0.01e q) was added dropwise via syringe. Reaction mixture turned into dark brown color in 5min. Stirred for another 15 min, then $NaIO_4$ was dissolved into 30 mL dioxane and 20 mL H_2O and added dropwise over 1h by titration funnel. Stirred for another 1h. Dilute the reaction mixture with 125 mL H_2O + 25 mL brine, extract with 150 mL DCM for three times, combined the organic layer, dried with anhydrous $MgSO_4$, filtered and concentrated. Residue was purified by silica gel column chromatography with Hexane : EtOAc = 1:1 through 1:2 to give product **15u**. Yield: 75%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 9.80 (d, *J* = 5.6 Hz, 1H), 8.23 (s, 1H), 7.57 (d, *J* = 8.1 Hz, 1H), 5.72 (ddd, *J* = 15.2, 8.5, 4.2 Hz, 1H), 5.67 (s, 1H), 5.01 (s, 1H), 4.44 (d, *J* = 4.8 Hz, 1H), 4.00 (dd, *J* = 9.4, 4.1 Hz, 1H), 3.59 (ddd, *J* = 15.0, 6.9, 2.8 Hz, 1H), 3.34 – 3.25 (m, 1H), 2.83 (dd, *J* = 19.2, 8.6 Hz, 1H), 2.68 (d, *J* = 15.6 Hz, 1H), 2.27 (m, 1H), 1.44 (s, 9H), 0.93 – 0.87 (m, 9H), 0.22 – -0.04 (m, 6H). HRMS: M+H$^+$ ($C_{22}H_{38}N_3O_7Si$): calculated: 484.24790, found: 484.24777

5. 3'-C-acetaldehyde-5'deoxy-5-tert-butyloxycarbonylamino-uridine ***16u***

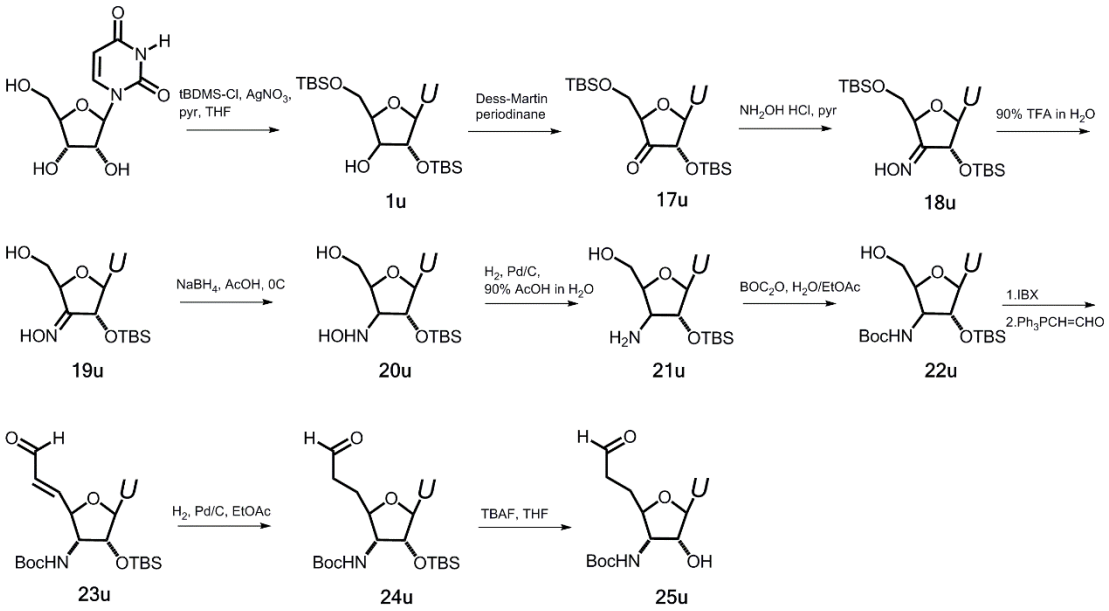


0.72 g **15u** (1.49 mmol) was dissolved in 10 mL THF at 0^{o}C. 1.56 mL TBAF (1.0 M in THF, 1.05 eq) was added dropwise at 0^{o}C. Warmed up the solution to 25^{o}C and stirred for 2 hrs under the protection of N_2. Solvent was removed under reduced pressure, and residue was purified by silica gel flash column using DCM : MeOH = 15:1 through

10:1 to give product **16u**. Yield: 62%. NMR shows product is a mixture of two diastereomers with 65:35 ratio. ^{1}H NMR (600 MHz, DMSO-d6) δ 11.35 (s, 1H), 7.63 (t, *J* = 7.5 Hz, 1H), 7.05 (s, 1H), 6.54 (d, *J* = 2.6 Hz, 0.65H) & 6.37 (d, *J* = 5.2 Hz, 0.35H), 5.91 (d, *J* = 2.8 Hz, 0.65H) & 5.72 (d, *J* = 1.7 Hz, 0.35H), 5.62 (dd, *J* = 8.0, 2.2 Hz, 0.65H) & 5.60 (dd, *J* = 8.0, 2.2 Hz, 0.35H), 5.54 (s, 0.35H) & 5.53 – 5.47 (m, 0.65H), 4.68 – 4.57 (m, 1H), 4.20 (m, *J* = 7.5 Hz, 0.65H) & 3.68 (m, 0.35H), 3.22 (dd, *J* = 12.5, 6.3 Hz, 1H), 3.15 – 3.05 (m, 1H), 2.84 (d, *J* = 7.7 Hz, 0.35H) & 2.79 (dd, *J* = 16.4, 8.0 Hz, 0.65H), 1.94 (m, 1H), 1.72 (m, 1H), 1.37 (s, 9H). HRMS: $M+H^+$ ($C_{16}H_{24}N_3O_7$): calculated: 370.16143, found: 370.16120.

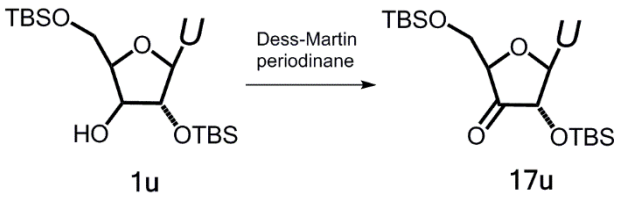
3.4.2 Synthesis of rANP Monomer rU2

Synthetic scheme of rU2 is summarized in **Scheme 3.5**. After protection of the 2'- and 5'- hydroxyl groups with TBDMS, the 3'- hydroxyl group is oxidized into ketone with Dess-Martin oxidation, then converted to oxime. After deprotection of 5'-TBS, the oxime is stereoselectively reduced into hydroxylamine using $NaBH_4$ in AcOH. The hydroxylamine is further reduced to amine using Pd/C catalyzed hydrogenation, and protected with Boc. The 5'-hydroxyl group is oxidized into aldehyde using IBX oxidation, followed by the Wittig ylide reaction to generate α,β-unsaturated aldehyde. The C=C double bond is selectively reduced by Pd/C catalyzed hydrogenation to yield the final product.



Scheme 3.5 Synthetic scheme of rU2.

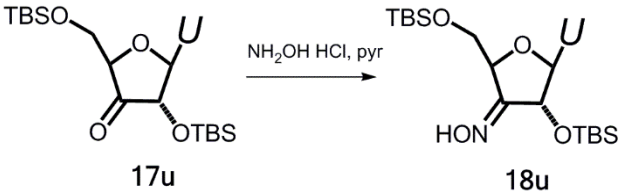
1. *2',5'-O-Bis-tert-butyldimethylsilyl-β-D-erythro-pentofuranos-3'-ulosyluracil* ***17u***[15]



1u was synthesized based on protocol described in Chapter 2. 5.8 g (12.29 mmol) **1u** was dissolved in 100 mL DCM, and the solution was cooled down to 0°C under N_2 protection. 7.82 g Dess-Martin periodinone (1.5 eq) was dissolved in 50 mL DCM and added into reaction mixture via syringe dropwise. The reaction mixture was warmed up to 25°C and stirred overnight. Diluted the reaction mixture with EtOAc, washed with saturated $NaHCO_3$ solution three times to remove acids, followed by saturated $Na_2S_2O_3$ solution three times to remove remaining oxidants, and water once. Combined organic layer and dried through anhydrous $MgSO_4$. Removed solvent under reduced pressure,

and residue was subject to silica gel column purification with Hexane : EtOAc = 3:1 to give product **17u**. Yield: 96%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8.48 (s, 1H), 7.83 (d, *J* = 8.2 Hz, 1H), 6.24 (d, *J* = 8.0 Hz, 1H), 5.82 – 5.77 (m, 1H), 4.22 (s, 1H), 4.15 (d, *J* = 8.0 Hz, 1H), 3.90 (qd, *J* = 11.2, 1.9 Hz, 2H), 0.94 – 0.81 (m, 18H), 0.12 – -0.04 (m, 12H). HRMS: M+H$^+$ ($C_{21}H_{39}N_2O_6Si_2$): calculated: 471.23467, found: 471.23455.

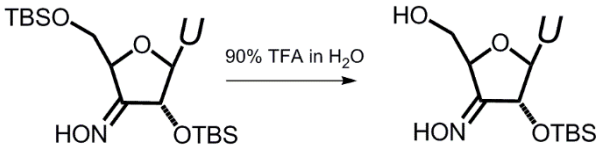
2. *1-[2,5-O-Bis(tert-butyldimethylsilyl)-3-deoxy-3-(hydroxy-imino-)-β-D-erythro-pentofuranossyl]uracil* ***18u***[16]



5.55 g (11.80 mmol) **17u** and 4.10 g hydroxylamine hydrochloride (5 eq) was dissolved into 150 mL pyridine. The reaction mixture was stirred overnight at room temperature under protection of N_2. The solvent was removed *in vacuuo*, and residue diluted with EtOAc. Washed with water twice followed by brine once, dried with anhydrous $MgSO_4$, filtered and concentrated. Residue was purified by silica gel flash column with Hexane : EtOAc = 3:1 to give product **18u**. Yield: 95%. NMR shows the product is a mixture of two diastereomers with 1:1 ratio. ^{1}H NMR (400 MHz, cdcl$_3$) δ 8.35 (s, 0.5H) & 8.30 (s, 0.5H), 8.08 (s, 0.5H) & 7.68 (s, 0.5H),7.95 (d, *J* = 8.2 Hz, 0.5H) & 7.80 (d, *J* = 8.1 Hz, 0.5H), 6.12 (d, *J* = 7.2 Hz, 0.5H) & 5.97 (d, *J* = 2.8 Hz, 0.5H), 5.77 (d, *J* = 6.6 Hz, 0.5H) & 5.70 (d, *J* = 7.8 Hz, 0.5H), 5.04 (s, 0.5H) & 4.90 (s, 0.5H), 4.81 (s, 0.5H) & 4.62 (d, *J* = 7.0 Hz, 0.5H), 4.19 – 4.06 (m, 1H), 3.96 (dd, *J* = 11.4, 3.1 Hz, 0.5H) & 3.91 – 3.83

(m, 0.5H), 0.98 – 0.83 (m, 18H), 0.16 – -0.02 (m, 12H). HRMS: M+H$^+$ ($C_{21}H_{40}N_3O_6Si_2$): calculated: 486.24557, found: 486.24508.

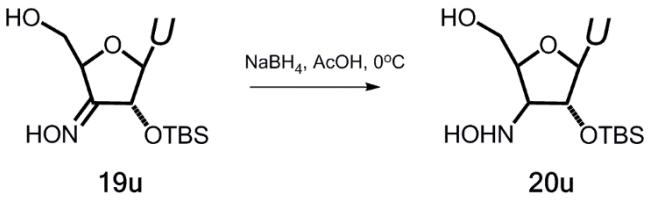
3. *1-[2-O-tert-butyldimethylsilyl-3-deoxy-3-(hydroxy-imino-)-β-D-erythro-pentofuranossyl]uracil* ***19u***[17]



18u **19u**

4.61 g **18u** was added into 50 mL 90% TFA in H_2O at 0^{o}C with vigorous stirring. Continued stirring for 20 min at 0^{o}C under N_2 protection. Removed TFA and H_2O as soon as possible under reduced pressure at room temperature. The residue was co-evaporated with toluene three times to further remove remaining solvent. Residue was purified by silica gel column chromatography using Hexane : EtOAc = 3:7 to give product **19u**. Yield: 73%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8,91 (s, 1H), 8.41 (s, 1H), 7.55 (d, *J* = 8.1 Hz, 1H), 5.83 (d, *J* = 8.1 Hz, 1H), 5.62 (d, *J* = 6.6 Hz, 1H), 5.06 (d, *J* = 6.7 Hz, 1H), 5.04 (m, 1H), 4.18 (dd, *J* = 12.0, 1.5 Hz, 1H), 3.99 (dd, *J* = 12.0, 2.0 Hz, 1H), 0.86 (s, 9H), 0.35 - 0.13 (m, 6H) HRMS: M+H$^+$ ($C_{15}H_{26}N_3O_6Si$): calculated: 372.15909, found: 372.15947.

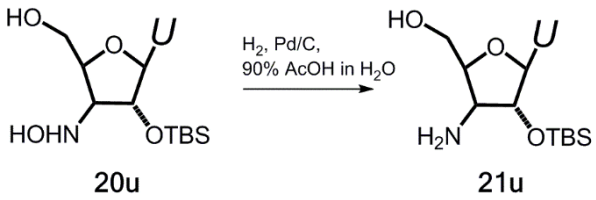
*4. 2'-O-tert-butyldimethylsilyl-3'-deoxy-3-(hydroxyamino)-uridine **20u***



19u **20u**

1.026 g (2.77 mmol) **19u** was dissolved into 10 mL AcOH. 524 mg $NaBH_4$ (5 eq) was dissolved in 25 mL AcOH and cooled down to 5°C to generate $NaAcBH_3$ *in situ*. To the AcOH solution of $NaBH_4$ was added AcOH solution of **19u** dropwise via syringe. Stirred the reaction mixture at 5°C for 2 hrs, then removed the solvent *in vacuuo*. The residue was diluted with EtOAc, washed with H_2O for three times then brine once, dried with anhydrous $MgSO_4$, filtered and removed the solvent under reduced pressure. The residue was purified by silica gel flash column with DCM : MeOH = 10:1 to give the product **20u**. Yield: 83%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 9.31 (s, 1H), 7.86 (d, *J* = 8.1 Hz, 1H), 5.72 (d, *J* = 8.1 Hz, 1H), 5.65 (d, *J* = 3.2 Hz, 1H), 4.64 (dd, *J* = 4.9, 3.5 Hz, 1H), 4.11 (d, *J* = 6.7 Hz, 1H), 4.01 (dd, *J* = 12.4, 1.9 Hz, 1H), 3.81 (dd, *J* = 12.4, 1.9 Hz, 1H), 3.66 – 3.61 (m, 1H), 0.91 (s, 9H), 0.16 – -0.05 (m, 6H). HRMS: $M+H^+$ ($C_{15}H_{28}N_3O_6Si$): calculated: 374.17474, found: 374.17406.

5. *2'-O-tert-butyldimethylsilyl-3'-deoxy-3-(amino)uridine* **21u**

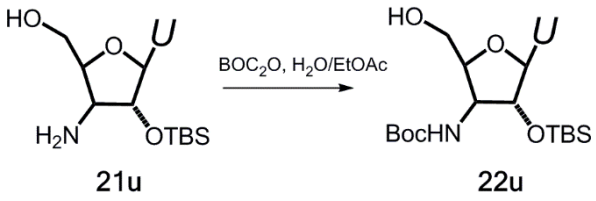


20u **21u**

132 mg **20u** (0.35 mmol) was dissolved into 15 mL of 90% AcOH in H_2O. 13 mg Pd/C

(10 wt%) was added and the reaction mixture was stirred under H_2 atmosphere overnight using a balloon. Monitor the reaction progress by TLC. After the reaction reached completion, Pd/C was removed by filtration and solvent was evaporated *in vacuuo*. The residue was co-evaporated with ethanol three times, and purified by column chromatography with DCM : MeOH = 10:1 to give product **21u**. Yield: 87%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 9.65 (s, 1H), 7.56 (d, *J* = 8.2 Hz, 1H), 5.83 (d, *J* = 8.1 Hz, 1H), 5.63 (d, *J* = 7.1 Hz, 1H), 4.17 (d, *J* = 13.1 Hz, 1H), 3.88 (d, *J* = 10.9 Hz, 1H), 3.73 (q, 1H), 0.86 (m, 9H), 0.10-0.02 (m, *J* = 37.4 Hz, 6H). HRMS: M+H$^+$ ($C_{15}H_{28}N_3O_5Si$): calculated: 358.17983, found: 358.17939.

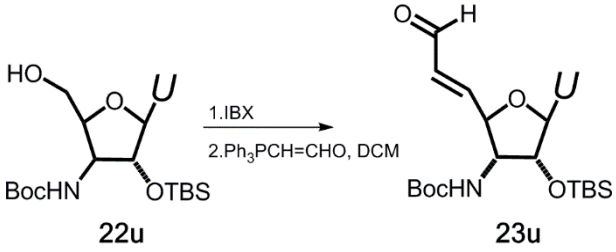
6. 2'-O-tert-butyldimethylsilyl-3'-deoxy-3-(tert-butyloxycarbonylamino)uridine ***22u***



21u 22u

2.815 g **21u** (7.88 mmol) was dissolved into 80 mL EtOAc + 80 mL H_2O. 1.730 g $(Boc)_2O$ was added and stirred overnight. The organic layer was separated, washed with saturated $NaHCO_3$ solution then water, dried by anhydrous $MgSO_4$ and concentrated to give product **22u**. Yield: 93%. ^{1}H NMR (400 MHz, cdcl$_3$) δ 8.35 (s, 1H), 8.12 (d, *J* = 8.2 Hz, 1H), 5.73 (d, *J* = 8.1 Hz, 1H), 5.69 (s, 1H), 4.91 (d, *J* = 8.3 Hz, 1H), 4.26 (d, *J* = 4.4 Hz, 1H), 4.06 (m, 2H), 3.79 (d, *J* = 13.7 Hz, 1H), 3.35 (m, 1H), 1.44 (s, 9H), 0.94 (m, *J* = 6.3 Hz, 9H), 0.23 (d, *J* = 37.4 Hz, 6H).HRMS: M+H$^+$ ($C_{20}H_{36}N_3O_7Si$): calculated: 458.23226, found: 458.23081.

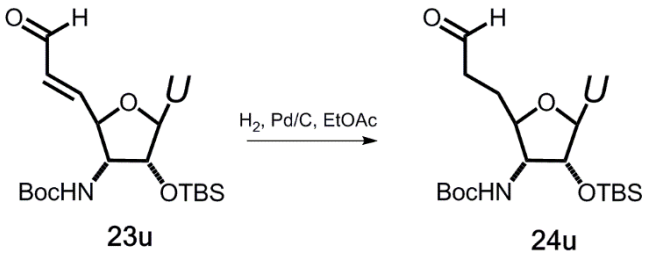
7. *5'(S)-[1'-(3'-deoxy-3-(tert-butyloxycarbonylamino)-2'-O-tert-butyldimethylsilyl-β-D-ribo-5'-penta-1',4'-furanosyl)-uracil] α,β-unsaturated δ-lactone* ***23u***[18-19]



22u **23u**

100 mg **22u** (0.28 mmol) was dissolved in 5mL acetonitrile. 340 mg SIBX (45 wt% IBX with stabilizer, 2.5 eq) was added into the solution. The suspension was refluxed at 80°C for 45 min under N_2 protection. After cooling down the suspension to 0°C, the insoluble substance was filtered off and washed with EtOAc. The filtrate was washed with saturated $NaHCO_3$ solution once, followed by water once. The organic layer was combined, dried with $MgSO_4$, and evaporated to dryness. Due to the instability of 5'-aldehyde uridine, the resulting product was directly subject to next step of reaction. After dissolving the aldehyde product into DCM, 97 mg of Wittig ylide reagent $Ph_3PCH{=}CHO$ (1.5 eq) was added into the solution, and the reaction mixture was stirred overnight at room temperature under N_2 protection. After reaction completed, the solvent was removed under reduced pressure, and the residue was purified by silica gel column chromatography with Hexane : EtOAc = 4:5 to give product **23u** containing triphenylphosphine oxide impurities. HRMS: $M+H^+$ ($C_{22}H_{36}N_3O_7Si$): calculated: 482.23226, found: 482.23236.

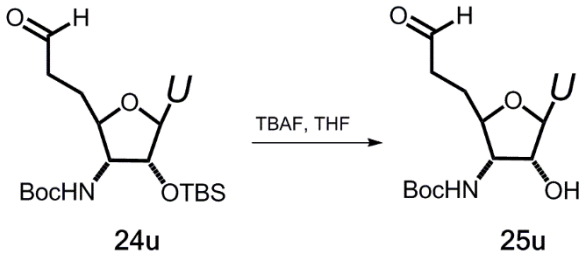
8. *5'(S)-[1'-(3'-deoxy-3-(tert-butyloxycarbonylamino)-2'-O-tert-butyldimethylsilyl-β-D-ribo-5'-penta-1',4'-furanosyl)-uracil]lactone* **24u**[20]



23u 24u

584 mg **23u** with triphenylphosphine oxide impurities was dissolved into 20 mL EtOAc. 58 mg Pd/C was added, and the reaction mixture was stirred under H_2 for 2 hours. After filtered off Pd/C, the filtrate was concentrated under reduced pressure and purified by silica gel column chromatography with Hexane : EtOAc = 3:5 to give **24u** containing triphenylphosphine oxide impurities. HRMS: $M+H^+$: ($C_{22}H_{38}N_3O_7Si$) calculated: 484.24790, found: 484.24744.

9. *5'(S)-[1'-(3'-deoxy-3-(tert-butyloxycarbonylamino)-β-D-ribo-5'-penta-1',4'-furanosyl)-uracil]lactone* **25u**



24u 25u

336 mg **24u** was dissolved in 10 mL THF. Cooled down the solution to 0°C, then added 0.73 mL TBAF (1.0 M in THF, 1.05eq) dropwise via a syringe. Warmed up the reaction mixture to room temperature and stirred for 2h. Removed the solvent *in vacuuo* and purified the residue by silica gel flash column with DCM : MeOH = 10:1. The product

was further purified by HPLC if necessary. Combined yield over 3 steps: 35%. ^{1}H NMR (400 MHz, cdcl$_3$) δ 10.16 (s, 1H), 9.83 (s, 1H), 7.56 (d, *J* = 8.1 Hz, 1H), 5.81 (d, *J* = 8.1 Hz, 1H), 5.69 (s, 1H), 5.41 (d, *J* = 8.7 Hz, 1H), 5.31 (s, 1H), 4.25 (d, *J* = 5.3 Hz, 1H), 4.01 (m, 1H), 3.82 (m, 1H), 2.70 (qd, *J* = 18.3, 9.2 Hz, 2H), 2.24 – 2.11 (m, 1H), 2.11 – 1.97 (m, 1H), 1.42 (d, *J* = 13.9 Hz, 11H). HRMS: M+H$^+$ ($C_{16}H_{24}N_3O_7$): calculated: 370.16143, found: 370.15103.

3.4.3 DNA Templated ANP Networks

Stock solution of rU1 is prepared by suspending **16u** in 1 M HCl aqueous solution overnight under N_2 at room temperature to deprotect Boc group. After deprotection, the solution was dried down under reduced pressure, and the residue was re-dissolved in DNAase-free water. The purity of product is confirmed by HPLC and mass spectrometry. Mass spectrometry gives a single peak of m/z = 270.10822, corresponds to rU1 (M+H$^+$: $C_{11}H_{16}N_3O_5$, calculated: 270.10899). Stock solution of rU2 is prepared the same way by deprotecting **25u**. ESI-MS (M+H$^+$: $C_{11}H_{16}N_3O_5$) found: 270.10868, calculated: 270.10899

DNA templates are purchased from Integrated DNA Technologies at standard desalting or HPLC purification grade. Standard desalting grade samples are further purified by reverse-phase HPLC purification using a C18 column, then dried by lyophilization. Dissolve DNA with DNAase free water to make stock solution of DNA templates.

Concentration of DNA template stock solution and rU1/rU2 stock solution is determined by a NanoDrop 2000 UV-Vis Spectrophotometer with absorption at 260 nm. Extinction coefficient of DNA template is provided by IDT, and extinction coefficient of rU1/rU2 is assumed to be the same as uridine.

DNA template stock solution and rANP monomer stock solution are diluted to indicated concentration using pH = 4.5 Acetate, pH = 6.5 MES or pH = 8.5 TAPS buffer (c = 50 mM, I = 100 mM or 500 mM). The mixture was slowly heated up to 75^{o}C, then cool down to room temperature and incubate at 4^{o}C.

DNA templated polymerization is monitored by Circular Dichroism coupled with a JASCO PFD-425S temperature control unit. CD spectra were recorded on a Jasco J-810 CD spectropolarimeter in 0.10 mm or 0.01 mm quartz cells. Spectra were recorded from 350 nm to 190 nm at a scanning rate of 100 nm/min and a resolution of 0.5 nm.

References

1. (a) Li, X.; Lynn, D. G., Polymerization on Solid Supports. *Angewandte Chemie* **2002,** *114* (23), 4749-4751; (b) Li, X.; Hernandez, A. F.; Grover, M. A.; Hud, N. V.; Lynn, D. G., Step-growth control in template-directed polymerization. *Heterocycles* **2011,** *82* (2), 1477-1488.

2. Bean, H. D.; Anet, F. A. L.; Gould, I. R.; Hud, N. V., Glyoxylate as a Backbone Linkage for a Prebiotic Ancestor of RNA. *Origins of Life and Evolution of Biospheres* **2006,** *36* (1), 39-63.

3. Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R., Self-Assembling Sequence-Adaptive Peptide Nucleic Acids. *Science* **2009,** *325* (5936), 73-77.

4. Hansen, D. J.; Manuguerra, I.; Kjelstrup, M. B.; Gothelf, K. V., Synthesis, Dynamic Combinatorial Chemistry, and PCR Amplification of 3′–5′ and 3′–6′ Disulfide-linked Oligonucleotides. *Angewandte Chemie International Edition* **2014,** *53* (52), 14415-14418.

5. (a) Barbeyron, R.; Vasseur, J.-J.; Smietana, M., pH-controlled DNA- and RNA-templated assembly of short oligomers. *Chemical Science* **2015,** *6* (1), 542-547; (b) Martin, A. R.; Barvik, I.; Luvino, D.; Smietana, M.; Vasseur, J.-J., Dynamic and Programmable DNA-Templated Boronic Ester Formation. *Angewandte Chemie International Edition* **2011,** *50* (18), 4193-4196.

6. (a) Eschenmoser, A., Etiology of Potentially Primordial Biomolecular Structures: From Vitamin B12 to the Nucleic Acids and an Inquiry into the Chemistry of Life's Origin: A Retrospective. *Angewandte Chemie International Edition* **2011,** *50* (52), 12412-12472; (b) Sreenivasachary, N.; Hickman, D. T.; Sarazin, D.; Lehn, J.-M., DyNAs: Constitutional Dynamic Nucleic Acid Analogues. *Chemistry – A European Journal* **2006,** *12* (33), 8581-8588.

7. Jones, R. J.; Swaminathan, S.; Milligan, J. F.; Wadwani, S.; Froehler, B. C.; Matteucci, M. D., Oligonucleotides containing a covalent conformationally restricted

phosphodiester analog for high-affinity triple helix formation: the riboacetal internucleotide linkage. *Journal of the American Chemical Society* **1993,** *115* (21), 9816-9817.

