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:

_____

Chenrui Chen Date

METASTABLE PHASES DIRECT POLYMER EMERGENCE AND EVOLUTION IN DYNAMIC CHEMICAL NETWORKS

By
Chenrui Chen
Doctor of Philosophy
Chemistry

_____

Dr. David G. Lynn
Advisor

_____

Dr. Vincent Conticello
Committee Member

_____

Dr. Emily Weinert
Committee Member

Accepted:

_____

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

_____ Date

METASTABLE PHASES DIRECT POLYMER EMERGENCE AND EVOLUTION IN DYNAMIC CHEMICAL NETWORKS

By

Chenrui Chen

B. S., Hunan University, 2006

M.S., Hunan University, 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

2015

Abstract

Metastable Phases Direct Polymer Emergence and Evolution in Dynamic Chemical Networks

By Chenrui Chen

The *de novo* construction of peptide surrogate polymers linked with reversible acetals is demonstrated in novel dynamic chemical networks. By changing the oxidation state of the amino acid, we demonstrate environmentally responsive dynamic peptide network that allows for specific oligomer synthesis and prion-like selection. Moreover, these dynamic networks are susceptible to infection by protein templates that template morphologically distinct supramolecular materials. This process crosses the critical chemical threshold necessary for self-organizing alternative polypeptide scaffolds and abiotic biopolymer evolution and provides the first empirical model for this polymer threshold of life.

METASTABLE PHASES DIRECT POLYMER EMERGENCE AND EVOLUTION IN DYNAMIC CHEMICAL NETWORKS

By

Chenrui Chen

B. S., Hunan University, 2006

M.S., Hunan University, 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

2015

For My Family

Acknowledgements

I would like to express the deepest thankfulness to my advisor, Dr. David Lynn, for all the support and encouragement he has given me through these years. He is a great scientist and mentor. His passion and curiosity in science, kindness and patience in students and his dedication to work provide an ideal environment for me to learn and grow as a scientist. I can always knock on his office door to discuss research with him, even on weekends or at night. He challenges me for critical thinking, designing experiments to test hypothesis and defending myself confidently to questions. His enthusiasm in teaching and out-reach teach me the importance of conveying science to the public and inspiring others. His optimism and great sense of humor cheers me up when I am down. All these will continue to impact me in the future.

I appreciate the feedback and insightful suggestions given by my committee members, Dr. Vincent Conticello and Dr. Emily Weinert, as well as my former committee member, Dr. Justin Gallivan. I learned a lot from Dr. Conticello's Biophysical class and Dr. Stefan Lutz's Biochemistry class.

I am appreciative to Dr. Anil Mehta for being a knowledgable and patient mentor and a great resource in the lab. He teaches the complex concept of NMR technique in a very smart and easily- understood way. He is always there whenever I need help and give insightful suggestions. I would like to thank Dr. Seth Childers, Dr. Rong Ni and Dr. James Simmons. When I joined the

lab, they taught me lots of lab skills including synthesizing peptides and using various analytical instruments.

Great thanks to our collaborators Dr. Martha Grover and Ming-Chien Hsieh. They help in modeling of the kinetics of the data and provide valuable insight into the analysis of phase transition in the chemical networks. Thanks to Dr. Facundo Fernandez and Manshui Zhou collaboration in analysis of chemical networks.

I am very grateful to Dr. Chengwei Li for teaching me all those organic synthesis techniques during my early years in the lab and Dr. Jay Goodwin for all the brain-storming discussions and great suggestions on experiments. Thanks to Dr. Junjun Tan for exchanging ideas and learning from each other's experiments. Her persistence always impresses and encourages me. Many thanks to Savannah Johnson, Tolu Omosun, we have been through so many things together and are emotional support for each other. I was lucky to know Yue Liu, Yi-Han Lin and Phoebe Young and enjoyed the time with them in the lab and in life.

I have enjoyed doing experiments with Fish (Ting) Pan, Li Zhang, Chen Liang, Rolando Rengifo and benefit from their contributions. I am also very fortunate to have Lisa Li, Noel Li and Allisandra Mowles as great officemates, who make coming to lab a true pleasure. I look forward to seeing all of your new discoveries! I am glad to have Dr. Jillian Smith and Daniel Pierce as colleagues who provide helpful feedback to my research. Many thanks to my friends Wenting Wu, Chris Xu, John Wang, Xuesong Yang, Ye Yang, Nianhui Song, etc.