8. Herdewijn, P., Targeting RNA with Conformationally Restricted Oligonucleotides. *Liebigs Annalen* **1996,** *1996* (9), 1337-1348.

9. Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K., New catalysts and procedures for the dimethoxytritylation and selective silylation of ribonucleosides. *Canadian Journal of Chemistry* **1982,** *60* (9), 1106-1113.

10. Meier, C.; Huynh-Dinh, T., Improved Conversion of Adenosine to 3′-Deoxyadenosine. *Synlett* **1991,** *1991* (04), 227-228.

11. (a) Fiandor, J.; Y. Tam, S., Synthesis of 3′-deoxy-3′-(2-propynyl)thymidine and 3′-cyanomethyl-3′-deoxythymidine, analogs of AZT. *Tetrahedron Letters* **1990,** *31* (5), 597-600; (b) Keck, G. E.; Enholm, E. J.; Yates, J. B.; Wiley, M. R., Selectivity and synthetic applications of radical reactionsOne electron C-C bond forming reactions VIA allylstannanes: Scope and limitations. *Tetrahedron* **1985,** *41* (19), 4079-4094.

12. Zhu, X.-F.; Williams, H. J.; Scott, A. I., Facile and highly selective 5[prime or minute]-desilylation of multisilylated nucleosides. *Journal of the Chemical Society, Perkin Transactions 1* **2000,** (15), 2305-2306.

13. (a) Hata, T.; Yamamoto, I.; Sekine, M., A simple method for the synthesis of 5'-azido-5'-deoxyribonucleosides. *Chemistry Letters* **1975,** *4* (9), 977-980; (b) Yamamoto, I.; Sekine, M.; Hata, T., One-step synthesis of 5'-azido-nucleosides. *Journal of the Chemical Society, Perkin Transactions 1* **1980,** (0), 306-310.

14. Mungall, W. S.; Greene, G. L.; Heavner, G. A.; Letsinger, R. L., Nucleotide chemistry. XX. Use of the azido group in the synthesis of 5'-terminal aminodeoxythymidine oligonucleotides. *The Journal of Organic Chemistry* **1975,** *40* (11), 1659-1662.

15. Robins, M. J.; Samano, V.; Johnson, M. D., Nucleic acid-related compounds. 58. Periodinane oxidation, selective primary deprotection, and remarkably stereoselective reduction of tert-butyldimethylsilyl-protected ribonucleosides. Synthesis of 9-(.beta.-D-xylofuranosyl)adenine or 3'-deuterioadenosine from adenosine. *The Journal of Organic Chemistry* **1990,** *55* (2), 410-412.

16. (a) Tronchet, J. M. J.; Benhamza, R.; Dolatshahi, N.; Geoffroy, M.; Türler, H., 3′-Deoxy-3′-Hydroxyamino-β-D-Xylofuranosyluracil and Derivatives Thereof. *Nucleosides and Nucleotides* **1988,** *7* (2), 249-269; (b) Ogawa, A.; Tanaka, M.; Sasaki, T.; Matsuda, A., Nucleosides and Nucleotides. 180. Synthesis and Antitumor Activity of Nucleosides That Have a Hydroxylamino Group Instead of a Hydroxyl Group at the 2‘- or 3‘-Position of the Sugar Moiety1. *Journal of Medicinal Chemistry* **1998,** *41* (25), 5094-5107.

17. Kojima, N.; Szabo, I. E.; Bruice, T. C., Synthesis of ribonucleic guanidine: replacement of the negative phosphodiester linkages of RNA with positive guanidinium linkages. *Tetrahedron* **2002,** *58* (5), 867-879.

18. Ozanne, A.; Pouységu, L.; Depernet, D.; François, B.; Quideau, S., A Stabilized Formulation of IBX (SIBX) for Safe Oxidation Reactions Including a New Oxidative Demethylation of Phenolic Methyl Aryl Ethers. *Organic Letters* **2003,** *5* (16), 2903-

2906.

19. Wnuk, S. F.; Ro, B.-O.; Valdez, C. A.; Lewandowska, E.; Valdez, N. X.; Sacasa, P. R.; Yin, D.; Zhang, J.; Borchardt, R. T.; De Clercq, E., Sugar-Modified Conjugated Diene Analogues of Adenosine and Uridine: Synthesis, Interaction with S-Adenosyl-l-homocysteine Hydrolase, and Antiviral and Cytostatic Effects. *Journal of Medicinal Chemistry* **2002,** *45* (12), 2651-2658.

20. Hidari, K. I. P. J.; Sanai, Y.; Nagai, Y.; Kanaoka, Y.; Hatanaka, Y.; Hashimoto, M., Synthesis of a carbon-linked CMP-NANA analog and its inhibitory effects on GM3 and GD3 synthases. *Heterocycles* **1996,** *43* (3), 531-534.

21. Johnson, W. C.; Tinoco, I., Circular dichroism of polynucleotides: A simple theory. *Biopolymers* **1969,** *7* (5), 727-749.

22. Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M., Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Research* **2009,** *37* (6), 1713-1725.

23. (a) Janssen, P. G. A.; Jabbari-Farouji, S.; Surin, M.; Vila, X.; Gielen, J. C.; de Greef, T. F. A.; Vos, M. R. J.; Bomans, P. H. H.; Sommerdijk, N. A. J. M.; Christianen, P. C. M.; Lecl ère, P.; Lazzaroni, R.; van der Schoot, P.; Meijer, E. W.; Schenning, A. P. H. J., Insights into Templated Supramolecular Polymerization: Binding of Naphthalene Derivatives to ssDNA Templates of Different Lengths. *Journal of the American Chemical Society* **2009,** *131* (3), 1222-1231; (b) Lin, J.; Surin, M.; Beljonne, D.; Lou, X.; van Dongen, J. L. J.; Schenning, A. P. H. J., On the mechanism of dynamic polymerization via recycled ss-DNA templated assembly of non-natural bases.

Chemical Science **2012,** *3* (9), 2732-2736.

24. (a) Fife, T. H.; Hagopian, L., Oxazolidine hydrolysis. Participation of solvent and buffer in ring opening. *Journal of the American Chemical Society* **1968,** *90* (4), 1007-1014; (b) Fife, T. H.; Hutchins, J. E. C., General-acid-catalyzed ring opening of oxazolidines. Hydrolysis of 2-[4-(dimethylamino)styryl]-N-phenyl-1,3-oxazolidine. *The Journal of Organic Chemistry* **1980,** *45* (11), 2099-2104.

25. Johansen, M.; Bundgaard, H., Prodrugs as Drug Delivery Systems XXV: Hydrolysis of Oxazolidines—A Potential New Prodrug Type. *Journal of Pharmaceutical Sciences* **1983,** *72* (11), 1294-1298.

Chapter Four

TTF-CHO Networks: Environmental Impacts and Self-Assembly

4.1 Introduction

Empowered by template-directed sequence replication of heritable information, nucleic acids have long been regarded as the basic material for reproduction, mutation and selection. However, this view has been challenged by the discovery of prions, an infectious and heritable proteins. Prion infectious propagation is also a template-directed replication process, immobilizing a protein of normal conformation on a misfolded surface and forcing it into the prion conformation[1]. Recent analysis found that prion infectious propagation also undergoes conformational mutation and selection, leading to evolution of different conformations under different conditions[2]. Mammalian prion proteins (PrP) are associated with diseases such as Creutzfeldt-Jacob disease[3], but now it is clear that many other proteins exhibits prion-like conformational infection behavior[4]. All these amyloid proteins self-assemble through a templated mechanism to generate and select for specific conformations in responsive to environmental factors, as shown in **Fig 4.1**[5].

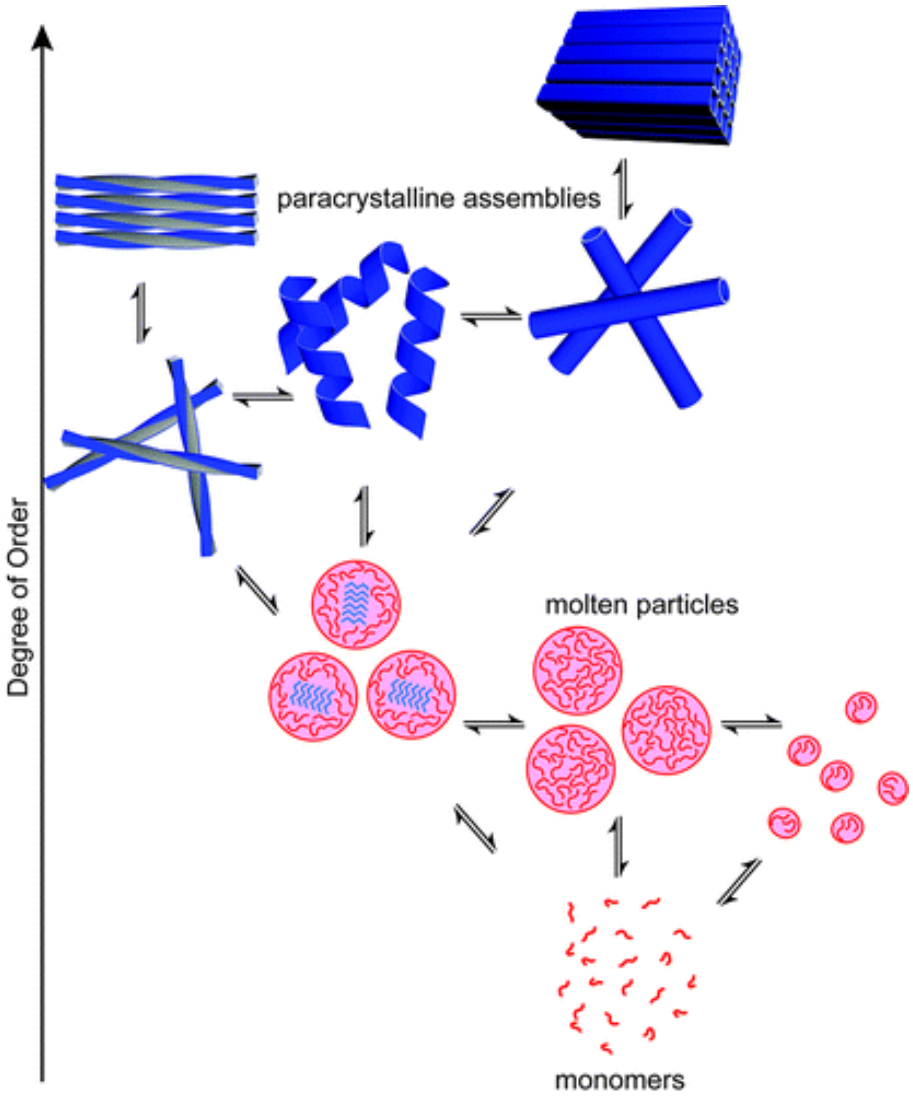


Fig 4.1 Phase network of amyloid peptide assembly[5]. Aβ(16–22) peptide can self-assemble into wide range of morphologies responsive to environment inputs and forms a dynamic phase network.

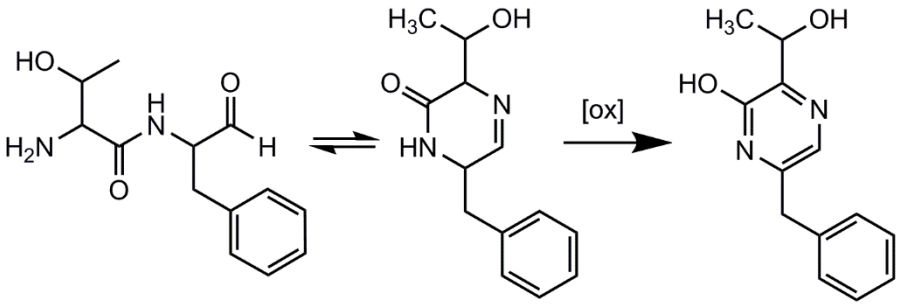
By coupling chemical transformation with peptide self-assembly, self-synthesizing peptides based on cross-β assembly[6] as well as self-assembling peptide DCNs[7] have been reported. However, the self-assembly peptide DCNs reported so far are limited by sequence diversity as well as choice of reversible linkages. With the reversible N,O-acetal linkage on the peptide backbone we established in Chapter 2, it is now possible

to incorporate this linkage into a peptide DCN that can select specifically for cross-β self-assembly propensity. This effort aims to expand the repertoire of dynamic linkages and sequence space of self-assembling peptide DCNs, as well as prove the generality of peptide DCNs based on template-directed conformation replication.

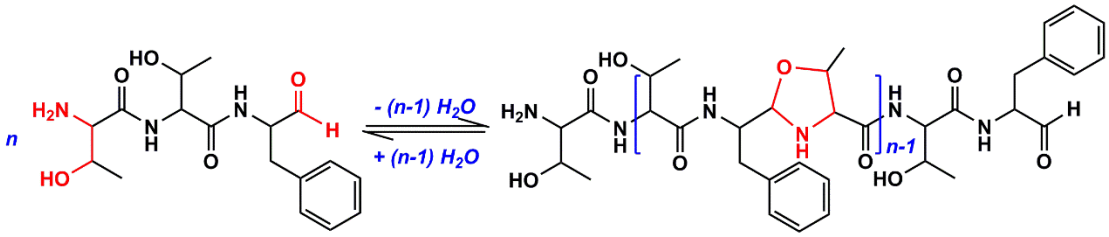
4.2 Results and Discussions

4.2.1 Construction of Dynamic Chemical Networks from TTF-CHO

As discussed in Chapter 2, Threonine and Phenylalanine were chosen as two amino acids for construction of dynamic chemical networks because the hydroxyl group of Threonine side chain is positioned to trap an intermediate imine condensation product as reversible cyclic N,O-acetal linkages, while Phenylalanine provides a high aggregation propensity[8]. In addition, a TF dipeptide has been shown to be selected in an enzymatic peptide DCN to exhibit strong self-assembly propensity[9]. Therefore, TF-CHO was initially chosen as the building block for construction of peptide DCN. Unfortunately, upon deprotection, TF-CHO quickly went through intramolecular cyclization followed by r irreversible oxidation, as shown in **Scheme 4.1**. To overcome this problem, O_2 must be completely excluded, or simply extending the building block one more residue, to which TTF has higher aggregation propensity compared to TFF based on previous simulations[10]. Through reversible N,O-acetal condensation, TTF-CHO should be able to generate the network with cyclic and linear species of different degrees of polymerization, as shown in **Scheme 4.2**.

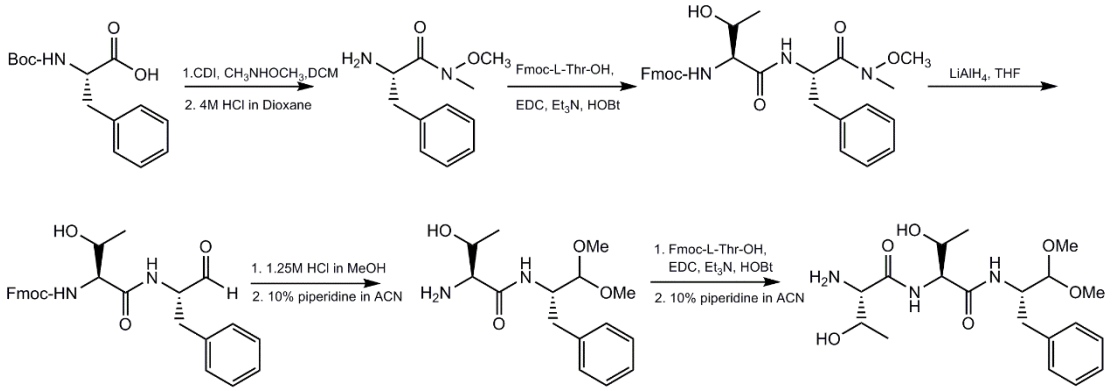


Scheme 4.1 TF-CHO undergoes intramolecular cyclization



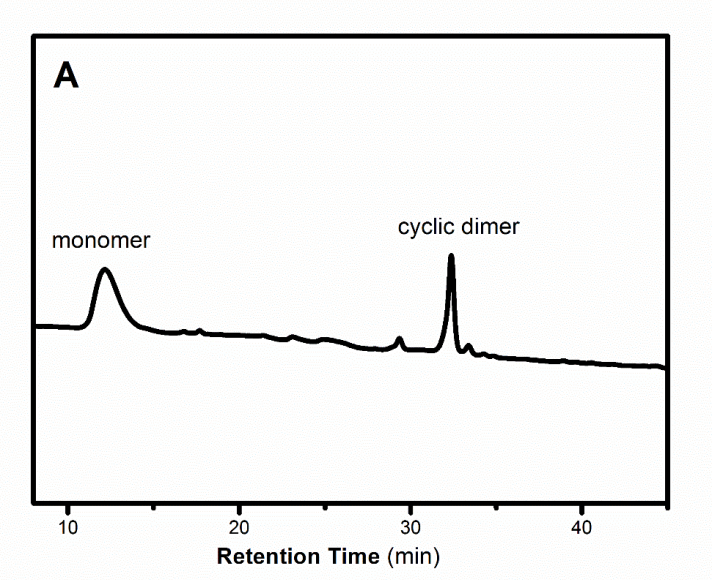
Scheme 4.2 TTF-CHO generates DCNs linked by reversible N,O-acetal bonds

The TTF-CHO synthetic route is shown in **Scheme 4.3**. Directly reducing C-terminus of TTF tripeptide was found unsuccessful, therefore the synthesis of TTF-CHO starts with Phenylalanine. The C-terminus of N-Boc Phenylalanine was converted to the Weinreb amide[11], reduced to aldehyde with LAH[12], and protected as the dimethyl acetal. Coupling of amino acids was achieved by standard EDC coupling protocol[13] using Fmoc-protected amino acids. The final product was purified by HPLC to give the dimethyl acetal protected TTF-CHO in 78% yield.



Scheme 4.3 Synthetic scheme of TTF-CHO

The TTF-CHO was generated *in situ* by deprotecting acetal linkage with 4% HCl in H_2O followed by removing the solvent *in vacuuo*. The residue was then re-dissolved as a 2 mM solution in 40% acetonitrile in water, and adjusted to pH = 7 using Et_3N. This solution was incubated in 4^oC and aliquots are removed for HPLC analysis over the course of 2 weeks. As shown in **Fig 4.2**, the cyclic dimer is the major component of the network. Over 2 weeks, ~ 75% of the monomer converted to cyclic dimer (**Fig 4.2B**). Mass spectrometry analysis of the whole network suggests formation of cyclic trimer and tetramer as well (**Table 4.1**).



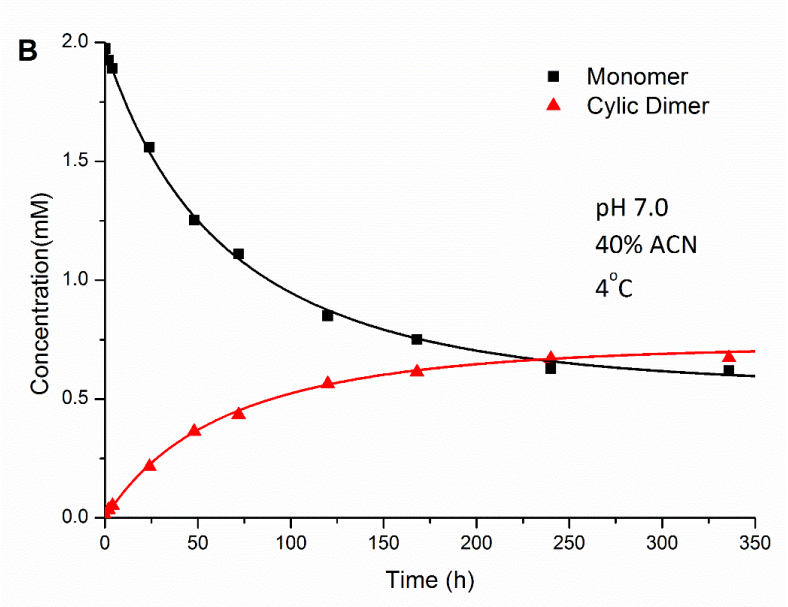


Fig 4.2 HPLC chromatogram and kinetics of TTF DCN in 40% ACN under pH = 7 (**A**) HPLC chromatogram of DCN at day 3 shows emergences of cyclic dimer as major product. (**B**) Kinetics of DCN over 14 days. Kinetics fitting is based on expression in **Scheme 4.4**. Kinetics constants are calculated as: $k_1 = 0.86 \times 10^{-6} (mol^{-1} \bullet sec^{-1})$, $k_{-1} = 0.37 \times 10^{-6}\ sec^{-1}$.

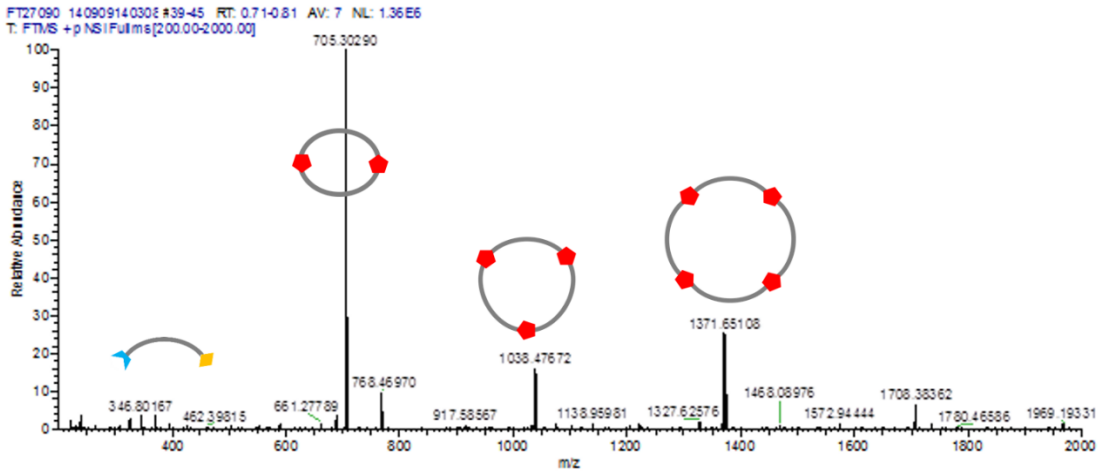
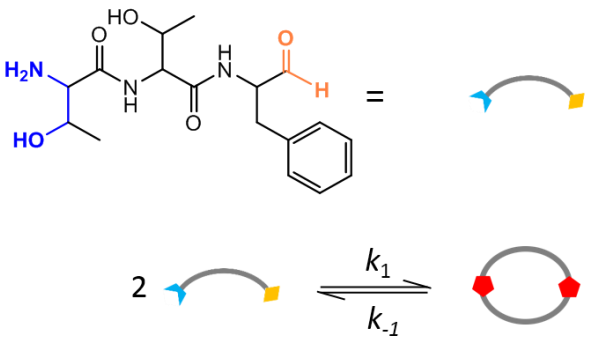


Fig 4.3 Mass spectrometry analysis of whole TTF-CHO DCN. m/z corresponding to monomer, cyclic dimer, cyclic trimer and cyclic tetramers can be detected.

Table 4.1 Mass Spectrometry identification of species generated in TTF-CHO DCN

species	**m/z found**	**m/z calculated**
TTF-CHO	352.1866	352.1872 ($M+H^+$: $C_{17}H_{26}N_3O_5$)
Cyclic dimer	667.3446	667.3455 ($M+H^+$: $C_{34}H_{47}N_6O_8$)
Cyclic trimer	1000.5136	1000.5144 ($M+H^+$: $C_{51}H_{70}N_9O_{12}$)
Cyclic tetramer	1371.6511	1371.6391($M+K^+$: $C_{68}H_{92}N_{12}O_{16}K$)

Under this condition, no obvious assembly structures were observed by TEM, suggesting the cyclic dimer, as well as other species, remain in the solution. The conversion of TTF-CHO monomer to cyclic dimer can be regarded as a simple second-order reaction, and the rate constants ware calculated based on the expression in **Scheme 4.4**. Fitting the experimental data to a second order reaction using minimum SSE, rate constants were calculated as $k_1 = 1.02\times10^{-6}$ and $k_{-1} = 0.130\times10^{-6}$. Corresponding kinetics fitting using minimum SSE is also shown in Fig **4.2B**.



Scheme 4.4 kinetics of TTF-CHO DCN with cyclic dimer as primary product

4.2.2 Environmental Conditions Control Behavior of TTF-CHO DCNs

Built on the models analyzed in Chapter 2 and literature reports[14], we learned that N,O-acetal condensation and hydrolysis are pH-dependent, thus pH is predicted to affect the behavior of DCNs. Here, 2 mM TTF-CHO was dissolved into pH = 6, 7, 8, while maintaining temperature and solvent condition (40% acetonitrile in water, 4°C). As shown in **Fig 4.4**, pH = 8 gave highest concentration of cyclic dimer, while at pH = 6, only very limited cyclic dimer was observed. Based on a second-order reaction model shown in **Scheme 4.4**, kinetic constants were calculated using minimum SSE approach and summarized in **Table 4.2**. The calculated rate constants suggest k_1 remains at same magnitude, and the change in the backward reaction rate k_{-1} dictates the pH-dependence of N,O-acetal concentration.

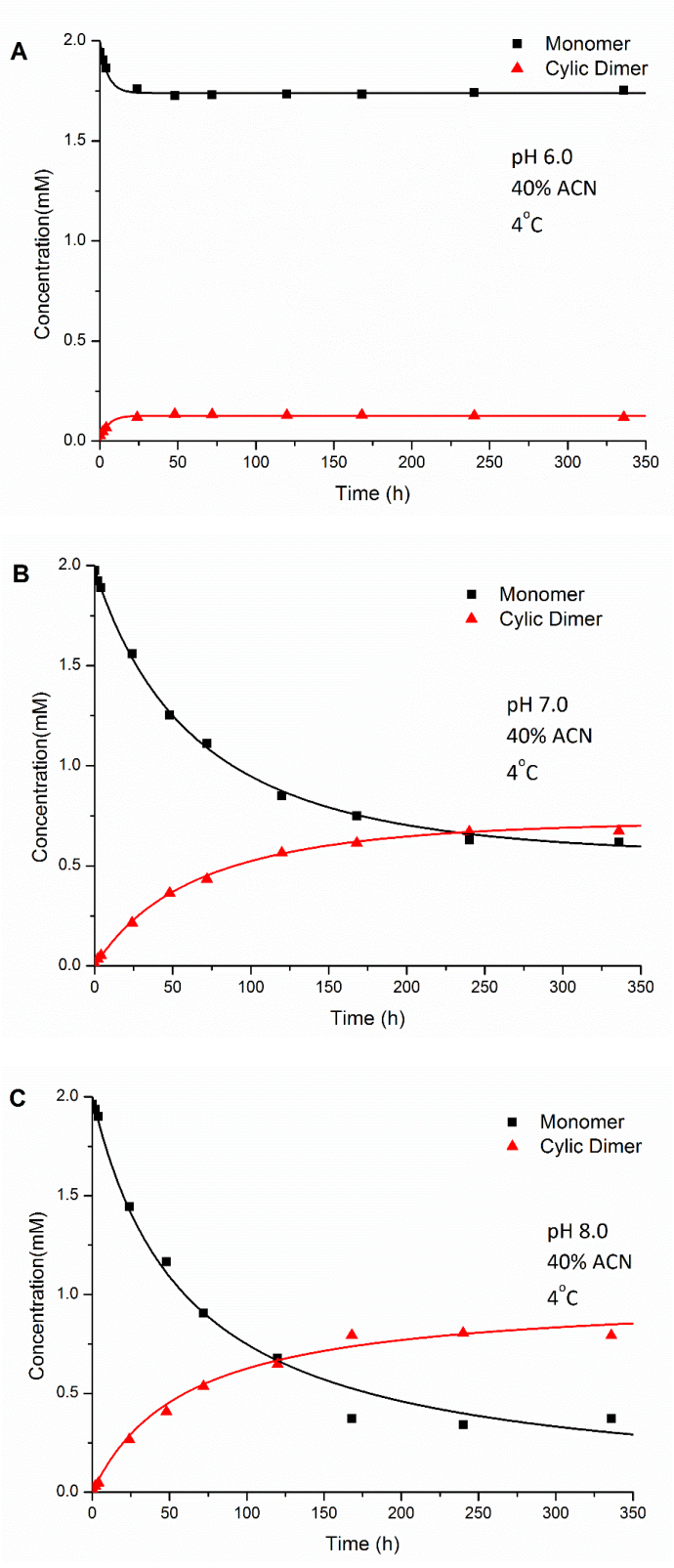


Fig 4.4 Kinetics of TTF-CHO DCNs under different pH. Condition: TTF-CHO = 2 mM, 40% acetonitrile in water, 4°C, (**A**) pH = 6, (**B**) pH = 7, (**C**) pH = 8. pH is adjusted by Et_3N. Kinetics fitting is based on expression in **Scheme 4.4** and calculated with experimental data using minimum SSE. Calculated rate constants are summarized in **Table 4.2**

Table 4.2 Rate and equilibrium constants of TTF-CHO DCNs under different pH

parameter	**pH = 6**	**pH = 7**	**pH = 8**
k_1/(mol^{-1}•sec^{-1})	1.72×10^{-6}	0.86×10^{-6}	1.02×10^{-6}
k_{-1}(sec^{-1})	41.04×10^{-6}	0.37×10^{-6}	0.130×10^{-6}
K/(mol^{-1})	4.19×10^{-2}	2.32	7.85

To evaluate the impact of H_2O on TTF-CHO, 2 mM DCNs in 5%, 20% and 40% acetonitrile were prepared at pH = 7 and 4^{o}C. As shown in **Fig 4.5**, increasing the amount of decreases cyclic dimer concentration, and again, this change is due to an increase in the hydrolysis rate k_{-1}, as shown in **Table 4.3**.

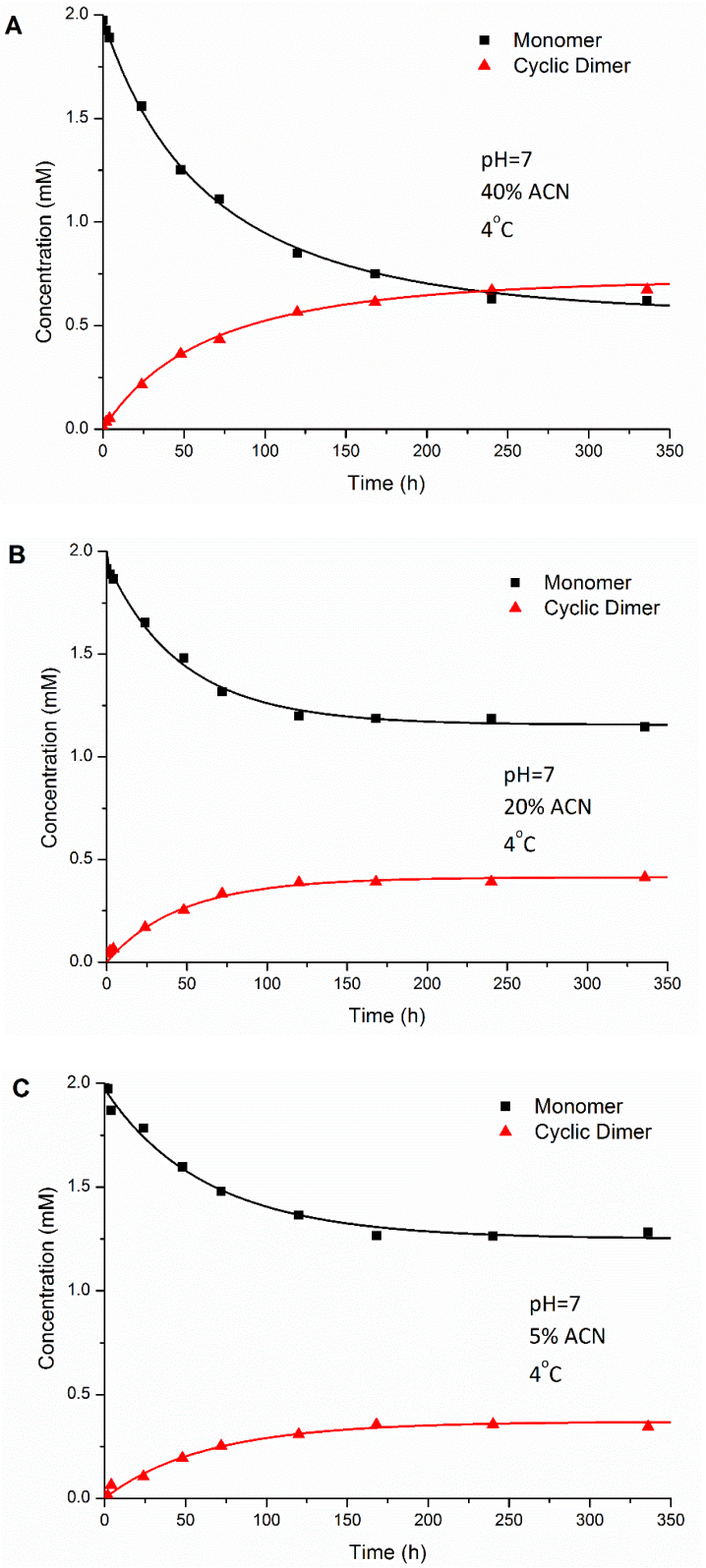


Fig 4.5 Kinetics of TTF-CHO DCNs in different solvents. Condition: TTF-CHO = 2 mM, pH = 7, 4ºC, (**A**) 40% acetonitrile in water, (**B**) 20% acetonitrile in water, (**C**) 5% acetonitrile in water. Kinetics fitting is based on expression in **Scheme 4.4** and calculated with experimental data using minimum SSE. Calculated rate constants are summarized in **Table 4.3**.

Table 4.3 Rate and equilibrium constants of TTF-CHO DCNs in different solvents

parameter	**40% ACN**	**20% ACN**	**5% ACN**
k_1/(mol^{-1}•sec^{-1})	0.87×10^{-6}	0.67×10^{-6}	0.43×10^{-6}
k_{-1}(sec^{-1})	0.38×10^{-6}	2.18×10^{-6}	1.83×10^{-6}
K/(mol^{-1})	2.29	0.307	0.235

However, higher temperatures increased the rate of both forward and backward reaction, as shown in **Fig 4.6** and **Table 4.4**. Increasing the temperature from 4ºC to 25ºC accelerates both rates to almost 10 fold. The equilibrium constant decreases at higher temperature, consistent with our observation in Chapter 2. As shown in **Fig 4.7**, higher temperature also increases the diversity of the DCNs.

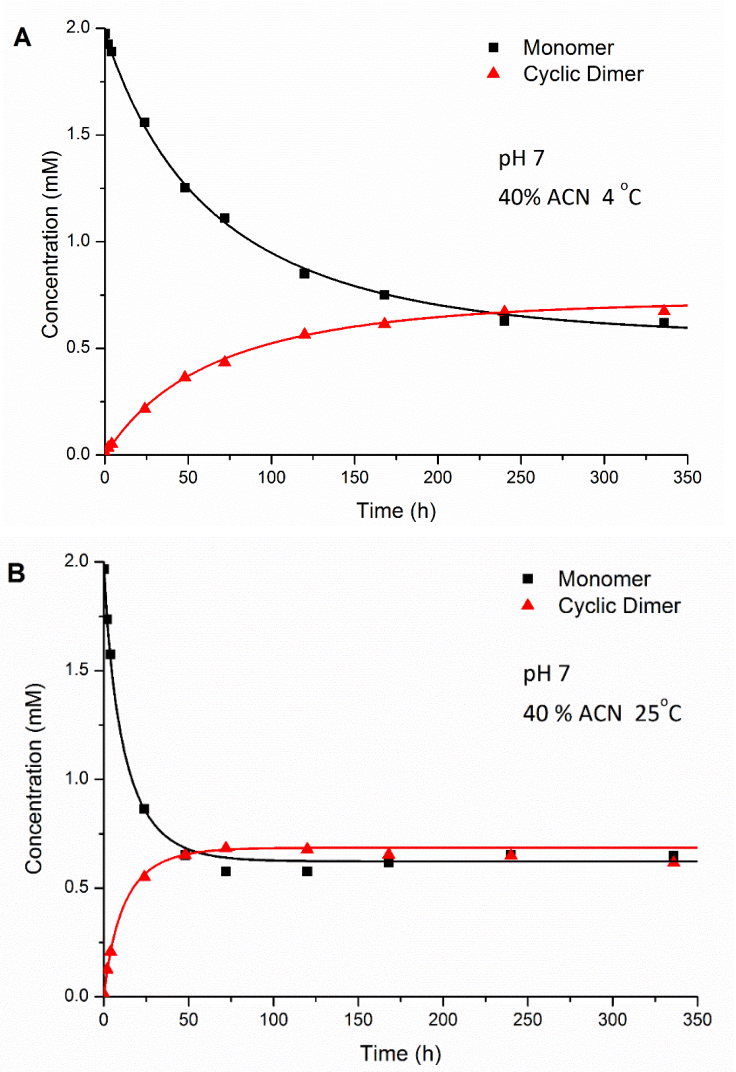


Fig 4.6 Kinetics of TTF-CHO DCNs under different temperature. Condition: TTF-CHO = 2 mM, pH = 7, 40% acetonitrile in water, (**A**) 4ºC, (**B**) 25 ºC. pH is adjusted by Et_3N. Kinetics fitting is based on expression in **Scheme 4.4** and calculated with experimental data using minimum SSE. Calculated rate constants are summarized in **Table 4.4.**

Table 4.4 Rate and equilibrium constant of TTF-CHO DCNs under 4°C and 25°C.

parameter	**4°C**	**25°C**
k_1/(mol^{-1}•sec^{-1})	0.86 ×10^{-6}	4.80 ×10^{-6}
k_{-1}(sec^{-1})	0.37 ×10^{-6}	2.72 ×10^{-6}
K/(mol^{-1})	2.32	1.76

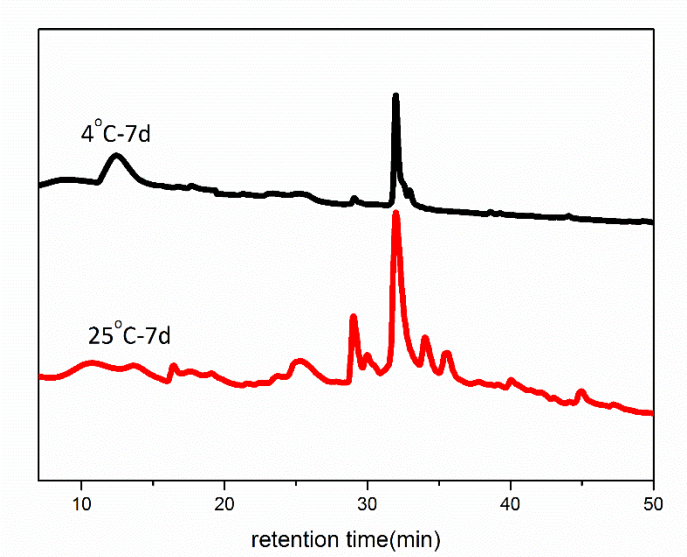
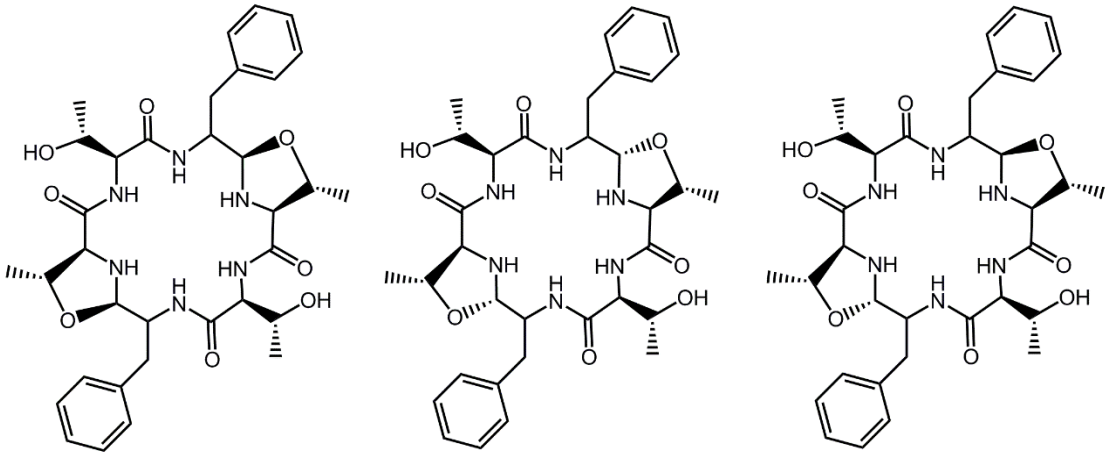


Fig 4.7 HPLC chromatograms of TTF-CHO TTF-CHO DCNs under 4°C and 25°C. Higher temperature increases the diversity of TTF-CHO network.

4.2.3 Characterization of Cyclic Dimer Structure

In Chapter 2, we found out that N,O-acetal generated between N-Boc-L-Phe-CHO and L-Thr-OMe predominantly adopts a *cis* conformation, however the N,O-acetal configuration preference is sensitive to many factors. The N,O-acetal linkages in the cyclic dimer may adopt different configurations due to ring strain and with two N,O-acetal linkages, there are three possible conformations for the cyclic dimer: (*cis*-*cis*),

(*trans-trans*) and (*cis-trans*), as shown in **Scheme 4.5.**



Scheme 4.5 Three possible cyclic dimer isomers

To probe N,O-acetal configuration in cyclic dimer, we prepared the ^{13}C-enriched monomer TTF-^{13}CHO, and generated DCN in 40% acetonitrile-d3 in D_2O (pD = 7.5). NMR Heteronuclear correlation experiment (HETCOR) was used to study the N,O-acetal configuration in cyclic dimer. At day 1, the crosspeak at 90.3 and 5.1 ppm is assigned as the correlation between ^{13}C and hydrated aldehyde proton respectively, as shown in **Fig 4.8A**. At day 7, when the DCN nears equilibrium, a new single crosspeak at 90.3 and 4.8 ppm is assigned as the to N,O-acetal on cyclic dimer (**Fig 4.8B**). The single resonance for the N,O-acetal linkage on cyclic dimer suggests the symmetry of two N,O-acetal linkages are identical, arguing against the possibility of (*trans-cis*) conformation.

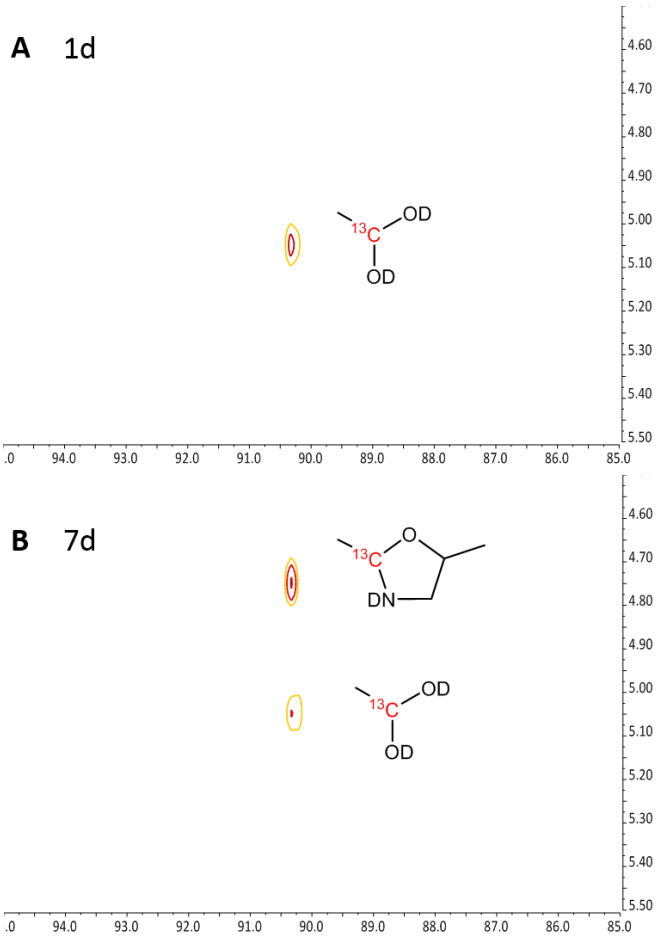


Fig 4.8 HETCOR experiment indicates two N,O-acetals on cyclic dimer are identical. (**A**) On day 1, the crosspeak (90.3, 5.1 ppm) corresponding to correlation between ^{13}C and hydrated aldehyde proton. (**B**) On day 7, only one new crosspeak (90.3, 4.8 ppm) corresponding to N,O-acetal emerged. The single resonance for the N,O-acetal linkage on cyclic dimer suggests two N,O-acetal linkages on cyclic dimer are identical, rules out (*cis-trans*) conformation.

Surface potential energy minimization using Maestro for both (*cis-cis*) and (*trans-trans*) cyclic dimers is shown in **Fig 4.9**. The (*trans-trans*) cyclic dimer is energetically more favored (-318.043 kJ/mol) than that of (*cis-cis*) (-281.569 kJ/mol), suggesting both N,O-acetal adopts a *trans* configuration to generate a (*trans-trans*)- cyclic dimer. These results suggest that the cyclic dimer structure favors the *trans* as stable acetal, while the *cis* configuration formed in model study is less stable in cyclic structures.

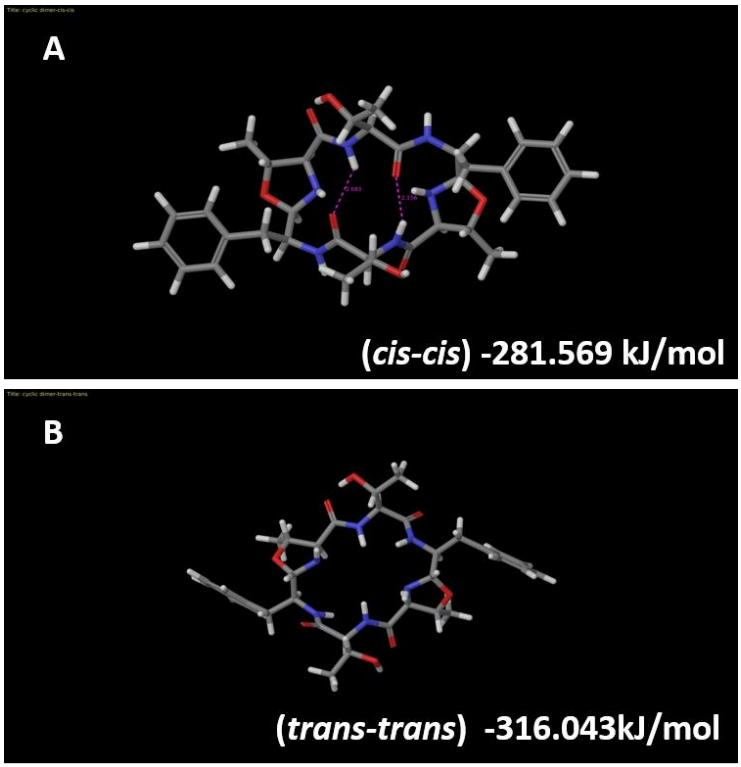


Fig 4.9 Simulation by surface energy minimization suggests cyclic dimer adopts (*tran-trans*) conformation. Energy minimization with Maestro calculates the (*trans-trans*) configuration is energetically more favored (-318.043 kJ/mol) than (*cis-cis*) configuration (-281.569 kJ/mol)

4.2.4 Cyclic Dimer Fibers Emerge in TTF-CHO DCN

Across almost all conditions we tested here, cyclic dimer is the major component of the DCN. No self-assembly structure can be detected in 40% or 20% acetonitrile in water, suggesting that the cyclic dimer has a low propensity for self-assembly. Interestingly, in 5% acetonitrile in water, pH = 7-7.5 and 4^{o}C, self-assembled fibers were detected by TEM. As shown in **Fig 4.10**, particles emerged after 24h and grew into fibers with width of 6.7 ± 1.0 nm (**Fig 4.11**). Even though only ~50% of TTF-CHO transformed into cyclic dimer after the DCN reached equilibrium, these assemblies can be pelleted by centrifuge, and as shown in **Fig 4.12**, re-injection of pellets into HPLC gives

predominantly cyclic dimers, suggesting cyclic dimer is the self-assembly species. No self-assembly can be detected under pH = 6 when DCN remains mostly monomers, nor under pH = 9, when DCN species is complicated with no dominant species.

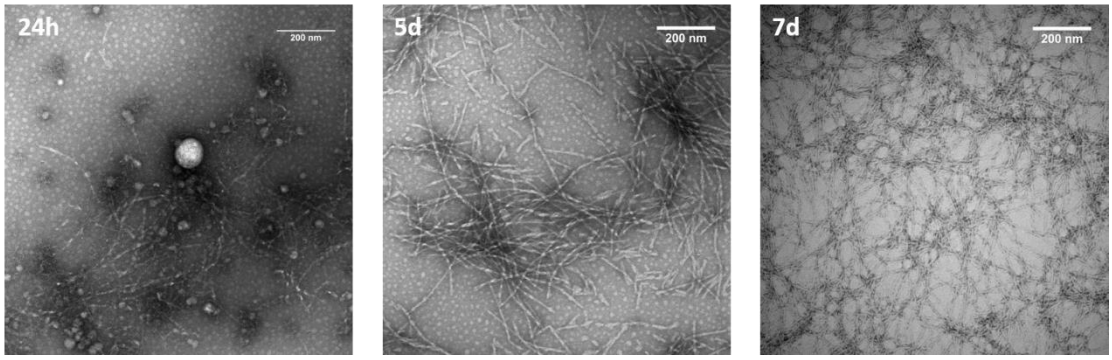


Fig 4.10 TEM image of self-assembled fibers in TTF-CHO DCN in 5% acetonitrile in water, at pH = 7 and 4°C. Particles as well as some fibers start to form at 24h, and gradually all transform into fibers.

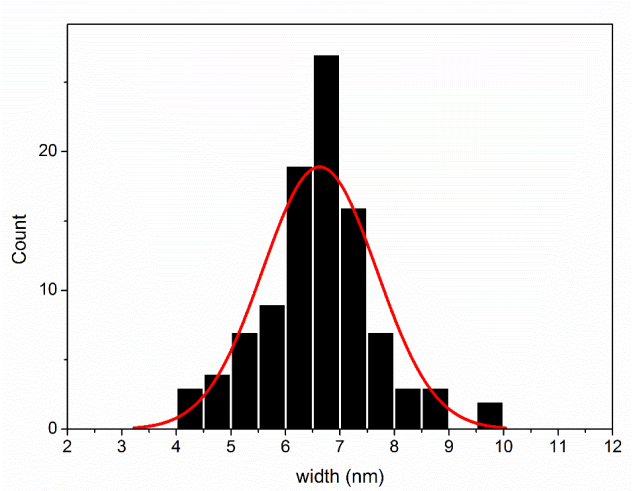


Fig 4.11 Histogram of cyclic dimer fiber width. Width is measured with ImageJ counting for 200 fibers, and fitted with a Gaussian distribution as shown in red line. The width is 6.7 ±1.0nm.

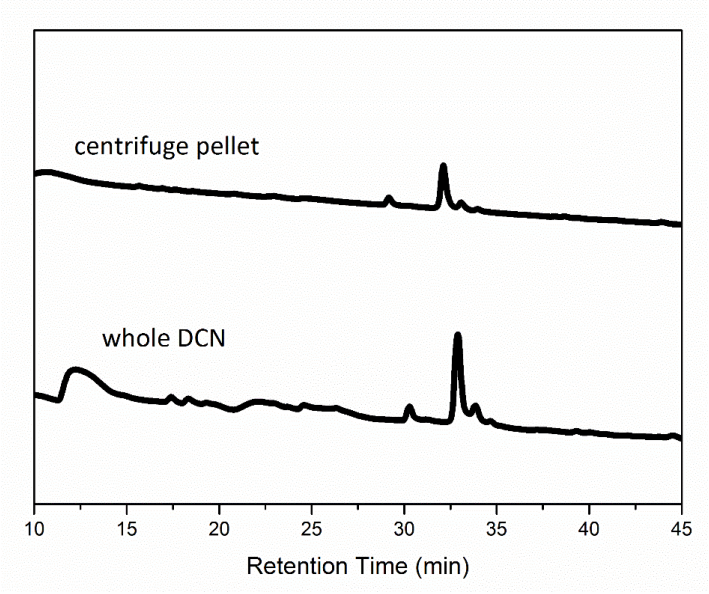
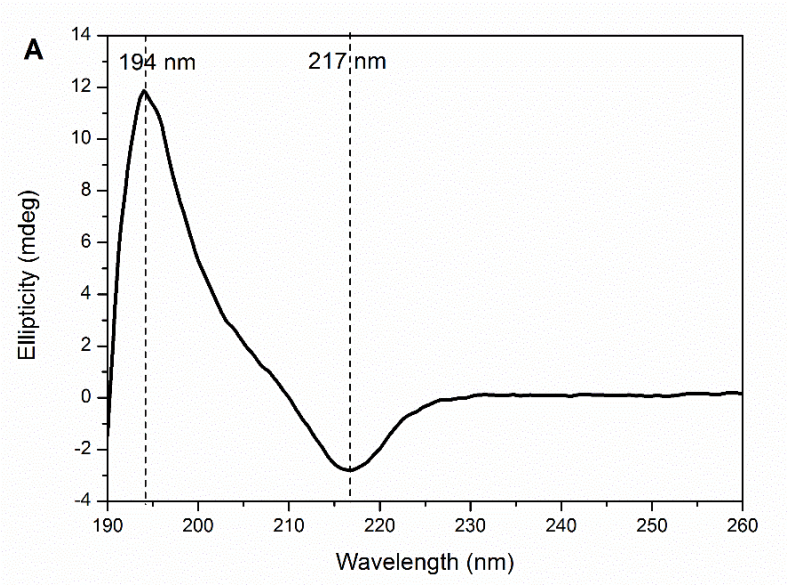
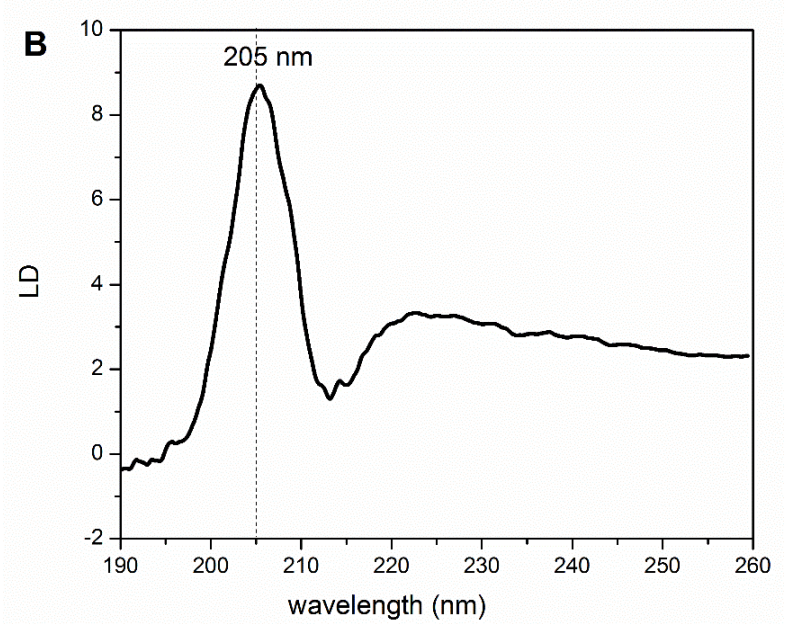


Fig 4.12 HPLC chromatograms of TTF-CHO whole DCN and centrifuged pellet. DCN incubation condition: 5% acetonitrile in water, pH = 7 and 4°C. Re-injection of centrifuged and re-suspended pellets gives cyclic dimer as dominant component.