I am grateful to Dr. Shaoxiong Wu and Dr. Bing Wang for helping with NMR experiments. Thanks to Hong Yi and Jeanette Taylor for helping with my transmission electron microscopy

experiments. Thanks to Fred Strobel for mass analysis, John Bacsa for X-ray diffraction. Special thanks to Neil Anthony from Physics for teaching me about fluorescence microscopy.

Finally, I am very appreciative that I have a loving and caring family. I cannot imagine my life without the unconditional love and endless support from my parents. Great thanks to my husband Foley for all the love, encouragement and support each and every day.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

AFM	atomic force microscopy
CD	circular dichroism
CR	congo red
DCC	dynamic combinatorial chemistry
DCL	dynamic combinatorial library
DCN	dynamic combinatorial network
DCM	dichloromethane
DIPEA	N,N-diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
EDT	1,2-ethanedithiol
FMOC	9-fluorenylmethoxycarbonyl
FRET	fluorescence resonance energy transfer
FT-IR	Fourier Transform Infrared Spectroscopy
HBTU	2-(1H-Benzotriazol-1-yl)-1,1,3,3-TetramethylUronium hexafluorophosphate
HFIP	hexafluoroisopropanol
HPLC	high performance liquid chromatography
IMS-MS	ion mobility spectrometry- mass spectrometry
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
MD	molecular dynamics

min	minute
mM	millimolar
mL	milliliter
Rho110	Rhodamine 110
RP-HPLC	reversed-phase high-performance liquid chromatography
TEM	transmission electron microscopy
ThT	Thioflavine T
TFA	trifluoroacetic acid
TLC	thin layer chromatography
UV	ultra-violet
μL	microliter
μM	micromolar

Chapter 1: Potential of Dynamic Chemical Network for Chemical Evolution

1.1 Chemical evolution – fundamental to Biological evolution

Charles Darwin's publication of "On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life" in 1859 defined biological evolution by natural selection, later described by Herbert Spencer (1864) as the "survival of the fittest". In the book, Darwin argued that species originated from a common ancestor even with no understanding of nucleic acid polymers serving as genetic information carriers. Later scientific discoveries suggested that all life forms originate from a common ancestor and life's essence emerged from simple chemistry. Today we can extend the analysis of how complex living forms originated from simple living forms, biological evolution, to address the pressing question of how the first living organism emerge at the very beginning, the physical process of chemical evolution. It is natural then to understand the origin of life on earth as two stages: 1) the **chemical evolution** stage, where inanimate matter generates a set of biopolymers to give simple living networks; 2) the **biological evolution** stage, where life forms grow from simple to more complex systems.

The arguments above maintain that living systems are a complex set of chemical reactions from which progressive emergent behaviors are possible. To gain more insight into the behavior of these networks of biopolymers and a fundamental understanding of biological evolution, we may

need to create such chemical networks. The basic hypothesis is if the biological species demonstrate evolution under selection pressure, then the biopolymers which constitute for biological species should also display these properties of adaptation and evolution under selection pressure. Mills et al. (Mills, Peterson, & Spiegelman, 1967) observed that Qβ RNA oligonucleotide replicated, mutated and evolved to a sequence containing just 17% of its original length when reacted with activated nucleotides in the presence of the replicase. The loss of 83% of the sequence, dispensable under the conditions of the experiment, conferred a selective advantage to increase the replication efficiency by 15 times fold. The process of replication, mutation, selection and evolution demonstrate by this early experiment outside of a living cell suggests that Darwinian behavior can also be reflected at this molecular level. Since then in vitro evolution procedures have been extended across a wide range of nucleic acid systems by Bartel and Szostak (Bartel & Szostak, 1993), Joyce et al. (Joyce, 2004; Wright & Joyce, 1997) and Johnston et al. (Johnston, Unrau, Lawrence, Glasner, & Bartel, 2001).