Structural characterization of cyclic dimer fibers using Circular Dichroism, Linear Dichroism and FT-IR are shown in **Fig 4.13**. Circular Dichroism analysis of centrifuge-enriched assemblies suggests classical β-sheet signature (**Fig 4.13A**), with a negative maximum at 217 nm and a positive maximum at 194 nm. Linear Dichroism using a microvolume cuvette at a rotation speed of 2000 rpm gives a positive π-π* transition at 205 nm, and a very weak n-π* transition at 220 nm (**Fig 4.13B**), very similar to LD signals observed for assembled Aβ(1-42) fibers[15]. Furthermore, FT-IR of centrifuge-enriched assembly shows a 1642cm^{-1} Amide I transition (**Fig 4.13C**), consistent with a hydrogen-bonded β-sheet type conformation[16]. Taken together, these data suggests the $(TTF)_2$ cyclic dimer fibers adopt β-sheet alignment, similar to previously reported cyclic peptide nanotubes[17]. The previously reported cyclic peptide assemblies

containing mixed amino acid conformations are driven almost exclusively by hydrogen bonding, and can form highly ordered assembly in pure organic solvent such as DCM. This cyclic $(TTF)_2$ fiber only assembles in 5% acetonitrile in water, but not in even slightly higher concentration of acetonitrile, suggesting hydrophobic interaction plays a critical role in stabilizing self-assembly structure.


A
194 nm
217 nm
Ellipticity (mdeg)
Wavelength (nm)


B
205 nm
LD
wavelength (nm)

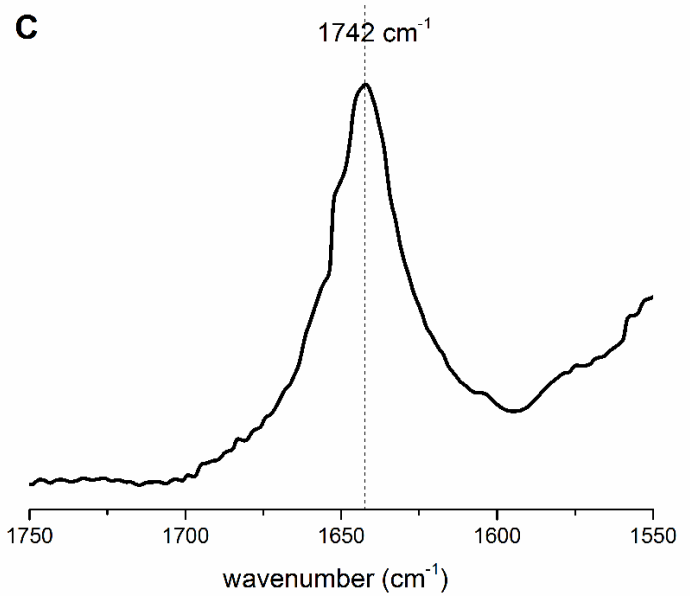

C
1742 cm-1
wavenumber (cm-1)

Fig 4.13 Spectroscopic characterization of cyclic dimer fibers suggests β-sheet conformation. (**A**) Circular Dichroism shows β-sheet signature with a negative maximum at 217 nm and a positive maximum at 194 nm. CD is performed with centrifuge enriched fibers. (**B**) Linear Dichroism gives a positive π-π* transition at 205 nm and a weak n-π* transition at 220 nm, similar to Aβ(1-42) fibers. A microvolume cuvette is used at a rotation speed of 2000 rpm. (**C**) FT-IR shows Amide I transition at 1642cm^{-1} diagnostic for cross-β peptide conformation. Fibers are enriched by centrifuge and resuspension for FT-IR.

4.3 Conclusion

The dynamic chemical networks constructed with TTF-CHO respond differently towards environment conditions including pH, temperature and solvent composition. In this network, cyclic dimer is selected as dominant species, with other cyclic species also present, as shown in **Fig 4.14**. Structure characterization using a combination of methods of cyclic dimer suggests N,O-acetals on cyclic dimer adopt (*trans-trans*) conformation, different from *cis* conformation on linear F-T dipeptide backbone we observed in Chapter 2. The cyclic ring strain could impact the relative stability and energy preferences of two configurations. In 40% or 20% acetonitrile in water, cyclic dimer doesn't self-assemble but still persists as major component of the DCNs, suggesting it is relatively stable against acetal hydrolysis.

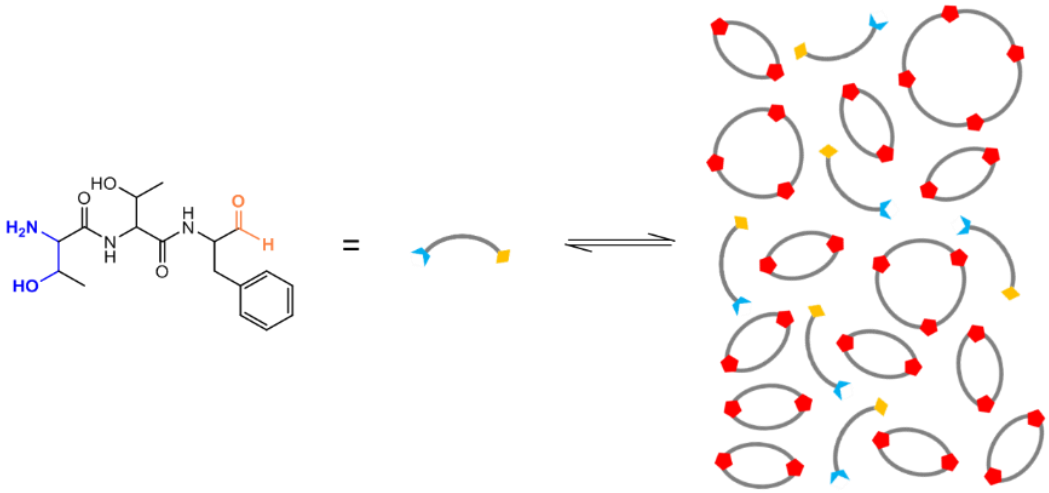


Fig 4.14 TTF-CHO DCNs. Cyclic dimer is selected as dominant species in this network, along with other cyclic species.

Interestingly, in 5% acetonitrile, cyclic dimers self-assembles into fibers, as shown in **Fig 4.15**. Structural characterization suggests β-sheet propagation in assembly, similar to previously reported cyclic peptide nanotubes. Hydrophobic interaction is a critical driven force for cyclic dimer self-assembly, clearly differentiate it from other cyclic peptide nanotubes primarily stabilized by hydrogen-bonding[18].

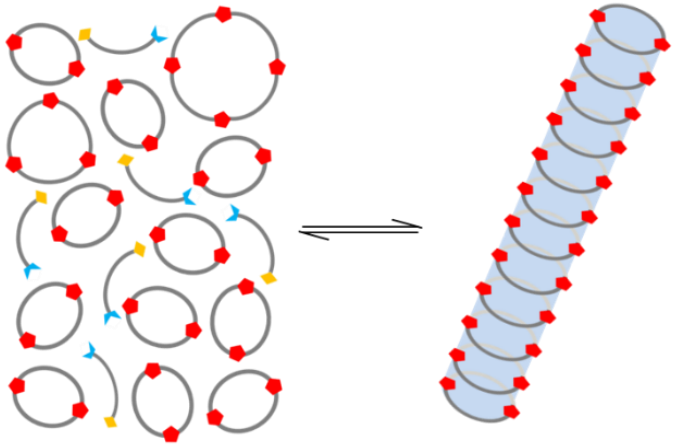


Fig 4.15 Selection of cyclic dimer self-assembly in TTF-CHO DCN

DCNs built from TTF-CHO building blocks establishes responsiveness towards environmental inputs. Dominance of cyclic dimer and lack of linear species prevent us from selecting for prion-like peptide assembly behavior, which are usually composed of linear peptides. This problem can be addressed by introducing a secondary component into the DCN, and is discussed in detail in Chapter 5.

4.4 Experimental

4.4.1 General Methods

All chemical reagents and solvents are purchased from Sigma-Aldrich, Fisher, AnalSpec, Alfa-Aesar and CIL without further purification unless otherwise noted. All reactions are performed under N_2 protection unless otherwise noted. All 1D and 2D NMR was performed on INOVA 400MHz or INOVA 600MHz NMR spectrometer.

Circular Dichroism Spectropolarimetry. CD spectra were recorded on a Jasco J-810 CD spectropolarimeter in 0.10 mm quartz cells. Spectra were recorded from 260 to 190 nm at a scanning rate of 100 nm/min and a resolution of 0.5 nm.

Flow Linear Dichroism Spectroscopy. Flow linear dichroism spectra were recorded on a JASCO J-810 circular dichroism spectropolarimeter using a microvolume cuvette with a path length of 50 μm and a rotation speed of 2000 rpm to establish Couette flow. The background scattering for each sample was obtained from the LD spectra of samples at 0 rpm. The LD spectra were measured after 5 min of rotation.

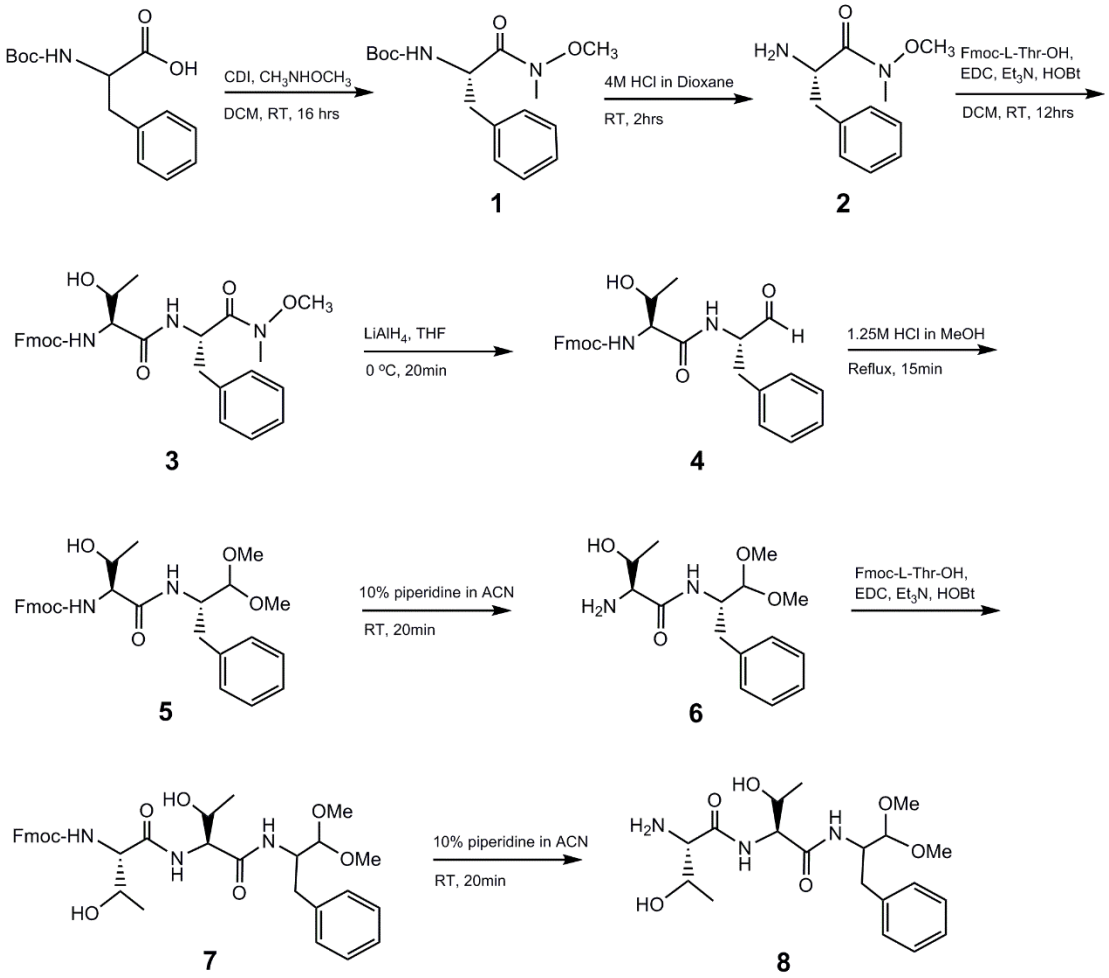
Transmission Electron Microscopy. All TEM are performed by using a Hitachi H-7500 transmission electron microscope to image the samples at 75kV. A TEM copper grid with a 200 mesh carbon support (Electron Microscopy Sciences) was covered with 10 μL of DCN for 1 min before wicking the excess solution with filter paper. 10 μL of 2% uranyl acetate was added and incubated for 2 min, excess solution was wicked away, and the grids were placed in desiccators to dry under vacuum overnight.

Attenuated Total Reflectance Fourier Transform Infrared (FT-IR).FT-IR spectra were acquired using a Jasco FT-IR 4100 at room temperature and averaging 500 to 1000 scans with 2 cm^{-1} resolution, using an MCT detector with a 5 mm aperture and a scanning speed of 4 mm/sec. Aliquots (10 μL) of DCN or peptide solution were dried as thin films on a Pike Galdi ATR diamond crystal.

X-ray powder diffraction (XRD). The samples were lyophilized to yield dry powder. The spectra were obtained by loading the powder into a 0.2 mm mylar capillary and the diffraction patterns acquired on a Bruker APEX-II diffractometer with graphite monochromated Cu radiation K-alpha radiation, λ=1.54184 Å, 40 kV and 35 mA, with a 0.5 pinhole collimator. Exposure times were typically 300s per frame. The data integration software XRD2DSCAN was used to convert the two dimensional data into a theta-2theta scan.

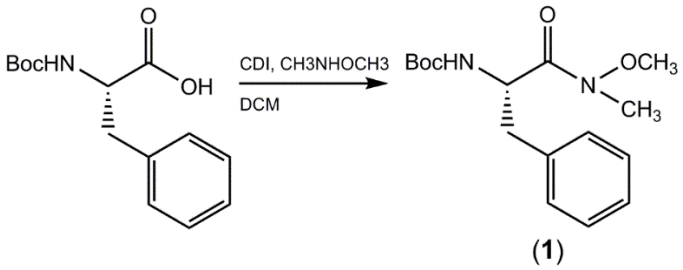
4.4.2 Synthesis of TTF-CHO

The synthetic route of TTF-CHO is summarized in **Scheme 4.6**. Boc-protected Phenylalanine was firstly converted to Weinreb Amide. After deprotecting Boc, EDC coupling was used to couple an Fmoc-protected Threonine. Then the TF-Weinreb Amide is reduced into TF-CHO using LAH reduction, and protected as acetal. Another molecule of Fmoc-Thr was coupled, and final product was generated after deprotection of Fmoc and purified by HPLC.



Scheme 4.6 Detailed synthetic route of TTF-acetal

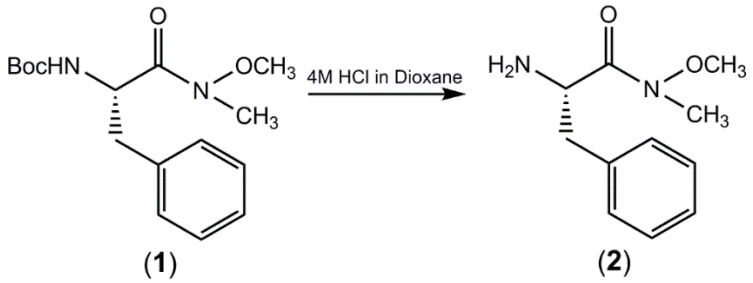
1. *Boc-Phe-Weinreb amide (**1**):*



(**1**)

To a solution of N-Boc-L-Phe 10.61 g (40 mmol) into DCM was added 1,1-carbonlydiimidazole 7.13 g (1.1 eq). After stirring for 1h, 4.29g (1.1 eq) N,O-dimethylhydroxyamine hydrochloride was added. After reaction mixture was stirred overnight under N_2 protection at room temperature, solvent was removed *in vacuuo*. Residue was diluted with EtOAc, washed with 1 M HCl (aq) three times, saturated $NaHCO_3$ solution twice and brine twice. Organic layer was dried with $MgSO_4$, then solvent was removed *in vacuuo* to yield **1.** Yield: 92%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 7.44 – 7.11 (m, 5H), 5.15 (d, *J* = 7.9 Hz, 1H), 4.94 (d, *J* = 6.6 Hz, 1H), 3.65 (s, 3H), 3.16 (s, 3H), 3.05 (dd, *J* = 13.3, 5.7 Hz, 1H), 2.85 (dt, *J* = 64.4, 32.2 Hz, 1H), 1.38 (s, 9H). HRMS for $C_{16}H_{25}N_2O_4$ (M+H): calculated: 309.18143, found: 309.18130.

2. *Phe-Weinreb amide (**2**):*

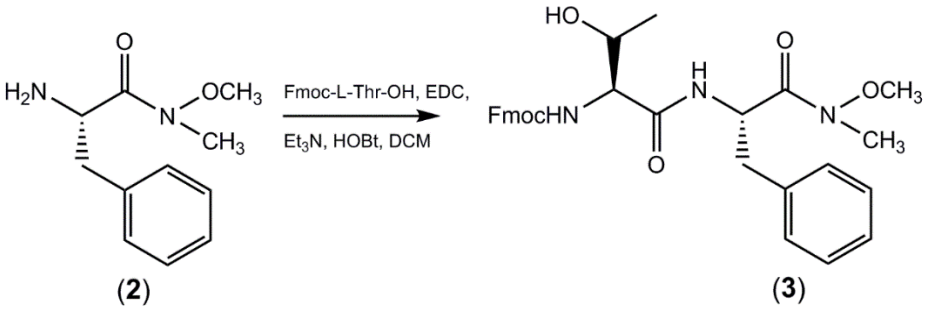


(**1**) (**2**)

1 was dissolved in 4 M HCl in dioxane and stir for 2h under room temperature then dried *in vacuuo* to give **2**. Yield: 100%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 8.56 (s, br), 7.34

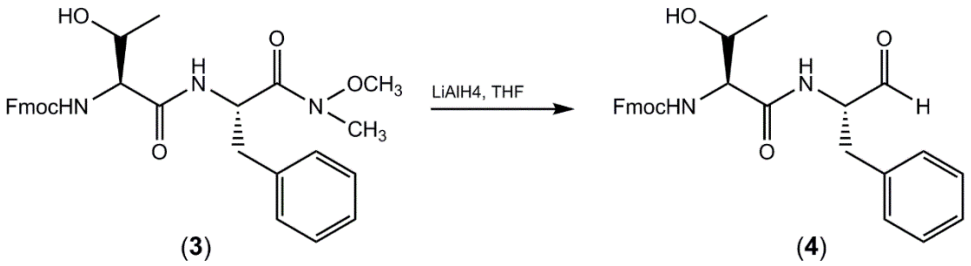
– 7.15 (m, 5H), 4.73 (d, *J* = 5.1 Hz, 1H), 3.64 (s, 1H), 3.35 (tt, *J* = 20.6, 6.8 Hz, 1H), 3.10 (s, 3H). HRMS for $C_{11}H_{17}N_2O_2$ (M+H): calculated: 209.12900; found: 209.12824

*3. Fmoc-Thr-Phe-Weinreb amide (**3**):*



1 g of **2** (4.09 mmol) was dissolved in DCM. 1.2 mL (2.1 eq) of Et_3N was added, followed by 0.8626 g (1.1 eq) of N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC). The solution turned slurry. 1.534g (1.1eq) of Fmoc-L-Thr-OH and 0.687g (1.1 eq) of HOBt were subsequently added and reaction mixture was stirred for 6h under RT. Wash with 1 M HCl (aq) twice, saturated $NaHCO_3$ solution twice then brine once, combined the organic layer, dried with $MgSO_4$ and removed the solvent under reduced pressure to give **3**. Yield: 80%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.59 (d, *J* = 7.3 Hz, 2H), 7.44 – 7.08 (m, 9H), 6.82 (d, *J* = 7.6 Hz, 1H), 5.56 (d, *J* = 7.9 Hz, 1H), 5.18 (m, 1H), 4.47 (dd, *J* = 10.6, 7.2 Hz, 1H), 4.36 – 4.29 (m, 1H), 4.25 (s, 1H), 4.21 (t, *J* = 7.0 Hz, 1H), 4.12 (dd, *J* = 11.3, 5.9 Hz, 1H), 3.76 (s, 3H), 3.21 (s, 3H), 3.14 (dd, *J* = 13.8, 5.0 Hz, 1H), 3.05 (s, 1H), 2.91 – 2.77 (m, 1H), 1.11 (t, *J* = 9.2 Hz, 3H). HRMS for $C_{30}H_{34}N_3O_6$ (M+H): calculated: 532.24476; found: 532.24424.

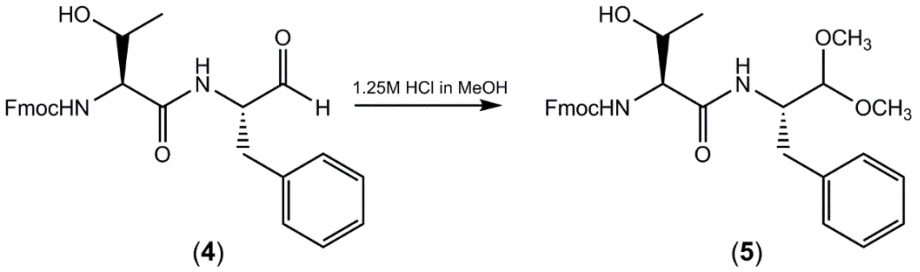
*4. Fmoc-Thr-Phe-aldehyde (**4**):*



(**3**) (**4**)

2.173 g **3** (4.08 mmol) was dissolved in THF under N_2 protection, then cooled down to -78ºC. 1.4 eq lithium aluminum hydride (1.0 M in THF, 5.7 mL) was added dropwise via syringe. The reaction mixture was warmed up to 0ºC and stirred for 30 min. Then, the solution was cooled down to -78ºC again and 57mL of 1 M $KHSO_4$ (aq) was added to quench the reaction. Removed THF *in vacuuo*, diluted residue with EtOAc, washed with saturated $NaHCO_3$ solution and brine. Combined the organic layer, dried with $MgSO_4$ then concentrated to yield aldehyde **4**. Yield: 91%. ^{1}H NMR (600 MHz, $cdcl_3$) δ 9.63 (s, 1H), 7.78 (d, *J* = 7.4 Hz, 2H), 7.60 (m, 2H), 7.39-7.10 (m, 9H), 5.61 (d, *J* = 8.0 Hz, 1H), 4.73 (dd, *J* = 13.8, 7.0 Hz, 1H), 4.48 (dd, *J* = 10.6, 6.9 Hz, 1H), 4.39 – 4.28 (m, 1H), 4.12 (dd, *J* = 14.2, 7.1 Hz, 1H), 3.12 (ddd, *J* = 21.5, 14.2, 6.8 Hz, 2H), 1.14 (t, *J* = 11.2 Hz, 3H). HRMS for $C_{28}H_{29}N_2O_5$ (M+H): calculated: 473.20765, found: 473.20789

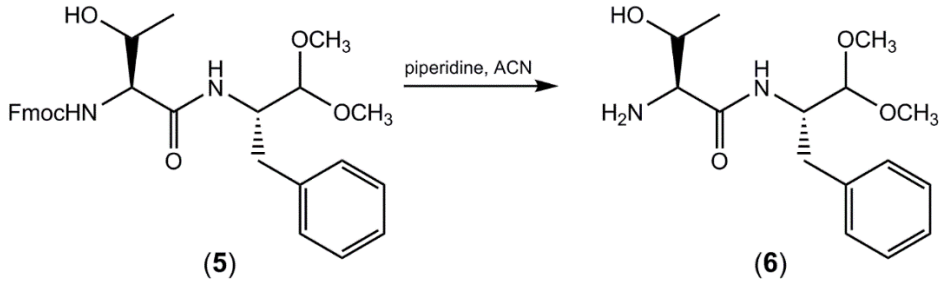
*5. Fmoc-Thr-Phe-acetal (**5**):*



(**4**) (**5**)

1.234 g **4** was dissolved in 20 mL 1.25 M HCl in MeOH, reflux at 75ºC for 15 min, then cool down to RT and quench with saturated $NaHCO_3$ aqueous solution. MeOH was removed. Diluted residue with EtOAc, washed with sat $NaHCO_3$ then water, dried, concentrated and purified by silica gel flash column (Hexanes : EtOAc = 1:3) to give **5**. Yield: 94%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 7.78 (d, *J* = 7.4 Hz, 2H), 7.60 (d, *J* = 7.1 Hz, 2H), 7.42 -7.15 (m, 9H), 6.43 (d, *J* = 8.9 Hz, 1H), 5.53 (d, *J* = 8.2 Hz, 1H), 4.50 (dd, *J* = 10.4, 7.0 Hz, 1H), 4.43 (tdd, *J* = 9.2, 5.7, 3.3 Hz, 1H), 4.34 (dd, *J* = 10.5, 7.1 Hz, 1H), 4.27 – 4.16 (m, 2H), 4.03 (d, *J* = 8.2 Hz, 1H), 3.44 (s, 3H), 3.39 (s, 3H), 2.95 (dd, *J* = 14.0, 5.9 Hz, 1H), 2.86 (s, 1H), 2.71 (dd, *J* = 13.9, 9.0 Hz, 1H), 1.11 (d, *J* = 6.4 Hz, 3H). HRMS for $C_{30}H_{35}N_2O_6$ (M+H): calculated: 519.24951, found: 519.24697.

*6. Thr-Phe-acetal (**6**):*

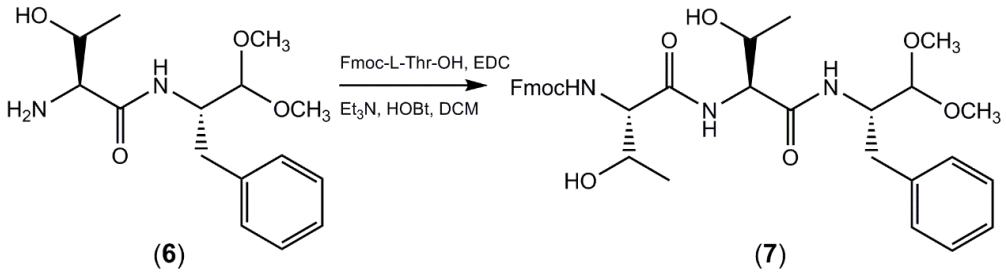


(**5**) (**6**)

1.104 g **5** was suspended in 10 mL acetonitrile. 2 mL piperidine was added and stirred for 20 min under room temperature. Removed solvent *in vacuuo* to obtain **6**. Yield:

70%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 7.45 (d, *J* = 9.7 Hz, 2H), 7.37 – 7.21 (m, 2H), 7.19 (dd, *J* = 10.4, 4.2 Hz, 3H), 4.48 – 4.39 (m, 1H), 4.18 (t, *J* = 41.9 Hz, 1H), 4.07 (dq, *J* = 9.6, 3.2 Hz, 1H), 3.46 (s, 3H), 3.42 (s, 3H), 3.14 (d, *J* = 3.1 Hz, 1H), 3.04 – 2.93 (m, 1H), 2.72 (dd, *J* = 14.0, 9.7 Hz, 1H), 1.01 (d, *J* = 6.5 Hz, 3H). HRMS for $C_{15}H_{25}N_2O_4$ (M+H): calculated: 297.18143, found: 297.18097.

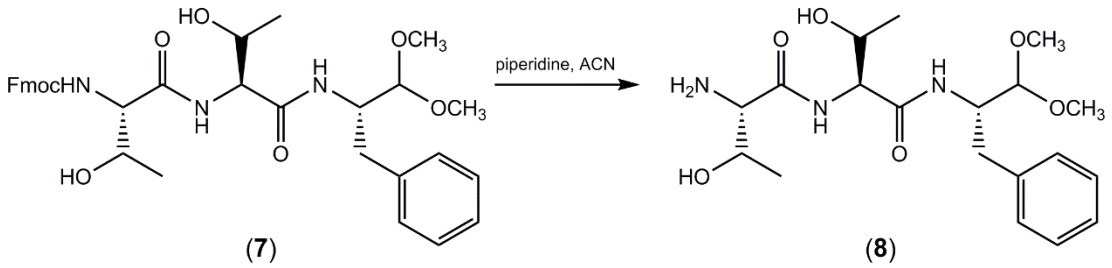
*7. Fmoc-Thr-Thr-Phe-acetal (**7**):*



(**6**) (**7**)

Same procedure as step 3. After removing the solvent, the product was purified by flash column with DCM : MeOH = 50:1 to 20:1. Yield: 74%. ^{1}H NMR (600 MHz, cdcl$_3$) δ 7.75 (d, *J* = 7.5 Hz, 2H), 7.57 (d, *J* = 7.4 Hz, 2H), 7.39-7.01 (m, 9H), 6.63 (d, *J* = 9.0 Hz, 1H), 5.72 (d, *J* = 7.7 Hz, 1H), 4.53 – 4.32 (m, 3H), 4.33 – 4.15 (m, 5H), 3.42 (s, 3H), 3.36 (s, 3H), 2.92 (dd, *J* = 14.0, 5.7 Hz, 1H), 2.74 (dd, *J* = 13.9, 8.8 Hz, 1H), 1.14 (d, *J* = 6.3 Hz, 3H), 1.06 (d, *J* = 6.4 Hz, 3H). HRMS for $C_{34}H_{42}N_3O_8$ (M+H): calculated: 620.29719, found: 620.29749.

*8. Thr-Thr-Phe-acetal (**8**):*



(**7**) (**8**)

Same procedure as step 6. Purified by HPLC to give final product **8**. Yield: 78%. ^{1}H NMR (400 MHz, dmso) δ 8.00 (d, *J* = 8.5 Hz, 1H), 7.79 (d, *J* = 8.9 Hz, 1H), 7.31 – 7.13 (m, 4H), 4.81 (d, *J* = 4.8 Hz, 1H), 4.76 (d, *J* = 4.0 Hz, 1H), 4.24 (d, *J* = 4.5 Hz, 1H), 4.17 (dd, *J* = 8.0, 4.0 Hz, 1H), 4.10 (ddd, *J* = 13.2, 9.1, 4.1 Hz, 1H), 4.00 – 3.92 (m, 1H), 3.89 (dd, *J* = 11.0, 4.9 Hz, 1H), 3.38 (s, 3H), 3.35 (s, 3H), 3.08 (d, *J* = 4.2 Hz, 1H), 2.88 (dd, *J* = 14.1, 3.5 Hz, 1H), 2.67 (dd, *J* = 14.1, 9.8 Hz, 1H), 1.07 (d, *J* = 6.4 Hz, 3H), 0.97 (t, *J* = 14.1 Hz, 3H). HRMS for $C_{19}H_{32}N_3O_6$ (M+H): calculated: 398.22911, found: 398.22925

4.4.3 Dynamic Chemical Networks

TTF-CHO was generated *in situ* by deprotecting TTF-acetal with 4% HCl in H_2O then dried *in vacuuo* at 20^{o}C. TTF-CHO then was dissolved into 2 mM solution with acetonitrile and water of desired ratio. The pH of solution was then adjusted by titrating aliquots of with 5% Et_3N in H_2O using a Fisher Scientific Accumet Basic AB15 pH meter. The solution was incubated under 4^{o}C or 25^{o}C to generate DCN.

Kinetics study of DCN was performed by taking out small fraction of DCN solution for

HPLC analysis over 2-4 weeks. All HPLC was performed on a Waters Delta 600 HPLC equipped with a photodiode array UV/Vis detector and a reverse-phase HPLC column (Kromasil 100-5C18, 4.6 x 250 mm). HPLC solvent gradient: acetonitrile: 10 mM TEAA in H_2O: 0-90 min: 10% to 100%, 90-100 min: 100%, 100-101 min: 100% to 10%, 101-110 min: 10%. UV absorbance at 258 nm was recorded and integrated for kinetics analysis, assuming absorption coefficient for phenylalanine is additive for all species. Species identification was conducted by ESI mass spectrometry at Emory University Mass Spectrometry Center.

References:

1. Chernoff, Y. O., Mutations and Natural Selection in the Protein World. *Journal of Molecular Biology* **2011,** *413* (3), 525-526.

2. Ghaemmaghami, S.; Watts, J. C.; Nguyen, H.-O.; Hayashi, S.; DeArmond, S. J.; Prusiner, S. B., Conformational Transformation and Selection of Synthetic Prion Strains. *Journal of Molecular Biology* **2011,** *413* (3), 527-542.

3. Weissmann, C., The state of the prion. *Nat Rev Micro* **2004,** *2* (11), 861-871.

4. Jahn, T. R.; Makin, O. S.; Morris, K. L.; Marshall, K. E.; Tian, P.; Sikorski, P.; Serpell, L. C., The Common Architecture of Cross-β Amyloid. *Journal of Molecular Biology* **2010,** *395* (4), 717-727.

5. Childers, W. S.; Anthony, N. R.; Mehta, A. K.; Berland, K. M.; Lynn, D. G., Phase

Networks of Cross-β Peptide Assemblies. *Langmuir* **2012,** *28* (15), 6386-6395.

6. (a) Takahashi, Y.; Mihara, H., Construction of a chemically and conformationally self-replicating system of amyloid-like fibrils. *Bioorganic & Medicinal Chemistry* **2004,** *12* (4), 693-699; (b) Rubinov, B.; Wagner, N.; Rapaport, H.; Ashkenasy, G., Self-Replicating Amphiphilic β-Sheet Peptides. *Angewandte Chemie International Edition* **2009,** *48* (36), 6683-6686; (c) Rajagopal, K.; Schneider, J. P., Self-assembling peptides and proteins for nanotechnological applications. *Current Opinion in Structural Biology* **2004,** *14* (4), 480-486; (d) Adler-Abramovich, L.; Gazit, E., The physical properties of supramolecular peptide assemblies: from building block association to technological applications. *Chemical Society Reviews* **2014,** *43* (20), 6881-6893.

7. Sadownik, J. W.; Ulijn, R. V., Dynamic covalent chemistry in aid of peptide self-assembly. *Current Opinion in Biotechnology* **2010,** *21* (4), 401-411.

8. Frederix, P. W. J. M.; Ulijn, R. V.; Hunt, N. T.; Tuttle, T., Virtual Screening for Dipeptide Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *The Journal of Physical Chemistry Letters* **2011,** *2* (19), 2380-2384.

9. Williams, R. J.; Smith, A. M.; Collins, R.; Hodson, N.; Das, A. K.; Ulijn, R. V., Enzyme-assisted self-assembly under thermodynamic control. *Nat Nano* **2009,** *4* (1), 19-24.

10. FrederixPim, W. J. M.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T., Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nat Chem* **2015,** *7* (1), 30-37.

11. Mentzel, M.; Hoffmann, H. M. R., N-methoxy-N-methylamides (Weinreb amides) in modern organic synthesis. *Journal für Praktische Chemie/Chemiker-Zeitung* **1997,** *339* (1), 517-524.

12. Nahm, S.; Weinreb, S. M., N-methoxy-n-methylamides as effective acylating agents. *Tetrahedron Letters* **1981,** *22* (39), 3815-3818.

13. Montalbetti, C. A. G. N.; Falque, V., Amide bond formation and peptide coupling. *Tetrahedron* **2005,** *61* (46), 10827-10852.

14. (a) Bergmann, E. D., The Oxazolidines. *Chemical Reviews* **1953,** *53* (2), 309-352; (b) Fife, T. H.; Hagopian, L., Oxazolidine hydrolysis. Participation of solvent and buffer in ring opening. *Journal of the American Chemical Society* **1968,** *90* (4), 1007-1014; (c) Fife, T. H.; Hutchins, J. E. C., General-acid-catalyzed ring opening of oxazolidines. Hydrolysis of 2-[4-(dimethylamino)styryl]-N-phenyl-1,3-oxazolidine. *The Journal of Organic Chemistry* **1980,** *45* (11), 2099-2104.

15. Dafforn, T. R.; Rajendra, J.; Halsall, D. J.; Serpell, L. C.; Rodger, A., Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell. *Biophysical Journal* **2004,** *86* (1), 404-410.

16. Kong, J.; Yu, S., Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochimica et Biophysica Sinica* **2007,** *39* (8), 549-559.

17. (a) Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N., Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **1993,** *366* (6453), 324-327; (b) Amor ń, M.; Castedo, L.; Granja, J. R., New Cyclic Peptide Assemblies with Hydrophobic Cavities: The Structural and Thermodynamic Basis of

a New Class of Peptide Nanotubes. *Journal of the American Chemical Society* **2003,** *125* (10), 2844-2845.

18. Amor ń, M.; Castedo, L.; Granja, J. R., Folding Control in Cyclic Peptides through N-Methylation Pattern Selection: Formation of Antiparallel β-Sheet Dimers, Double Reverse Turns and Supramolecular Helices by 3α,γ Cyclic Peptides. *Chemistry – A European Journal* **2008,** *14* (7), 2100-2111.

Chapter Five

TTF-CHO + $(TTF)_2$ Network: Self-templated and External Templated Replication

5.1 Introduction

In Chapter 4, we constructed dynamic chemical networks built from modified tripeptide monomer Thr-Thr-Phe-CHO (TTF-CHO) and demonstrated network responsiveness towards diverse environmental inputs. In this chapter, we introduced a second component, linear peptide $(TTF)_2$ into the TTF-CHO network to generate the linear product TTF*ox*TTFTTF (*ox* represents oxazolidine, or N,O-acetal) for selection of self-templated and external templated conformational replication mimicking prion-like infectious propagation.

Prion-like infectious propagation is a template-directed conformational replication process that incorporates peptides into cross-β supramolecular assemblies. Otto reported a self-replicating peptide network that selects for different assembly species in response to different shear forces[1] using reversible disulfide exchange. This network generates diversity on macrocycle size but not on peptide sequences. Recently, our lab incorporated reversible N,N-acetal on peptide backbones to construct NFF-CHO peptide DCNs[2]. As shown in **Fig 5.1**, in this network, diversity is generated on sequence level, and linear trimer is selected as self-assembly species, undergoing stages

of phase transition to assemble into fibers.

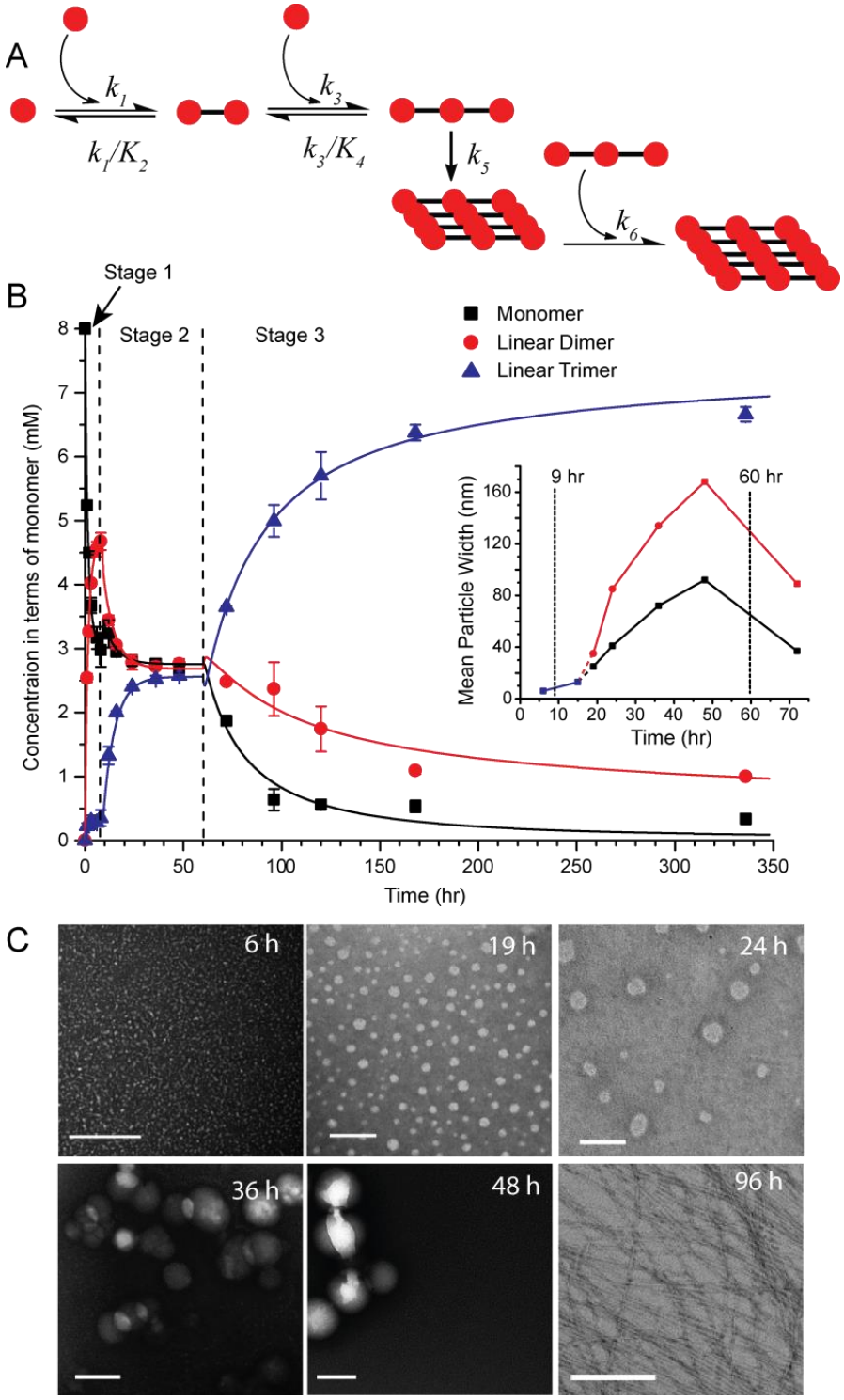


Fig 5.1 NFF-CHO dynamic peptide network. (**A**) Kinetic model for the NFF-CHO network. (**B**) Kinetic fits of network member concentrations NFF-CHO network. Inset shows the mean widths of the two particle populations (black and red) over time. (**C**) TEM micrographs of samples at the indicated time points (scale bar = 200 nm).

In this chapter, we generated linear TTF*ox*TTFTTF in a peptide DCN for selection of prion-like infectious propagation. The incorporation of single N,O-acetal linkage lead to self-assembly structure as parallel fibers, different from anti-parallel nanotubes generated by all peptide sequence $(TTF)_3$. We also explored the network responsiveness towards template diversity by introducing external templates.

5.2 Results and Discussion

5.2.1 Understanding $(TTF)_n$ Peptide Self-Assembly

A series of $(TTF)_n$ peptides (n = 2, 3, 4) were prepared as control for the cross-β assembly propensity of these sequences. When incubated at 2 mM under pH = 8, 40% acetonitrile in water at 4^oC, no self-assembly of $(TTF)_2$ peptide can be detected. $(TTF)_3$ peptide however forms ribbon-like assemblies structure within 24h, and these ribbons gradually transforms into nanotubes (**Fig 5.2**) with diameter of 38.4 ± 7.4 nm (**Fig 5.3**). Further elongating peptide sequence into $(TTF)_4$ were not further explored due to limitation in solubility.

Findings that $(TTF)_2$ peptide does not assemble while $(TTF)_3$ peptide assembles into tubes suggest an approach to test the impact of a single N,O-acetal on assembly propensity. We predict that if TTF-CHO and $(TTF)_2$ peptide can condensate into N,O-acetal linked linear trimer, TTF*ox*TTFTTF (*ox* represents oxazolidine, or N,O-acetal linkage), this product would be selected as self-propagating species, and assemble into nanotubes, similar to $(TTF)_3$ peptide.

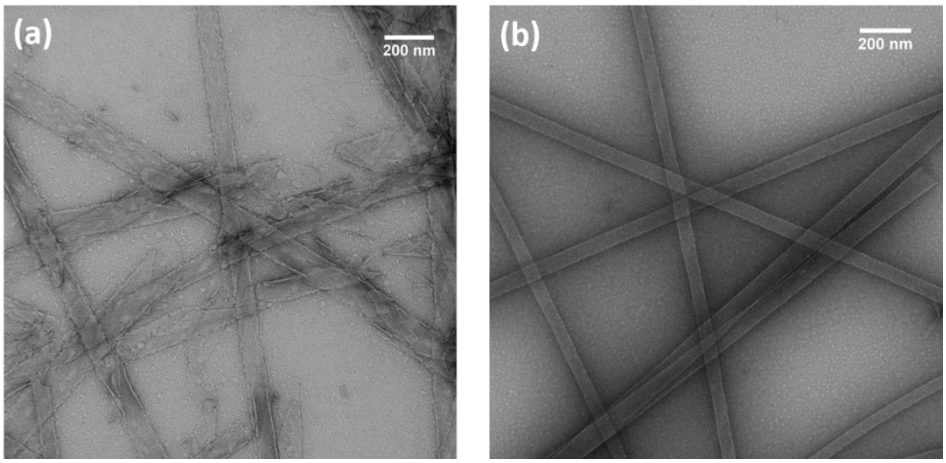


Fig 5.2 TEM images show $(TTF)_3$ peptide self-assemble into nanotubes. Under pH = 7-8, 40% acetonitrile in water at 4°C, $(TTF)_3$ peptide forms (**a**) ribbons at day 1, then transforms into mature nanotubes (**b**) after 2 - 4 weeks.

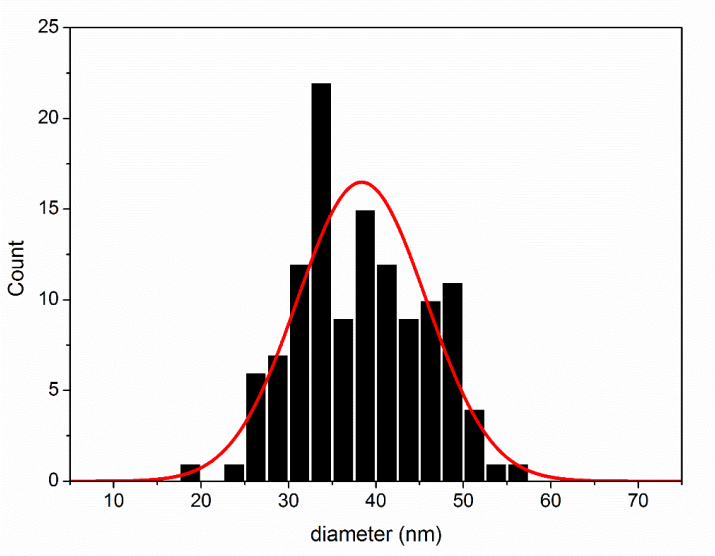


Fig 5.3 Histogram and Gaussian distribution of $(TTF)_3$ nanotube diameter. Nanotube width is measured with ImageJ counting 120 tubes, and calculate back to diameter. Gaussian distribution calculates the average diameter to be 38.4 ± 7.4 nm.

5.2.2 Structural Characterization of $(TTF)_3$ nanotubes

As shown in **Fig 5.4**, AFM defines the flattened tube height to be 6.8 ± 0.1 nm, thinner than E22L peptide nanotubes previously reported by our group considering the longer peptide sequence[3]. The thinner wall might suggest this nanotube is composed by peptide monolayer instead of bilayer, or the bilayer is orientated with tilt angle similar to E22V nanotubes[4].

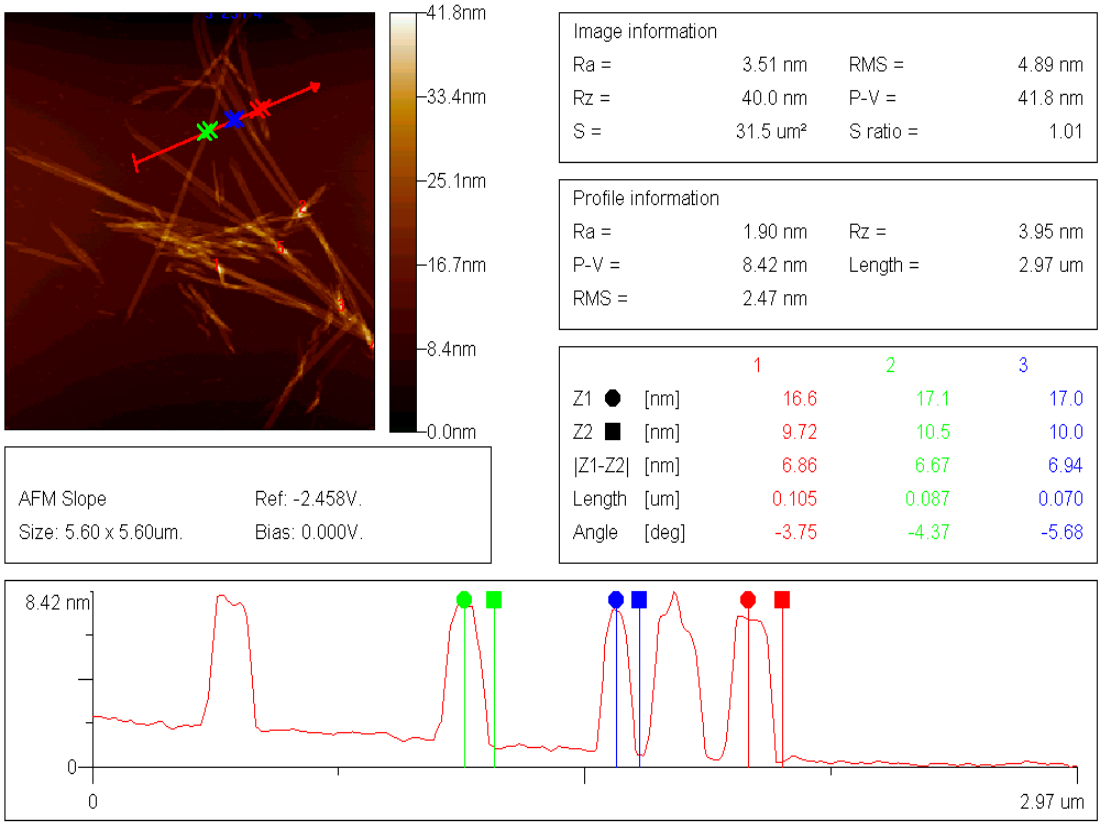


Fig 5.4 AFM measurement of $(TTF)_3$ nanotube wall thickness. Nanotube height is measured to be 6.8 ± 0.1 nm, corresponding to wall thickness of 3.4 nm.

X-ray powder diffraction, FT-IR and CD were used to investigate peptide conformation and orientation. As shown in **Fig 5.5A**, FT-IR gives a 1621 cm^{-1} Amide I transition, characteristic of β-sheets conformation[5]. The 1692 cm^{-1} overtone is proven to be

diagnostic of anti-parallel peptide chains within the assembly[6]. X-ray powder diffraction pattern supports two strong reflections signals (**Fig 5.5B**), 4.7 Å corresponding to hydrogen bonding distance in β-sheet structure, and 10.1 Å for β-sheet stacking repeat distance most characteristic of the cross-β assembly (lamination distance)[7]. The $(TTF)_3$ nanotubes give a unique CD spectrum (**Fig 5.5 C**), with a negative maximum at 197 nm, and a very strong positive maximum at 217 nm, almost a mirror image of typical β-sheet. With FT-IR and X-ray diffraction data consistent with cross-β conformation, it is unclear what causes this unique CD ellipticity. Similar CD signal was previously observed in E22V peptide assemblies[4].

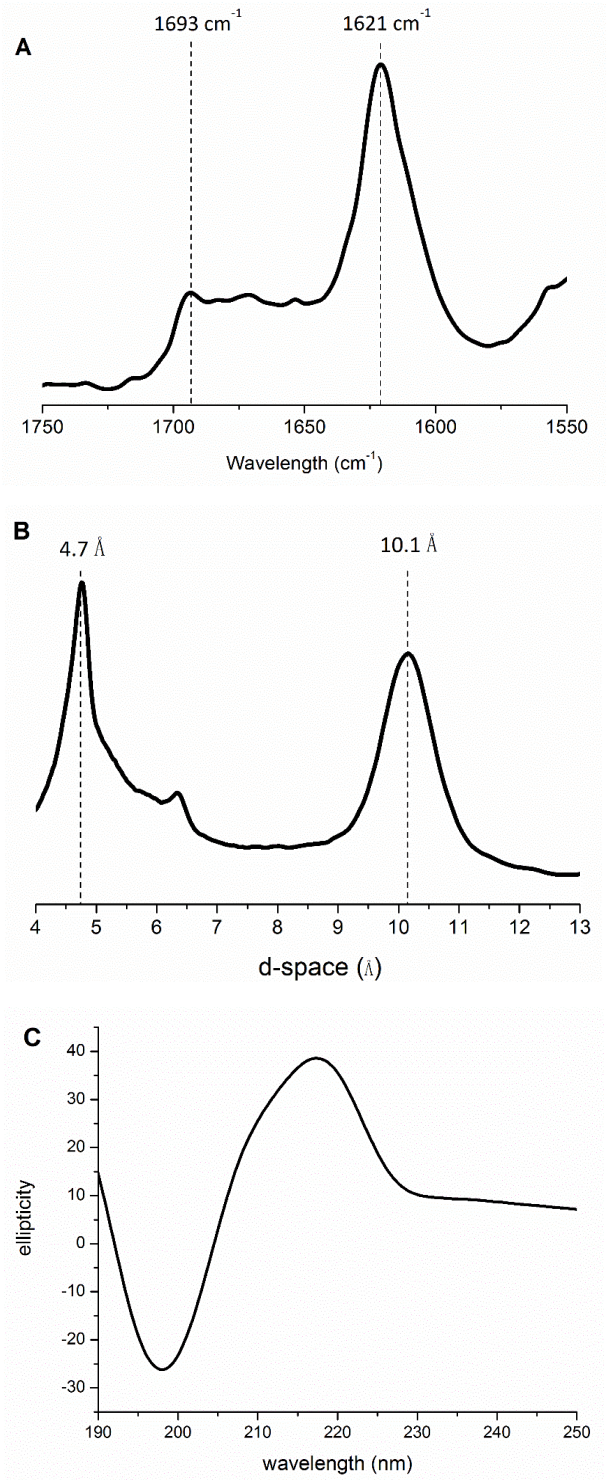


Fig 5.5 $(TTF)_3$ nanotube conformation characterization by FT-IR, X-ray powder diffraction and Circular Dichroism. (**A**) FT-IR shows Amide I band at 1621 cm^{-1} indicative of β-sheet and a transition at 1693 cm^{-1} suggesting anti-parallel orientation. (**B**) X-ray powder diffraction gives reflection at d-spacing of 4.7 Å and 10.1 Å, correspondent to Hydrogen bonding distance in β-sheet and lamination distance. (**C**) Circular Dichroism of $(TTF)_3$ nanotubes shows a negative maximum at 197 nm and a strong positive maximum at 217 nm.

Positively charged AuNPs and negatively charged AuNPs were added into solution containing mature $(TTF)_3$ nanotubes to study the surface charge of $(TTF)_3$ nanotubes. The observation that negatively charged AuNPs bind to nanotube surfaces (**Fig 5.6A**), and positively charged AuNPs do not bind (**Fig 5.6B**), suggests the surface of $(TTF)_3$ nanotubes is positively charged, confirms the N-terminus of the peptide exposed on the tube surface, as is seen in the E22L tubes[8].

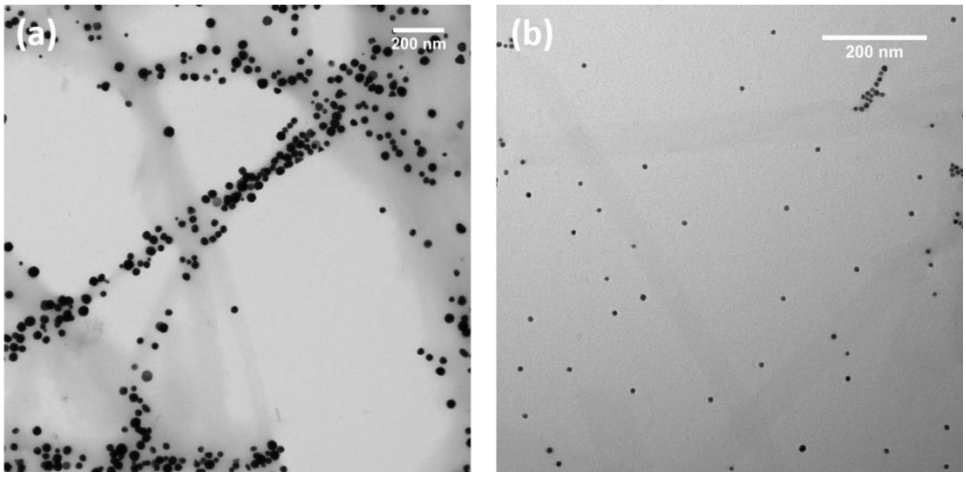


Fig 5.6 AuNPs binding experiments visualized by TEM without staining suggest $(TTF)_3$ nanotube surface is positively charged, with N-terminus exposed on surface. (**A**) Negatively charged AuNPs bind to the surface of nanotubes. (**B**) Positively charged AuNPs have no binding on nanotubes.

Taken together, these data suggests $(TTF)_3$ nanotube has cross-β peptide assembly structure and align in anti-parallel fashion, as shown in **Fig 5.7**. The thinner tube wall is more consistent with peptide monolayer instead of bilayer we observed in E22L.

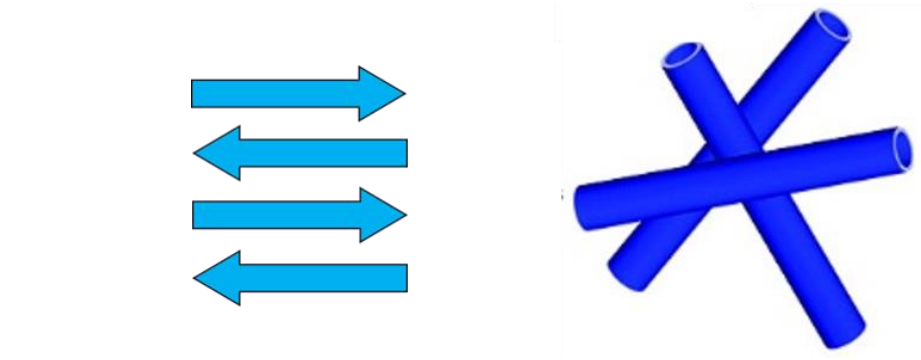


Fig 5.7 $(TTF)_3$ peptide assemble into nanotubes with anti-parallel cross-β structure.

5.2.2 Selecting Linear Trimer TTFoxTTFTTF from TTF-CHO + $(TTF)_2$ Network

Introducing $(TTF)_2$ peptide into TTF-CHO network can generate TTF*ox*TTFTTF, a linear $(TTF)_3$ peptide analog with one N, O-acetal linkage replacing amide bond. This product is expected to self-assemble into anti-parallel cross-β nanotubes, similar to its native peptide counterpart. Other species such as cyclic dimer are also expected in the network, and the proposed network diversity is shown in **Fig 5.8**.

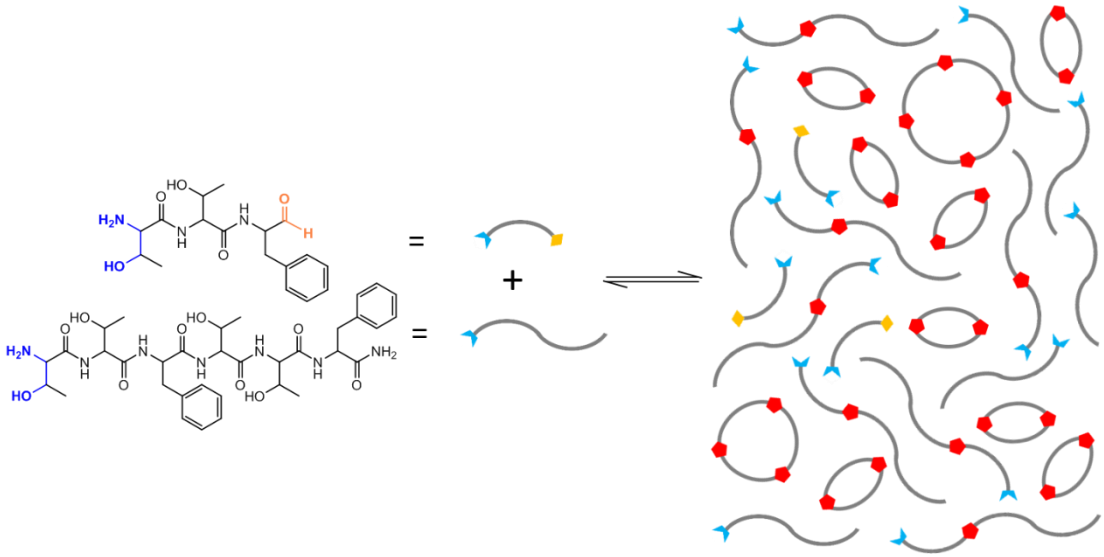


Fig 5.8 Illustration of TTF-CHO + $(TTF)_2$ dynamic chemical network.

To test the proposed assembly, $(TTF)_2$ peptide is added into TTF-CHO network at a 1:1 ratio, 2 mM each in 40% acetonitrile.at pH = 8 and incubated at 4°C. As HPLC chromatogram in **Fig 5.9** shows, two major new peaks corresponding to new network members appeared after 3 days. Mass spectrometry identifies the first species (retention time 28 min) as linear trimer product TTF*ox*TTFTTF (**Fig 5.10**), and the second (retention time 32 min) as the cyclic dimer. Detailed mass spectrometry identification of each component is summarized in **Table 5.1**.

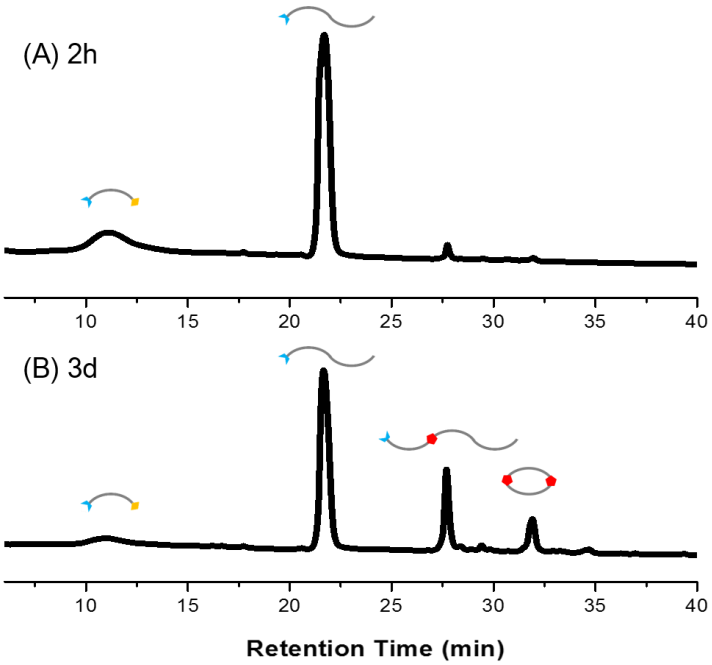


Fig 5.9 HPLC chromatograms of TTF-CHO + $(TTF)_2$ network exhibit emergence of TTF*ox*TTFTTF and cyclic dimer. Retention time for monomer: 11 in; $(TTF)_2$ peptide: 22 min; TTF*ox*TTFTTF: 28 min; cyclic dimer: 32 min.

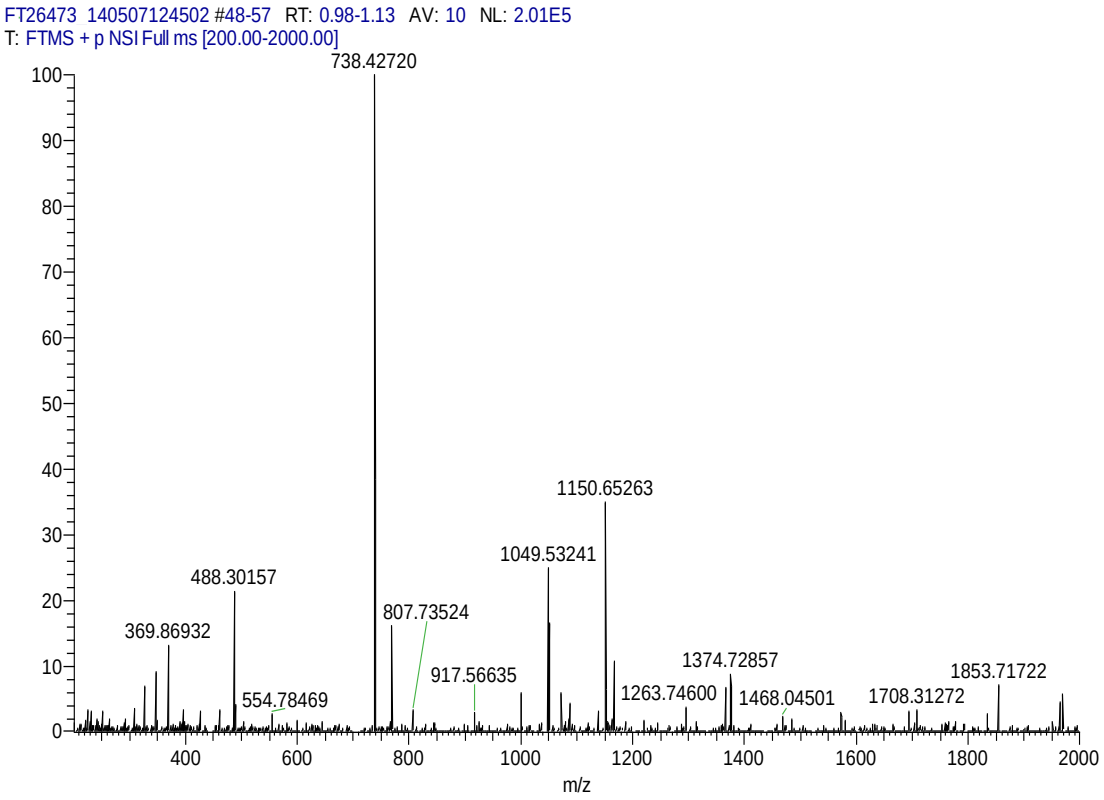


Fig 5.10 High resolution mass spectrometry confirms formation of linear trimer product TTF*ox*TTFTTF: $[M+H]^+$: 1049.53241; $[M+Et_3N+H]^+$ 1150.65263

Table 5.1 Summary of Mass Spectrometry identification of network species

species	**m/z found**	**m/z calculated**
TTF-CHO	352.1866	352.1872 ($M+H^+$: $C_{17}H_{26}N_3O_5$)
$(TTF)_2$	817.4837	817.4824 ($M+Et_3N+H^+$: $C_{40}H_{65}N_8O_{10}$)
TTF*ox*TTFTTF	1049.5330	1049.5308 ($M+H^+$: $C_{51}H_{73}N_{10}O_{14}$)
Cyclic dimer	667.3446	667.3455 ($M+H^+$: $C_{34}H_{47}N_6O_8$)

5.2.3 Self-Templated Propagation in TTF-CHO + $(TTF)_2$ Network

In order to identify the supramolecular assembly of the network, the whole network was sedimented at 16,000 × g for 30 min. The pellet was re-suspended and analyzed by HPLC. Enrichment of TTF*ox*TTFTTF in self-assembly, as shown in **Fig 5.11** (**b**), and spontaneously decrease of this species in supernatant (**Fig 5.11** (**c**)) suggests that TTF*ox*TTFTTF is a dominant component in the self-assembly. The residual $(TTF)_2$ peptide and cyclic dimer could due to the difficulty of completely removing residual solution in the gel-like assemblies or the disassembly of TTF*ox*TTFTTF during HPLC analysis.

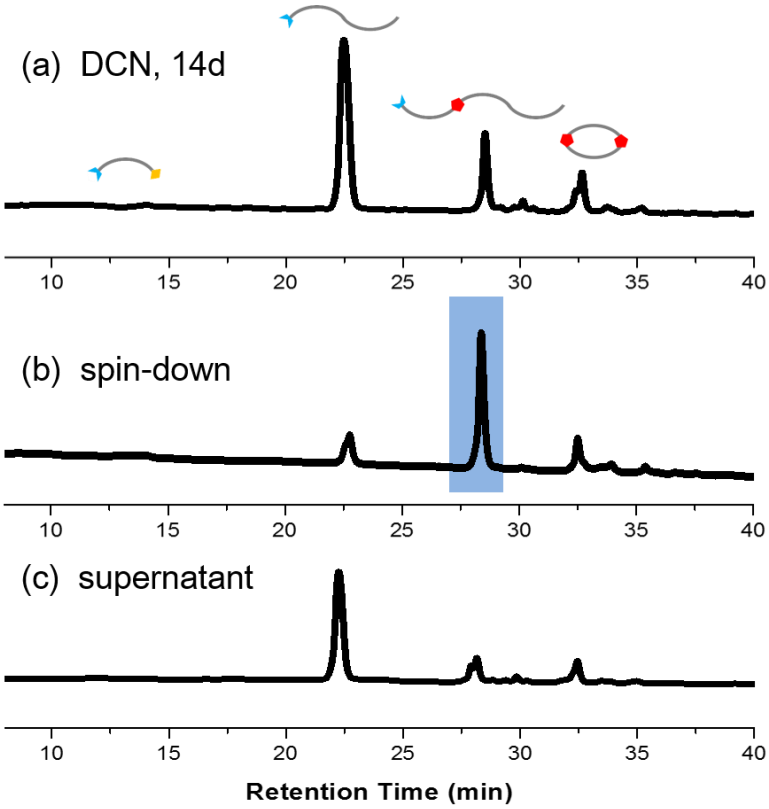


Fig 5.11 HPLC identifies TTFoxTTFTTF as dominant self-assembly species in DCN. (**A**) HPLC chromatogram of full network at 14d, (**B**) HPLC chromatogram of reinjection of self-assembly species. HPLC chromatogram shows a single peak correspondent to linear product TTF*ox*TTFTTF. Enrichment of self-assembly species is achieved by sedimentation at 16,000 × g for 30 min, followed by resuspension with 40% acetonitrile in water. (**C**) HPLC chromatogram of supernatant after sedimentation shows attenuation of linear trimer.

The morphology of TTF*ox*TTFTTF self-assembly in the network is different from the $(TTF)_3$ nanotubes (**Fig 5.12**). TEM provides evidence for particles formation after 24 hours, with diameter of 29.0 ± 3.0 nm (**Fig 5.12A & Fig 5.13A**), and after 3 days, particles began to transform into twisted fibers (**Fig 5.12B**). After 7 days, the particles disappeared and TTF*ox*TTFTTF assemblies matured into homogenous fibers with widths of 12.03 ± 1.90 nm, as shown in **Fig 5.12 C & D** and **Fig 5.13B**.

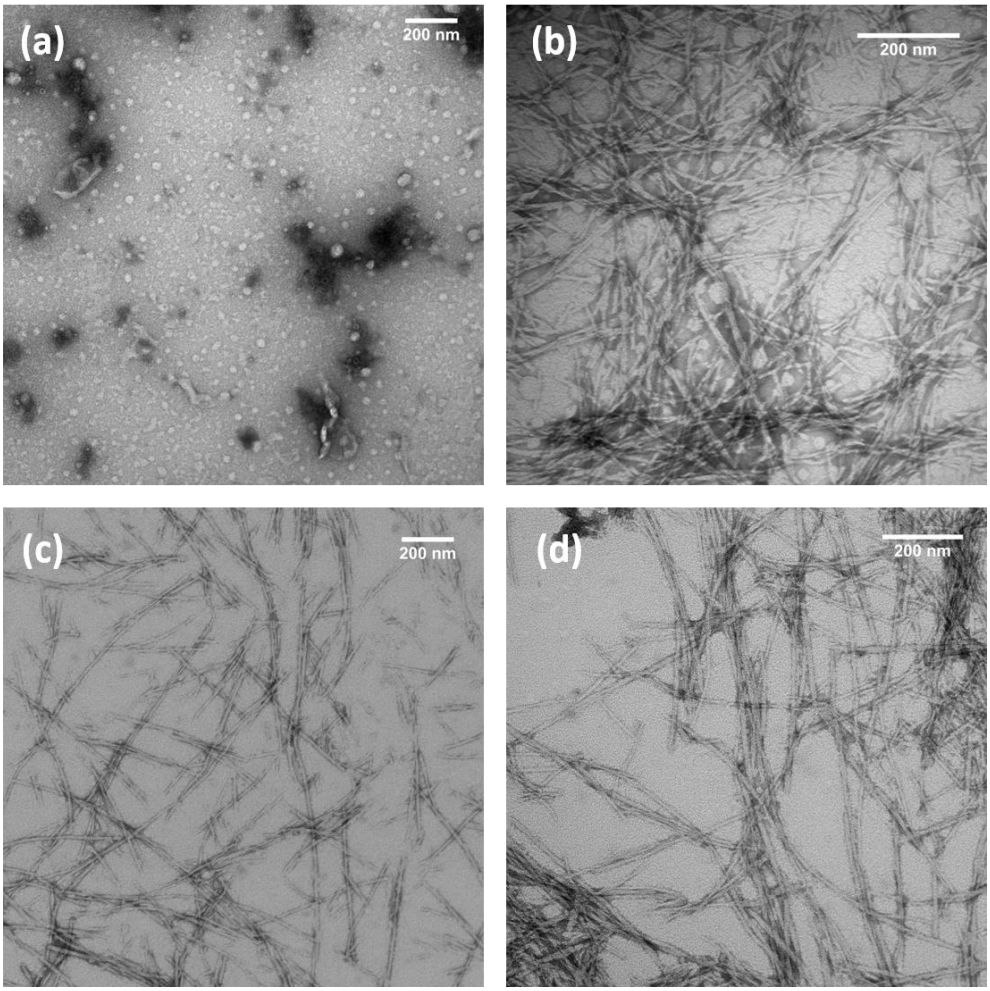


Fig 5.12 TEM images of self-assembly in TTF-CHO + $(TTF)_2$ network show phase transition in self-assembly. (**a**) Particles emerged at 24h; (**b**) after 3 days, particles starts to transform into fibers; (**c**) 7d and (**d**) 10d, particles disappeared and all transform into fibers

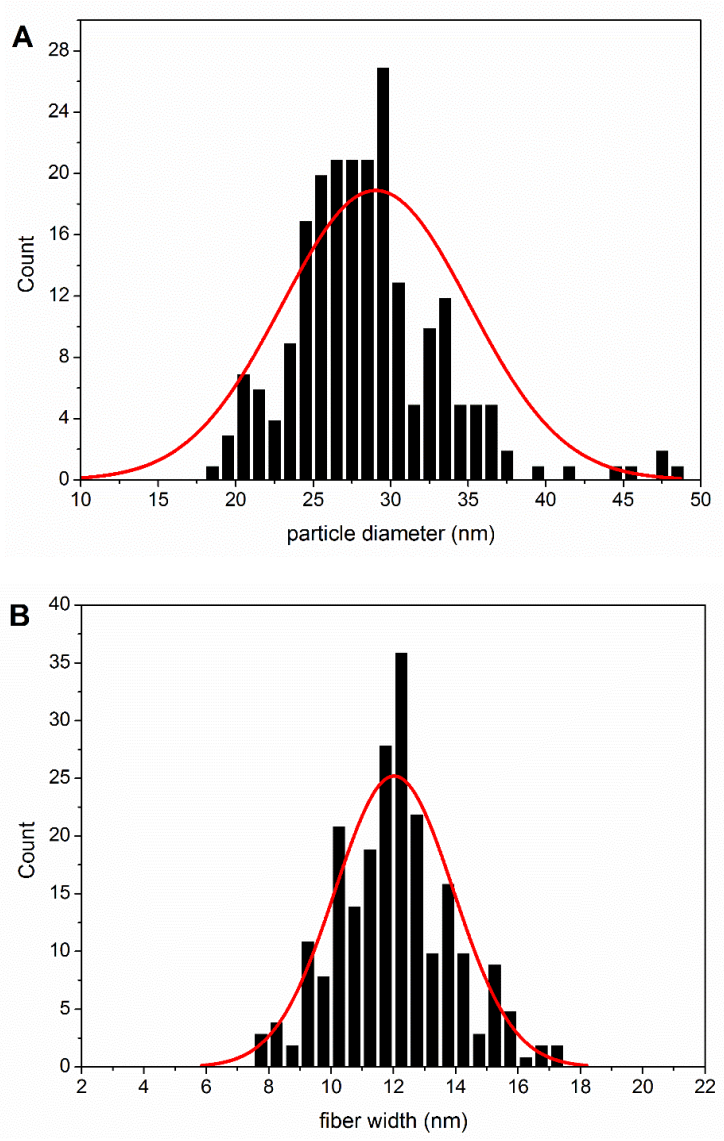


Fig 5.13 Histograms and Gaussian distribution of particle and fiber sizes. (**A**) Particles at 24h, with diameter of 29.0 ± 3.0 nm. (B) Mature fibers at 10 days, with width of 12.03 ± 1.90 nm

Structural characterization of self-assembly fibers is shown in **Fig 5.14**. As shown in **Fig 5.14A**, on FT-IR, the amide I transition is shifted to 1645 cm^{-1}, a phenomenon previously observed in peptide cross-β assembly with single site backbone disruption[9]. The X-ray powder diffraction contains reflections at 4.7 Å and 10 Å, typically of cross-β assembly. But the lamination reflection is much broader and weaker than $(TTF)_3$ nanotubes and more similar to the XRD pattern of fibers (**Fig 5.14B**) and that seem in

the E22V nanotubes, both containing smaller number of lamination repeats. CD spectrum again is significantly different from the typical β-sheet, with a positive maximum at 220 nm very similar to $(TTF)_3$ control peptides (**Fig 5.14C**).

Taken together, these data suggests that cross-β structures in TTF*ox*TTFTTF fibers are slightly disrupted, but largely reflect similar β-sheet secondary structure with native $(TTF)_3$ peptide even though they differ greatly in overall morphology.

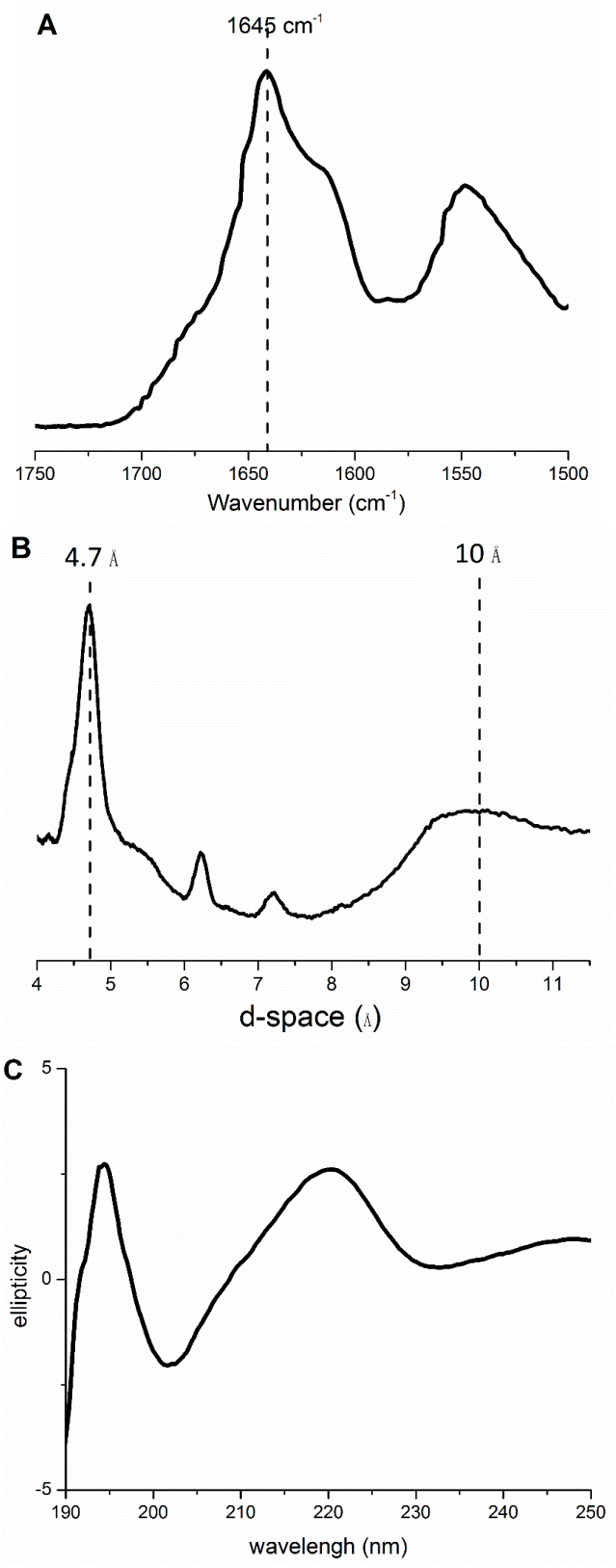


Fig 5.14 Conformation characterization of linear trimer fibers using FT-IR, XRD and CD. (**A**) FT-IR shows Amide I transition at 1645 cm^{-1} (**B**) XRD gives sharp 4.7 Å peak correspondent to hydrogen bonding, and weak, broad 10 Å peak correspondent to lamination distance. (**C**) CD gives positive maximums at 195 nm and 220 nm and negative maximum at 202 nm

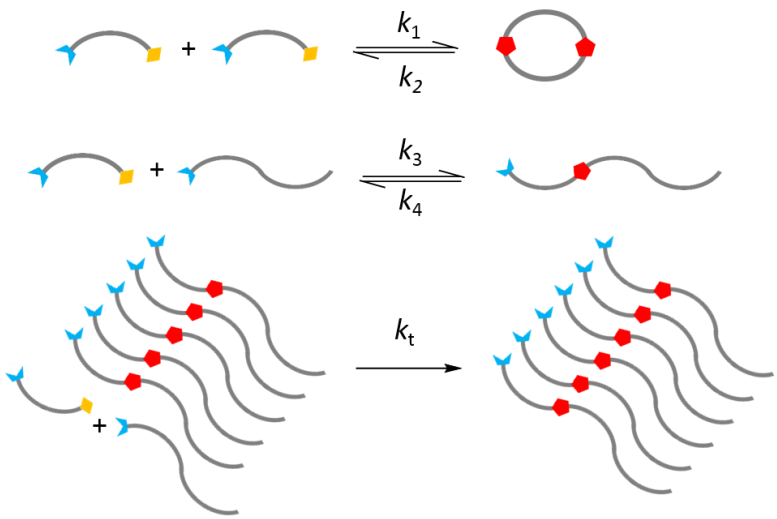
5.2.4 Kinetics of TTF-CHO + $(TTF)_2$ DCN

The self-templated replication in the TTF-CHO + $(TTF)_2$ network contains both chemical transition (reversible covalent bond formation) and physical transition (template-directed conformational propagation). Combining experimental data and mathematical modeling, we investigated the kinetics of templated replication in this network. The maturation of the TTF-CHO + $(TTF)_2$ network was followed by HPLC, and each network member was quantified by integration on the chromatogram components at 258 nm.

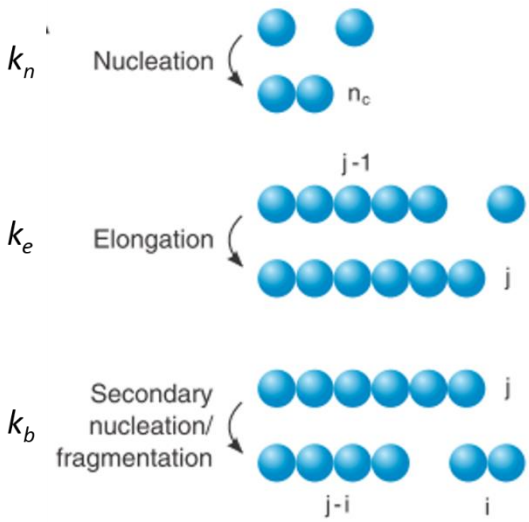
The kinetic model includes both the chemical reactions and the physical transitions. As shown in **Scheme 5.1**, the chemical process includes reversible N,O-acetal condensation between TTF-CHO and $(TTF)_2$ peptide, as well as cyclization of TTF-CHO to form cyclic dimers**.** The rate constants k_1, k_2, k_3 and k_4 each represent the forward and back reaction rate constants, together with an extra templated reaction at the fiber ends with rate constant k_t to account for the templated self-replication process[10].

The physical transition describes TTF*ox*TTFTTF templated assembly into fibers, as shown in **Scheme 5.2**. The linear trimer species TTF*ox*TTFTTF self-assembly is modeled with published approach[11], and contains three physical transition events: k_n is the rate constant for nucleation, describes the formation of a stable propagating nucleus; k_e is the rate constant for filament elongation, representing fiber growth; and

k_b is the filament breakage rate constant, the process generating new template surfaces for propagation.



Scheme 5.1 The chemical transition in kinetics modeling, where k_1 and k_2 as rate constants for formation and hydrolysis of cyclic dimer, k_3 and k_4 as rate constants for formation and hydrolysis of TTF*ox*TTFTTF linear trimer, and k_t as rate constant for template-directed TTF*ox*TTFTTF formation.



Scheme 5.2 Physical transition in kinetics modeling. k_n captures the nucleation process, k_e represents the fiber elongation, and k_b represents the breakage of fibers.

Parameter optimization was carried out with the *fmincon* function in MATLAB 2012a (The Mathworks, Inc) by minimizing the sum of square error (SSE) between the calculated and experimental values[12]. The best fit and the corresponding parameter values are summarized in **Figure 5.15** and **Table 5.2**, respectively. Given 8 parameters, this model nicely captures the kinetics and describes the interactions between species. As shown in the **Fig 5.15**, the DCN reaches equilibrium after 250 hours, consistent with all reactions being reversible, as specified in our design. This model also captures the participation of the templating reaction as reflected in k_t. Removal of this templating rate from the model results in poor data fitting. At equilibrium, the remaining TTF-CHO and $(TTF)_2$ suggests that the assemblies are less stable under this condition and stay equilibrium with other species. This differs from Otto's system[1] and N,N-acetal linked peptide DCNs from our group[2], in which self-assemblies eventually dominated and consumed the rest of species.

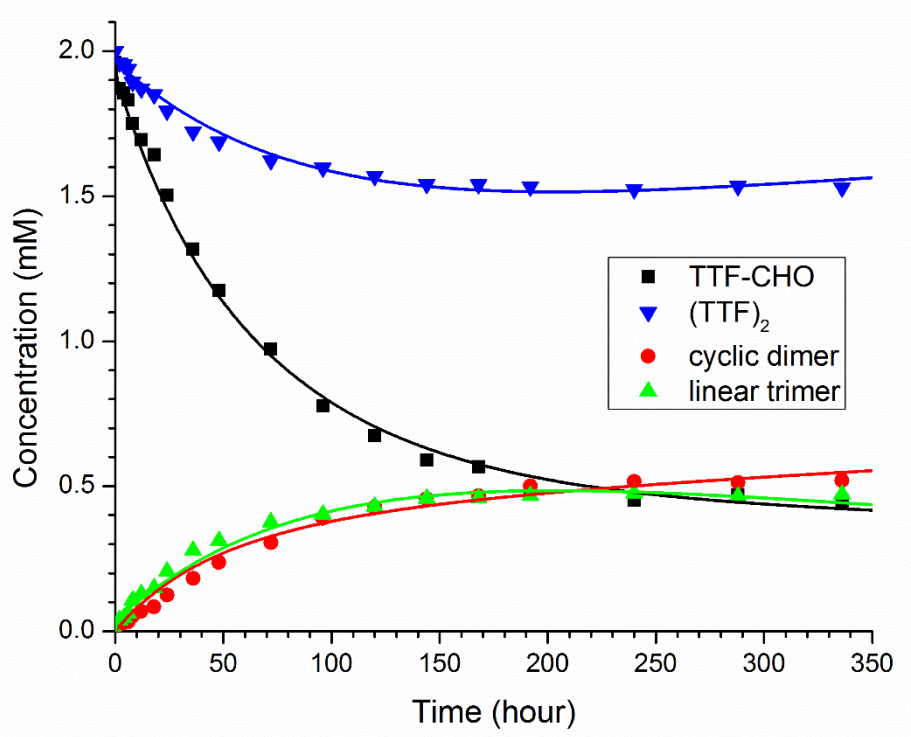


Fig 5.15. Experimental kinetics data and mathematical fitting of the TTF-CHO + $(TTF)_2$ peptide network.

Table 5.2 Rate and equilibrium constant**s** from best fit to TTF-CHO + $(TTF)_2$ network

parameter	**Value**
k_1 ($mM^{-1}sec^{-1}$)	6.77×10^{-7}
k_2 (sec^{-1})	5.71×10^{-10}
k_3 ($mM^{-1}sec^{-1}$)	7.63×10^{-6}
k_4 (sec^{-1})	6.08×10^{-4}
k_t ($mM^{-1}sec^{-1}$)	2.82×10^{-6}
k_n ($mM^{-1}sec^{-1}$)	4.67×10^{-4}
k_e ($mM^{-1}sec^{-1}$)	1.70×10^{-8}
k_b (sec^{-1})	2.54×10^{-6}

5.2.5 External Template Guides DCN to New Self-Assembly Structure

The DCN we explored above exhibits template-directed conformational replication, and selects for a propagating species that assembles into fibers with parallel cross-β structure. Peptide assembly structures other than fibers can serve as external template and provide different surfaces for conformational replication. Here we explored the template diversity, and plasticity of self-assembly species in terms of adopting different conformation by external templating effect.

$(TTF)_3$ peptide nanotube was chosen as logical next choice for external template, due to its identical sequence with self-assembly species in the DCN but one mutation site on backbone. In order to provide more templating surface to exert sufficient impact on the behavior of DCN, $(TTF)_3$ nanotubes were treated in bath sonicator for 1 hour to generate fragmented nanotube seeds as template, as shown in **Fig 5.18**. These nanotube seeds were added as templates into the TTF-CHO + $(TTF)_2$ network in the ratio of template : TTF-CHO : $(TTF)_2$ peptide = 0.4 : 2 : 2 mM. The templated network was incubated in 40% acetonitrile in water at pH = 8 and 4^{o}C. The templated network was monitored by HPLC over the course of 2 weeks and as shown in **Fig 5.16**, after 3 days, the linear trimer species TTF*ox*TTFTTF (retention time overlap with $(TTF)_3$ peptide) and cyclic dimer emerged. The time-dependence of DCN species distribution is plotted in **Fig 5.17**, but unfortunately the network is viscous and difficult to sample as a homogenous solution. Nevertheless, the distribution of species upon equilibrium is not significantly different from the DCN without the $(TTF)_3$ templates.

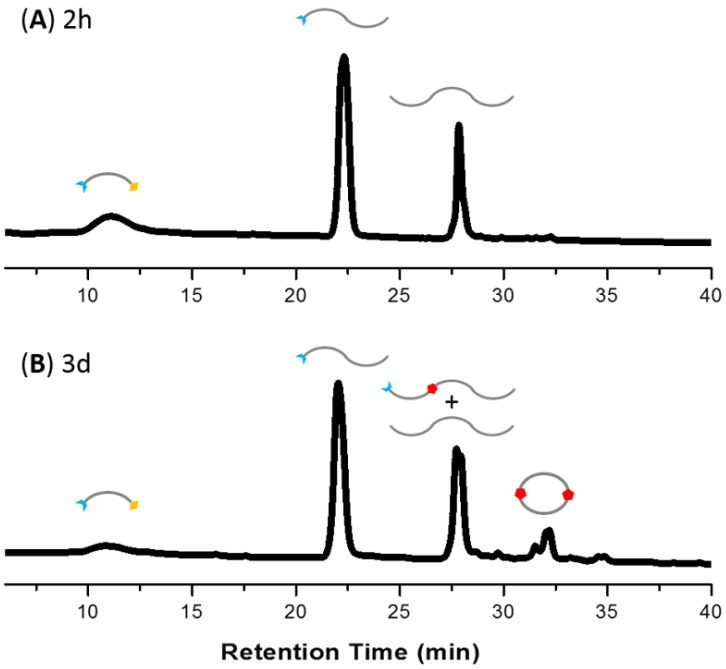


Fig 5.16 HPLC chromatogram of templated TTF-CHO + $(TTF)_2$ network shows emergence of . TTF*ox*TTFTTF and cyclic dimer, similar to self-templated network. (**A**) shows 3 peaks correspondent to 3 initial component in DCN: TTF-CHO monomer, $(TTF)_2$ peptide and $(TTF)_3$ peptide. (**B**) At day 3, both cyclic dimer and linear trimer emerged within network. The linear trimer TTF*ox*TTFTTF has similar retention time with $(TTF)_3$ peptide, and two peaks overlaps on chromatogram.

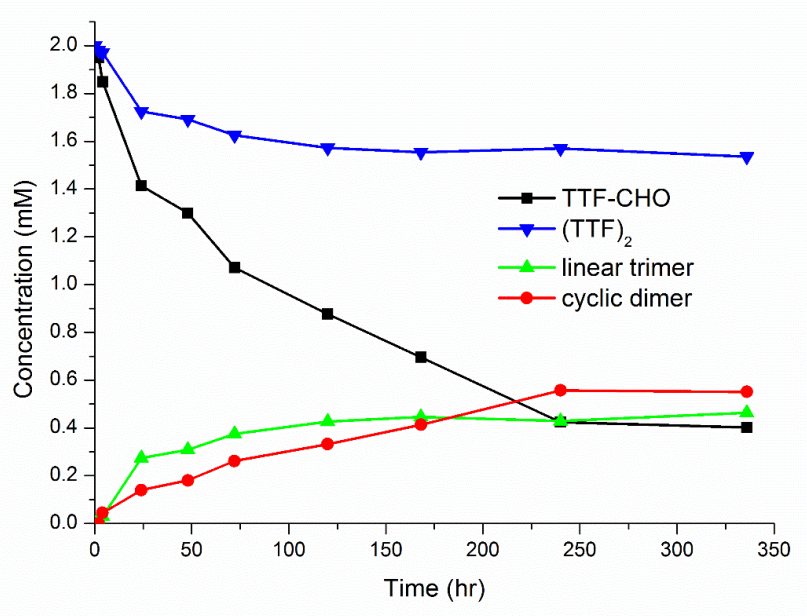


Fig 5.17 Kinetics of $(TTF)_3$ templated TTF-CHO + $(TTF)_2$ network shows network equilibrium is not significantly different from self-templated network. Due to the sampling difficulty associated with viscosity of the network, the kinetic data is less informative.

With the external template $(TTF)_3$, new assembly structure was formed. As shown in **Fig 5.18**, particles emerged after 24h, similar to non-templated network, but instead of transition into fibers, ribbons and nanotubes were observed. These particles and nanotubes co-existed for a period of time, but after 10 days, only long nanotubes similar to $(TTF)_3$ peptides nanotubes were observed.

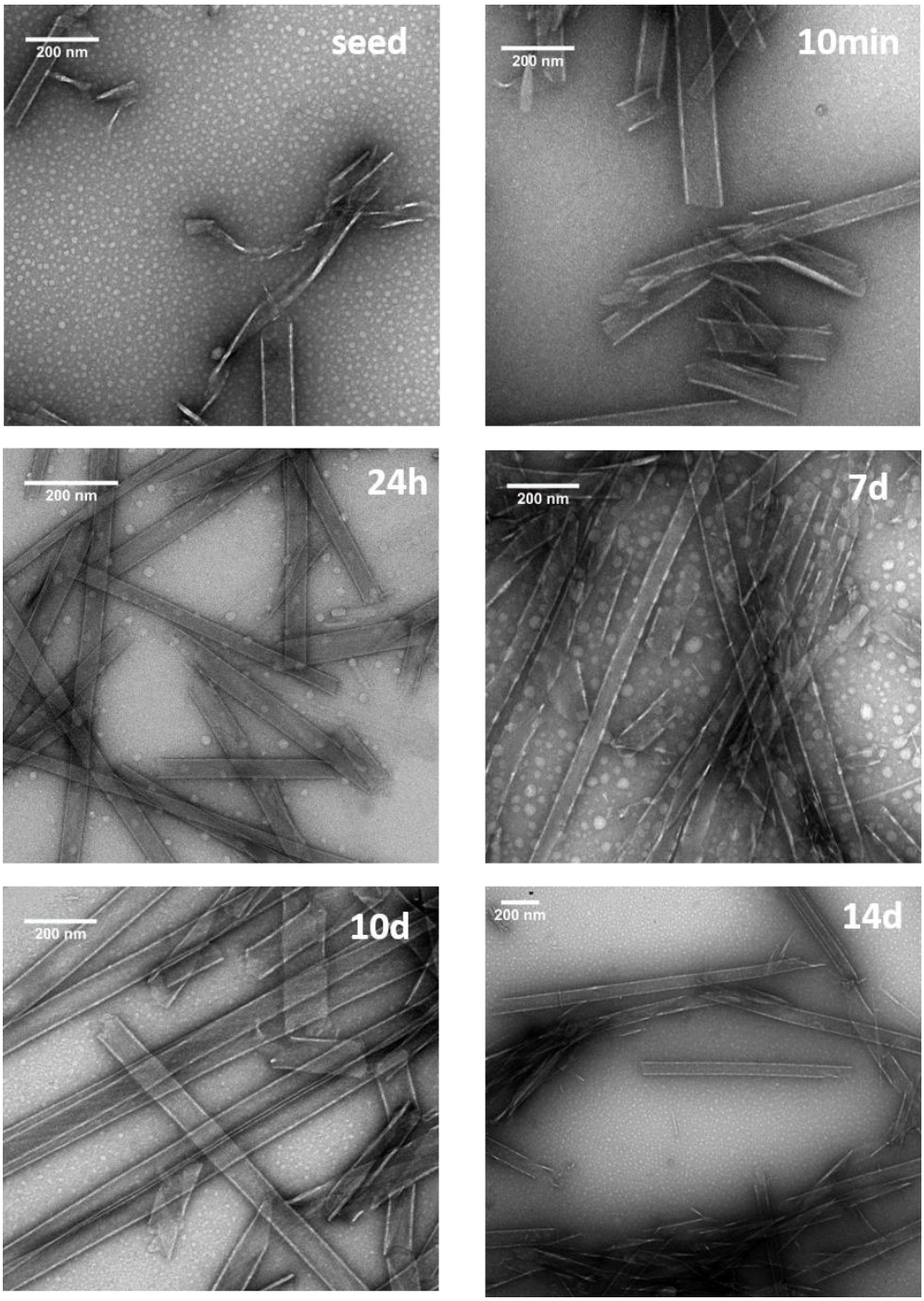


Fig 5.18 TEM images of $(TTF)_3$ nanotube templated DCN suggests new assembly structure formation. Templates were generated by sonicating $(TTF)_3$ nanotubes into fragments. After addition of templates into DCN, particles were detected after 24 h. Elongation of nanotubes was concomitant with gradual disappearance of particles, no fibers can be observed. After 10 d, only long nanotubes can be detected.

The assemblies were enriched by sedimentation at 16,000 × g for 30 min and resuspended with 40% acetonitrile in water for HPLC analysis (**Fig 5.19**). The peak corresponding to self-assembly species contains both $(TTF)_3$ template peptide and TTF*ox*TTFTTF. The co-existence of $(TTF)_3$ and TTF*ox*TTFTTF peptides within assembly is further supported by mass spectrometry, shown in **Fig 5.20** and **Table 5.3**, but do not resolve how the two structures are mixed in the nanotubes.

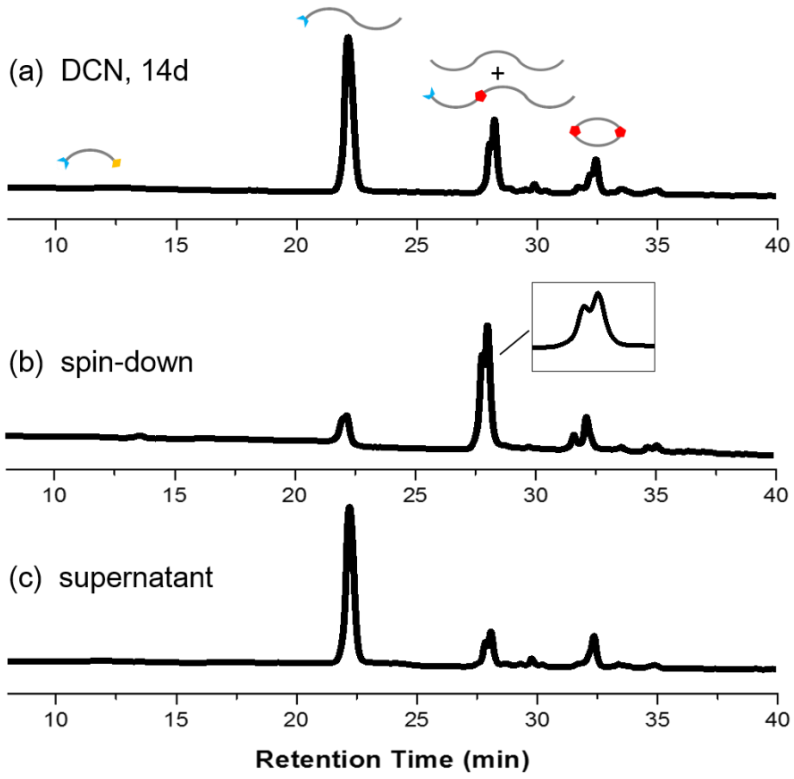


Fig 5.19 HPLC identification of self-assembly species in templated DCN suggests co-assembly of $(TTF)_3$ peptide and TTF*ox*TTFTTF. (**A**) HPLC chromatogram of full network at 14d, (**B**) HPLC chromatogram of reinjection of self-assembly species. HPLC trace shows overlap of two peaks (zoom-in), correspondent to TTF*ox*TTFTTF and $(TTF)_3$. (**C**) HPLC chromatogram of supernatant after sedimentation shows attenuation of linear trimer and $(TTF)_3$.

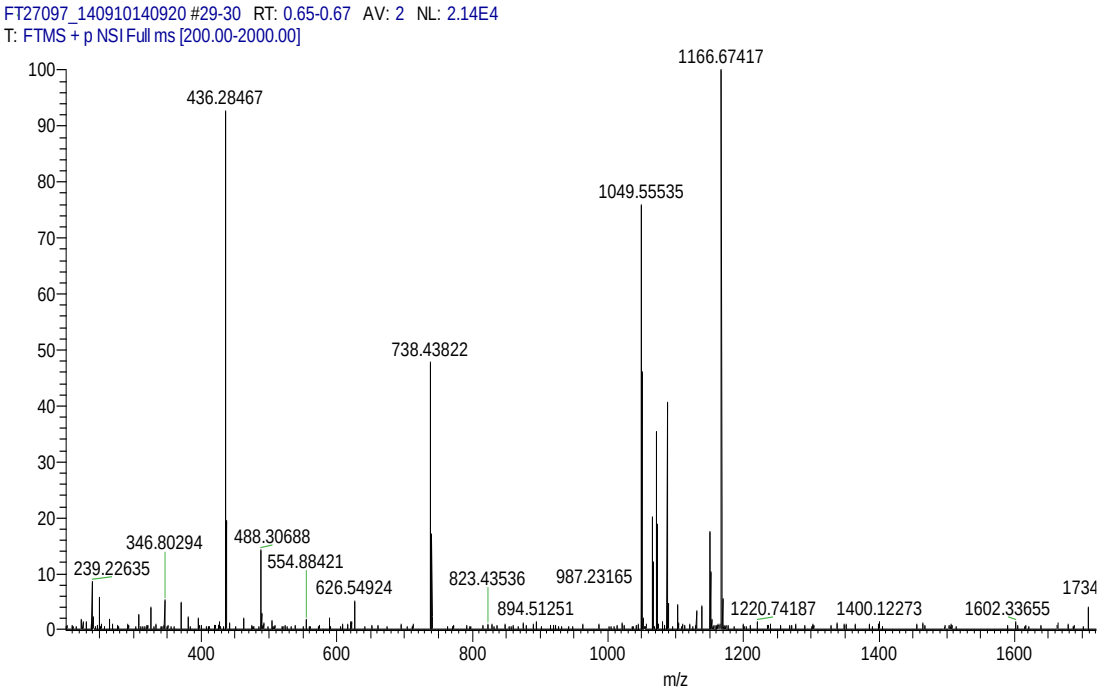


Fig 5.20 ESI high resolution mass spectrometry of enriched self-assembly species supports co-existence of (TTF)3 peptide and TTF*ox*TTFTTF in assembly. TTF*ox*TTFTTF: 1049.5330 ($M+H^+$), $(TTF)_3$: 1166.6481 ($M+Et_3N+H^+$)

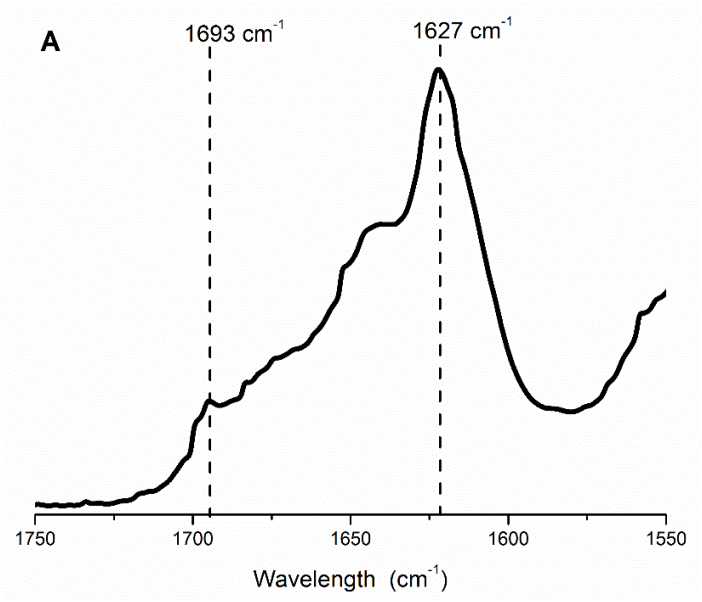
Table 5.3 Mass Spectrometry identification of assembly species in tube-templated DCN

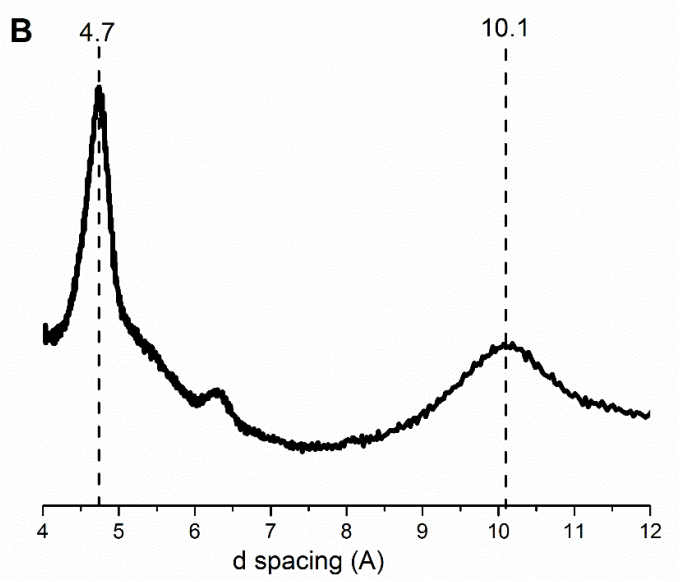
species	**m/z found**	**m/z calculated**
TTF*ox*TTFTTF	1049.5330	1049.5308 ($M+H^+$: $C_{51}H_{73}N_{10}O_{14}$)
$(TTF)_3$ peptide	1166.6481	1166.6461 ($M+Et_3N+H^+$: $C_{57}H_{88}N_{11}O_{15}$)

5.2.6 Structural Characterization of Nanotubes in External Templated DCN.

As shown in **Fig 5.21**, the amide I band in the FT-IR spectrum occurs at 1621 cm^{-1}, and a 1692cm^{-1} overtone assigned as anti-parallel alignment is also present(**Fig 5.21A**). X-ray powder diffraction provides evidence for the cross-β peptide alignment with d-spacing of 4.7 Å as well as the lamination distance at 10.1 Å (**Fig 5.21B**). The 10.1 Å

reflection is sharper and stronger compared to TTF*ox*TTFTTF fibers, consistent with more repeating laminates in nanotubes rather than in fibers. Circular Dichroism contains a complex pattern, with a negative ellipticity maximum at 203 nm, a positive maximum at 221 nm, and another negative maximum at 229 nm (**Fig 5.21C**), which may reflect mixed assemblies that differ from either pure assembly.


A
1693 cm-1
1627 cm-1
1750
1700
1650
1600
1550
Wavelength (cm-1)


B
4.7
10.1
4
5
6
7
8
9
10
11
12
d spacing (A)

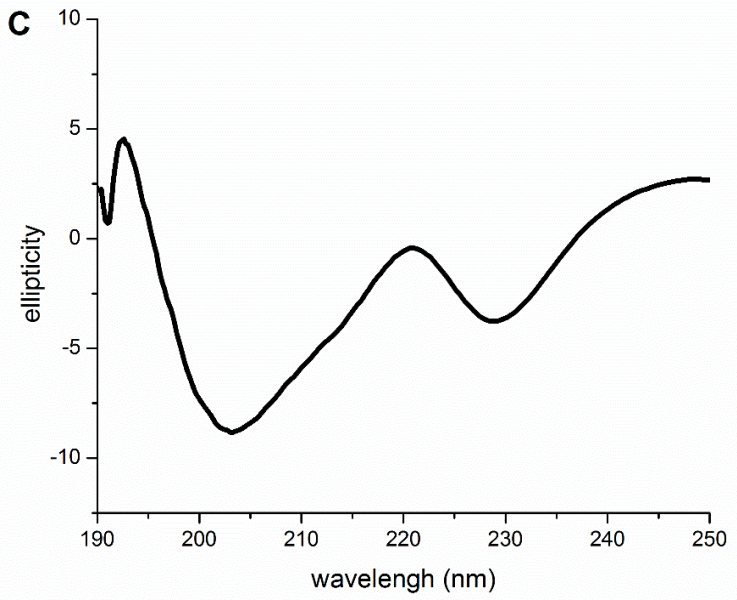

C
10
5
0
-5
-10
ellipticity
190
200
210
220
230
240
250
wavelengh (nm)

Fig 5.21 FT-IR, XRD and CD characterization of nanotubes in templated DCN suggests anti-parallel cross-β assembly. (**A**) FT-IR shows 1621 cm^{-1} amide I transition and 1692cm^{-1} overtone, suggesting anti-parallel β-sheet. (**B**) XRD pattern gives d-spacing of 4.7 Å for hydrogen-bonding distance and stronger 10.1 Å for lamination distance. (**C**) Circular Dichroism shows complicated spectrum likely due to mixed assembly, with negative maxima at 203 nm and 229 nm and positive maximum at 221 nm.

The data presented above demonstrates that in the presence of external template, linear trimer TTF*ox*TTFTTF is capable of being incorporated on the template surface and assemble into new structure as anti-parallel nanotubes, instead of separately grow as parallel fibers. This result suggests the diversity of network can be further enriched by introducing different templates.

5.3 Conclusion

In this chapter, we described a minimal dynamic self-replication peptide network and its response to templating by a pre-assembled peptide nanotubes. This TTF-CHO + $(TTF)_2$ peptide network is able to self-template for the selection of TTF*ox*TTFTTF fibers through phase transition (**Fig 5.12**), different from the all peptide $(TTF)_3$ control which assembles into nanotubes. Combined experimental analysis and mathematical modeling show that the kinetics of both chemical and physical transformation are involved in this self-templated DCN. All chemical reactions are reversible, an inherent property of dynamic chemical networks, and templated reaction is involved in directing formation of self-assemblies. The physical phase transition contains three events: nucleation, elongation and breakage of self-assembly fibers.

In the presence of $(TTF)_3$ nanotubes as an external template, the N,O-acetal linkage appears to be incorporated into self-assembly and adopt to morphology. The new self-assembly structure becomes nanotubes, with anti-parallel cross-β structure (**Fig 5.22**) instead of parallel fibers.

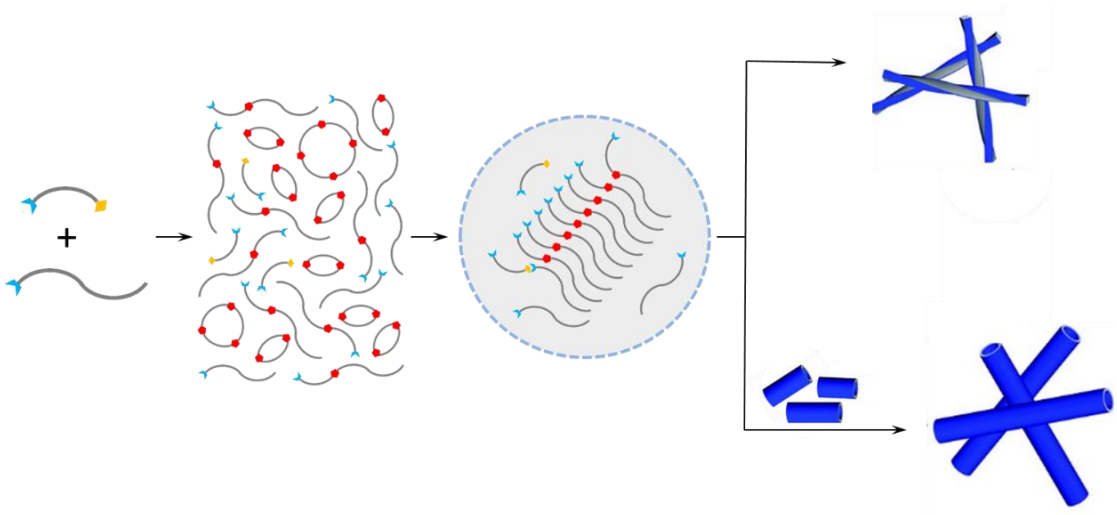


Fig 5.22 Self-templated and external templated replication in peptide DCN. Through dynamic covalent linkages, DCNs generate species diversity. Among all network members, TTF*ox*TTFTTF undergoes through phase transitions from solution to particles. With of self-templating (top) or external templating (bottom) effect, particles further transit into cross-β fibers or nanotubes.

5.4 Experimental

5.4.1 General Methods

Peptide Synthesis. Peptides H-TTFTTF-NH_2 and H-TTFTTFTTF-NH_2 were synthesized on a Liberty CEM Microwave Automated Peptide Synthesizer utilizing Fmoc-Rink Amide MBHA Resins purchased from AnaSpec. All Fmoc protected amino acids were purchased from Anaspec, and other chemicals from Sigma-Aldrich or Fisher

Scientific. Each peptide synthesis was performed at 0.1 mmol using a 45 mL reaction vessel at a scale of 0.1mmol. Fmoc-Rink Amide MBHA Resin was initially swollen using ~7 ml dimethylformamide (DMF) for 15 minutes. Fmoc deprotection was achieved by addition of 20% piperdine 0.1M N-Hydroxybenzotriazole (HOBt) in DMF with microwave power set to maintain temperature between 45-55℃ for 180 sec, followed by 3X flushing with DMF. Each coupling step is performed using 0.1 M Fmoc protected amino acid, and activated with 0.1 M 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate (HBTU) and 0.2 M N,N-Diisopropylethylamine (DIEA) in DMF. Coupling temperatures are maintained between 75-82 ℃ by optimizing microwave power for 300 sec. After coupling, the resin is rinsed with three aliquots of DMF. Peptides are cleaved from the resin using trifluoroacetic acid/thioanisole/1,2- ethanedithiol/anisole (90: 5 : 3 : 2, v/v/v/v) at room temperature for 3 hrs. The cleaved peptide- TFA solution is filtered, and precipitated by dropwise addition of cold (-20℃) diethyl ether. Precipitated product is centrifuged at 3000 rpm for 15 min, and the pellet is subjected to 3 additional rounds of washing with cold diethyl ether, followed by desiccating overnight. Dried peptides are dissolved in minimal volume of 40% acetonitrile / 60% water and purified by RP-HPLC (Water Delta 600) using a C18-reverse phase column with an acetonitrile-water (0.1% TFA) gradient. The molecular weight of each peptide is verified by mass spectrometry.

Transmission Electron Microscopy. All TEM are performed by using a Hitachi H-7500 transmission electron microscope to image the samples at 75kV. A TEM copper grid

with a 200 mesh carbon support (Electron Microscopy Sciences) was covered with 10 μL of DCN for 1 min before wicking the excess solution with filter paper. 10 μL of 2% uranyl acetate was added and incubated for 2 min, excess solution was wicked away, and the grids were placed in desiccators to dry under vacuum overnight.

Peptide Assembly. Peptide powders were dissolved in 40% acetonitrile/water, adjusted to pH 7 or 8 with Et_3N using a Fisher Scientific Accumet Basic AB15 pH meter. Incubate at 4 ℃ for 2-4 weeks till sample maturation.

Attenuated Total Reflectance Fourier Transform Infrared (FT-IR). FT-IR spectra were acquired using a Jasco FT-IR 4100 at room temperature and averaging 500 to 1000 scans with 2 cm^{-1} resolution, using an MCT detector with a 5 mm aperture and a scanning speed of 4 mm/sec. Aliquots (10μL) of DCN or peptide solution were dried as thin films on a Pike Galdi ATR diamond crystal.

X-ray powder diffraction (XRD). The samples were lyophilized to yield dry powder. The spectra were obtained by loading the powder into a 0.2 mm mylar capillary and the diffraction patterns acquired on a Bruker APEX-II diffractometer with graphite monochromated Cu radiation K-alpha radiation, λ=1.54184 Å, 40 kV and 35 mA, with a 0.5 pinhole collimator. Exposure times were typically 300s per frame. The data integration software XRD2DSCAN was used to convert the two dimensional data into a theta-2theta scan.

Atomic Force Microscopy (AFM). 20 µL of assembled peptide solutions were diluted to desired concentrations and placed on a silicon chip (4'' diameter diced silicon wafer (TedPella, Inc, Redding, CA, USA) previously cleaned by sonication in methanol for 20 min, for one minute at room temperature. Excess solution was removed with filter paper and the chip was rinsed with distilled H_2O. Tapping mode analysis on a JEOL JSPM- 4210 scanning probe microscope employed ultra-sharp non-contact silicon cantilevers (MikroMasch, Wilsonville, OR, USA) with typical frequencies between 240 and 350 kHz. Images were collected on dry samples

Negatively-charged gold nanoparticles were synthesized[13] by addition of freshly prepared 3 mM sodium citrate solution (500 µL) to 0.3 mM aqueous solution of $HAuCl_4$ (4.4 mL), and stirred for ~10 min. Freshly prepared, ice-cold, 0.3 M $NaBH_4$ (60 µL) was added while stirring. The solution immediately turned pink, indicating the formation of gold nanoparticles. The concentration ratio of [Au]/[capping agent]/[$NaBH_4$] was 1:1:12. The spherical-shaped nanoparticles of dimensions 4 ±1 nm showed a localized surface plasmon resonance (SPR) transition at 505 nm, as measured with a Jasco V- 530 UV spectrophotometer, and were further characterized by TEM. Positively-charge gold 10 nm cationic gold nanoparticles (Cytodiagnostics, Burlington, Ontario, Canada) were surface functionalized with (11-Mercaptoundecyl)-N,N,N-trimethylammonium bromide. For gold binding studies, 200 µL of preformed gold colloid (0.3 mM) was mixed with 5 µL mature peptide nanotubes (1.3 mM) to a final

9.2:1 ratio of gold to peptide. The mixture was incubated at room temperature for 3 hrs, until a purple red precipitate gradually formed. After centrifugation, the collected pellet was resuspended in the assembly solution. Then 10 μL of sample was applied to TEM grid for 2 min, and excess solvent was removed with filter paper. The sample grid was stored in vacuum desiccator overnight before imaging.

5.4.2 Dynamic Chemical Networks

TTF-CHO was generated in situ by deprotecting TTF-acetal with 4% HCl in H_2O then dried *in vacuuo* at 20^oC. TTF-CHO then was re-dissolved with $(TTF)_2$ peptide into 2 mM : 2 mM solution in 40% acetonitrile in H_2O. The pH of solution was adjusted by titrating aliquots of 5% Et_3N in H_2O using a Fisher Scientific Accumet Basic AB15 pH meter. The solution was incubated in 4^oC to generate DCN. For template seeding experiment, $(TTF)_3$ nanotube seeds was generated by sonicating $(TTF)_3$ nanotubes for 1hr then added into TTF-CHO + $(TTF)_2$ solution after pH adjustment. Solution pH was re-measured and readjusted if necessary. All other conditions are the same.

Kinetics study of DCN was performed by taking out small fraction of DCN solution for HPLC analysis over 2-4 weeks. All HPLC was performed on a Waters Delta 600 HPLC equipped with a photodiode array UV/Vis detector and a reverse-phase HPLC column (Kromasil 100-5C18, 4.6 x 250 mm). HPLC solvent gradient: acetonitrile: 10 mM TEAA in H_2O: 0-90 min: 10% to 100%, 90-100 min: 100%, 100-101 min: 100% to 10%, 101-110 min: 10%. UV absorbance at 258 nm was recorded and integrated for

kinetics analysis, assuming absorption coefficient for phenylalanine is additive for all species. Species identification was conducted by ESI mass spectrometry at Emory University Mass Spectrometry Center.

References

1. Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S., Mechanosensitive Self-Replication Driven by Self-Organization. *Science* **2010,** *327* (5972), 1502-1506.

2. Chen, C., Metastable Phases Direct Polymer Emergence and Evolution in Dynamic Chemical Networks. *Emory University Doctoral Thesis* **2015**.

3. Childers, W. S.; Mehta, A. K.; Ni, R.; Taylor, J. V.; Lynn, D. G., Peptides Organized as Bilayer Membranes. *Angewandte Chemie International Edition* **2010,** *49* (24), 4104-4107.

4. Childers, W. S., Amyloid: Merging the Properties of Membranes and Enzymes. *Emory University Doctoral Thesis* **2010**.

5. (a) Kong, J.; Yu, S., Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochimica et Biophysica Sinica* **2007,** *39* (8), 549-559; (b) Jackson, M.; Mantsch, H. H., The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Critical Reviews in Biochemistry and Molecular Biology* **1995,** *30* (2), 95-120.

6. Miyazawa, T., Perturbation Treatment of the Characteristic Vibrations of Polypeptide Chains in Various Configurations. *The Journal of Chemical Physics* **1960,** *32* (6), 1647-1652.

7. (a) EANES, E. D.; GLENNER, G. G., X-RAY DIFFRACTION STUDIES ON AMYLOID FILAMENTS. *Journal of Histochemistry & Cytochemistry* **1968,** *16* (11), 673-677; (b) Geddes, A. J.; Parker, K. D.; Atkins, E. D. T.; Beighton, E., "Cross-β" conformation in proteins. *Journal of Molecular Biology* **1968,** *32* (2), 343-358.

8. Mehta, A. K.; Lu, K.; Childers, W. S.; Liang, Y.; Dublin, S. N.; Dong, J.; Snyder, J. P.; Pingali, S. V.; Thiyagarajan, P.; Lynn, D. G., Facial Symmetry in Protein Self-Assembly. *Journal of the American Chemical Society* **2008,** *130* (30), 9829-9835.

9. Elgersma, R. C.; Meijneke, T.; de Jong, R.; Brouwer, A. J.; Posthuma, G.; Rijkers, D. T. S.; Liskamp, R. M. J., Synthesis and structural investigations of N-alkylated [small beta]-peptidosulfonamide-peptide hybrids of the amyloidogenic amylin(20-29) sequence: implications of supramolecular folding for the design of peptide-based bionanomaterials. *Organic & Biomolecular Chemistry* **2006,** *4* (19), 3587-3597.

10. Bourbo, V.; Matmor, M.; Shtelman, E.; Rubinov, B.; Ashkenasy, N.; Ashkenasy, G., Self-assembly and Self-replication of Short Amphiphilic β-sheet Peptides. *Origins of Life and Evolution of Biospheres* **2011,** *41* (6), 563-567.

11. Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M., An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* **2009,** *326* (5959), 1533-1537.

12. Miller, S. M.; Rawlings, J. B., Model identification and control strategies for batch cooling crystallizers. *AIChE Journal* **1994,** *40* (8), 1312-1327.

13. Maiti, S.; Dutta, S.; Das, P. K., Unmodified "GNP–Oligonucleotide" Nanobiohybrids: A Simple Route for Emission Enhancement of DNA Intercalators. *Chemistry – A European Journal* **2011,** *17* (27), 7538-7548.

Chapter Six

Summary and Outlook

6.1 Summary

Biological information transfer, either DNA sequence replication defined by the Central Dogma, or the prion infectious conformational replication, is relied on a template-directed replication mechanism. The templated replication is the molecular foundation for biology to reproduce, mutate and evolve. The significance of templated-replication in biological evolution, polymer chemistry and pathology attracts researchers from diverse backgrounds to understand the process and re-invent templated-replication in a simpler, non-enzymatic fashion. However, low fidelity and lack of effective ligation methods impede the efficient non-enzymatic replication to take place. The booming field of dynamic chemical networks (DCNs) provides a unique opportunity for us to address such problems and create fully dynamic replication systems responsive to different environments and templates. By taking advantage reversible N,O-acetal linkages, we re-invented both DNA templated sequence information transfer and peptide self-templated conformational replication through dynamic chemical networks.

We investigated dynamic feature of N,O-acetal linkages on both nucleoside and peptide backbones. As discussed in Chapter 2, environmental factors (pH, temperature, solvent) affect the dynamics of N,O-acetal condensation and hydrolysis. We also studied stereo-

chemistry of N,O-acetal on both nucleoside and peptide backbones using spectroscopic analysis.

With insight into the dynamic properties of the N,O-acetal bond, fully functionalized networks were constructed for selection of templated replication behaviors. In Chapter 3, we extended the DNA template chemistry to rANP networks. Two uridine-derived monomers were synthesized, one with 3'-aldehyde and 5'-amine (rU1), the other with 3'-amine and 5'-aldehyde (rU2), as building blocks of the networks. DNA templates selected for DNA/ANP duplex structure within the networks and the resulting products showed high thermal stability. Initial evidence suggested that duplex formation is driven by Watson-Crick base-pairing, with non-complementary poly(dT) template unable to drive formation of duplexes.

Chapter 4 explored DCNs based on functionalized tripeptide TTF-CHO. In this network, cyclic dimer is selected as major network species. With 2D-NMR and simulations, the stereochemistry of the two new N,O-acetals are assigned as (*trans*, *trans*). In 40% and 20% acetonitrile, no self-assembly occurs, however, in 5% acetonitrile, this cyclic dimer assembles into fibers with initial evidence suggesting to be β-sheet.

In Chapter 5, we achieved templated conformational replication of TTF*ox*TTFTTF in peptide DCNs. The linear trimer assembles into fibers, different from nanotube-like structures formed by control $(TTF)_3$ peptide. Structural analyses suggest cross-β

structure in fibers is slightly perturbed with the introduction of N,O-acetal. Kinetics study highlights reversibility of all chemical process, and involvement of templated ligation. This network is responsive towards external template effect, with $(TTF)_3$ template selecting for same self-assembly species but different morphology, establishing analog conformational replication.

Both rANP network and peptide network are capable of producing informational polymers in response to templates. The templated replication combines chemical and physical transitions to achieve emergence of high order supramolecular structures. With the remarkable power to select for sequence specificity and high order structure, it sets the stage for further exploration of emergent properties and evolvable systems.

6.2 Outlook

6.2.1 One Common Linkage for Two Biopolymers

In biology, the two informational biopolymers, nucleic acids and proteins, are two separated biopolymers with different linkages, but form a highly synergic network through cooperative functions and co-assemblies. For example, the ribosome itself is a multi-component supramolecular assembly containing RNA subunits and ribosomal proteins[1], and functions as a digital-to-analog converter to transfer information from nucleic acids to proteins. In prion infectious propagation, cross-β assembly surface can recruit nucleic acids binding, and the binding in return stimulate and facilitate prion growth[2].

In this thesis, using reversible N,O-acetal linkages as the common linkage to construct both nucleic acid and peptide dynamic chemical networks, we significantly simplify current biology. This common linkage in two informational biopolymers allows for minimal form of nucleic acid/amino acid symbiosis. Incubating both ANPs and peptide-aldehydes in one network enables interaction between two building blocks based on common N,O-acetal linkages along with non-covalent interactions, and potentially selecting for collaboration, co-assembly or chimeric structures, leading to minimal form biopolymer co-evolution.

6.2.2 Enriching Inventory Complexity of DCN Building Blocks

Templated duplex formation between homopolymer $(dA)_{16}$ and uridine derived rANP monomers, and peptide replication based on N,O-acetal linked $(TTF)_n$ sequences are minimal systems able to illustrate the essentials of templated replication in dynamic chemical networks. To further diversify the inventory and extend these networks, other building blocks can now be introduced into the networks.

To enrich the inventory of ANP networks, monomers with different nucleobases can be synthesized. More general methods for synthesis of ANP monomers bearing different nucleobases can be developed starting from sugars or other backbone molecules. Reported synthetic routes can be readily modified for the synthesis of ANP monomers, such as the synthesis of 3'-amino-5'-carboxymethyl nucleosides from D-glucose[3]

along with other methods[4]. Synthesis of nucleic acid analogs capable of forming other types of dynamic linkages are known[5], and with these general methods for synthesis, non-canonical nucleobases can be incorporated to further enrich the pool of DCN building blocks.

The sequence and backbone diversity of peptide networks can also be extended. For peptide replication in DCNs, 20 amino acids along with many non-natural amino acids, allows for exploration of a broad range of sequence diversity. In addition, as discussed in Chapter 2, diverse amino acid residues also provide broad chemical space for dynamic linkages. Using more than one building blocks to construct peptide DCNs enables generating peptide sequences with multiple different dynamic linkages, further increasing the diversity level of the networks.

6.2.3 Selection of Functions in Dynamic Chemical Networks

Template-directed replication in DCNs allows for selection of oligomers with high order structures such as duplexes or cross-β assemblies. Theses structures potentially dictate functions, such as catalytic activity, as emergent property in evolvable systems.

Catalytic nucleic acids are regarded as first functional biopolymers in prebiotic evolution, supported by catalytic activities reported in natural[6] and modified[7] nucleic acids. Our lab found out that functionalized ANP linkages is catalytic and can function as artificial DNA ligases[8]. Compared to natural nucleic acids, amine or acetal linkages

can be easily functionalized to achieve more versatile catalytic activity. Therefore, selection for ANPs with catalytic activity (ANPzymes) is critical step for ANP networks to achieve functional diversity beyond structural diversity, leading to emergent functions in such networks.

Similarly, in peptide networks, cross-β assembly structures exhibit potential for diverse catalytic activity that generate complexity and feedback in the networks. Cross-β assembly provides binding grooves for small molecules[9], and upon binding, catalysis can be achieved[10]. Introducing dynamic linkages into peptide self-assembly structure allows for structure diversity, as we have shown in Chapter 5, which potentially leads to on-and-off catalytic activity and feedbacks upon catalysis, as different structures lead to difference in catalytic activity.

References

1. Schmeing, T. M.; Voorhees, R. M.; Kelley, A. C.; Gao, Y.-G.; Murphy, F. V.; Weir, J. R.; Ramakrishnan, V., The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA. *Science* **2009,** *326* (5953), 688-694.

2. Deleault, N. R.; Lucassen, R. W.; Supattapone, S., RNA molecules stimulate prion protein conversion. *Nature* **2003,** *425* (6959), 717-720.

3. (a) Chakraborty, T. K.; Gajula, P. K.; Koley, D., Studies Directed toward the

Development of Amide-Linked RNA Mimics: Synthesis of the Monomeric Building Blocks. *The Journal of Organic Chemistry* **2008,** *73* (17), 6916-6919; (b) Sunkari, Y. K.; Pal, C.; Reddy, T. J.; Chakraborty, T. K., 3′-Amino-5′-carboxymethyl-3′,5′-dideoxy nucleosides for the synthesis of fully amide-linked RNA mimics. *Tetrahedron* **2014,** *70* (35), 5455-5462.

4. (a) Robins, M. J.; Doboszewski, B.; Timoshchuk, V. A.; Peterson, M. A., Glucose-Derived 3'-(Carboxymethyl)-3'-deoxyribonucleosides and 2',3'-Lactones as Synthetic Precursors for Amide-Linked Oligonucleotide Analogues1. *The Journal of Organic Chemistry* **2000,** *65* (10), 2939-2945; (b) Rozners, E.; Liu, Y., Toward Amide-Linked RNA Mimics: Total Synthesis of 3'-C Branched Uridine Azido Acid via an Ene−Iodolactonization Approach. *Organic Letters* **2003,** *5* (2), 181-184; (c) Rozners, E.; Liu, Y., Monomers for Preparation of Amide-Linked RNA: Asymmetric Synthesis of All Four Nucleoside 5‘-Azido 3‘-Carboxylic Acids. *The Journal of Organic Chemistry* **2005,** *70* (24), 9841-9848.

5. Hickman, D. T.; Sreenivasachary, N.; Lehn, J.-M., Synthesis of Components for the Generation of Constitutional Dynamic Analogues of Nucleic Acids. *Helvetica Chimica Acta* **2008,** *91* (1), 1-20.

6. Silverman, S. K., DNA as a Versatile Chemical Component for Catalysis, Encoding, and Stereocontrol. *Angewandte Chemie International Edition* **2010,** *49* (40), 7180-7201.

7. Taylor, A. I.; Pinheiro, V. B.; Smola, M. J.; Morgunov, A. S.; Peak-Chew, S.; Cozens, C.; Weeks, K. M.; Herdewijn, P.; Holliger, P., Catalysts from synthetic genetic polymers. *Nature* **2015,** *518* (7539), 427-430.

8. Ye, J.; Gat, Y.; Lynn, D. G., Catalyst for DNA Ligation: Towards a Two-Stage Replication Cycle. *Angewandte Chemie International Edition* **2000,** *39* (20), 3641-3643.

9. Childers, W. S.; Mehta, A. K.; Lu, K.; Lynn, D. G., Templating Molecular Arrays in Amyloid's Cross-β Grooves. *Journal of the American Chemical Society* **2009,** *131* (29), 10165-10172.

10. (a) Miao, X.; Cao, W.; Zheng, W.; Wang, J.; Zhang, X.; Gao, J.; Yang, C.; Kong, D.; Xu, H.; Wang, L.; Yang, Z., Switchable Catalytic Activity: Selenium-Containing Peptides with Redox-Controllable Self-Assembly Properties. *Angewandte Chemie International Edition* **2013,** *52* (30), 7781-7785; (b) Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stöhr, J.; Smith, T. A.; Hu, X.; DeGrado, W. F.; Korendovych, I. V., Short peptides self-assemble to produce catalytic amyloids. *Nat Chem* **2014,** *6* (4), 303-309; (c) Zhang, C.; Xue, X.; Luo, Q.; Li, Y.; Yang, K.; Zhuang, X.; Jiang, Y.; Zhang, J.; Liu, J.; Zou, G.; Liang, X.-J., Self-Assembled Peptide Nanofibers Designed as Biological Enzymes for Catalyzing Ester Hydrolysis. *ACS Nano* **2014,** *8* (11), 11715-11723; (d) Lee, K. S.; Parquette, J. R., A self-assembled nanotube for the direct aldol reaction in water. *Chemical Communications* **2015,** *51* (86), 15653-15656.