The in vitro evolution is not only seen in a single species but also expressed in a more striking expression of natural selection between species that compete for resources. Joyce and coworkers (Voytek & Joyce, 2009) reported on two RNA enzymes challenged to compete for a same substrate, extinction occurs. However, when the two enzymes are incubated with 5 potential substrates, each enzyme adapted to different substrate so that they can co-evolve in the same environment. Even the simple systems demonstrate competitive and mutualistic behaviors at the chemical level and mimic the evolutionary process of living system, genetic materials are selected by environmental inputs (Sulloway, 1982).

1.2 Two types of information transfer in evolution: digital and analog (genetic and epigenetic)

Nucleic and amino acid polymers are the two different types of informational biopolymers. As outlined in (Jay T. Goodwin, Mehta, & Lynn, 2012), the genetic information encoded in DNA is considered "digital" as nucleic acid base-paring demonstrates folding behavior with geometrically discrete hydrogen bonding patterns to assure replication fidelity of complementary strands. In contrast, protein folding is demonstrated by many context dependent forces, including van der Waals, aromatic, ionic and hydrogen-bonding interactions provided by the amino acid side chains and amide backbone, which make its folded states more "analog-like" in their behavior.

Chemical evolution uses both strategies represented by these two biopolymers to achieve information transfer: in digital fashion (viral evolution) and in analog fashion (peptide/ protein evolution) (Jay T. Goodwin et al., 2012). In viral evolution (**Figure 1-1 A**), the digital information encoded as linear sequences of nucleobases is replicated with high fidelity to its complement through base-paring. Diversity is generated through copying errors and mutations. Selection of a specific nucleic acid sequence is determined by functional benefits of the mutants and the fittest species will survive to make more copies at the expense of others. However, in prion evolution (**Figure 1-1 B**), the information encoded in a cross-β architecture conformation is replicated through templated associations. Diversity is achieved in the template-directed propagation of the protein conformation, and selection again is dictated by replication fitness. Li et al. demonstrate Darwinian evolution of prions in cell culture (J. Li, Browning, Mahal, Oelschlegel, & Weissmann, 2010), where cloned prion populations gradually show heritable phenotypic properties changes resulted from mutations, and the emergence of distinct

populations in different environments resulted from selective pressures. This epigenetic information encoded in protein sequence can induce heritable changes without involving nucleic acids and is dictated by sets of intra- and intermolecular interactions highly dependent on complex environments. Thus, these proteins act as genes by templating their conformation, just as DNA templates its sequence (Wickner et al., 2014).

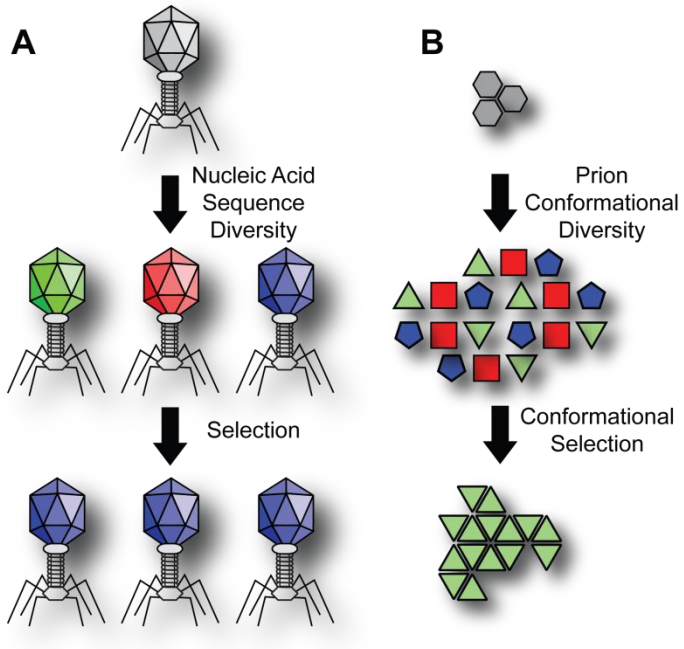


Figure 1-1 Two chemical evolution strategies. Digital evolution is represented by (A) viruses where nucleic acid sequence is varied by mutation and selection is from a diverse population of distinct molecules. In (B), prion evolution creates conformational diversity and selection occurs from the ability of those conformations to propagate through templated conformational events. Colors indicate distinct species resulting from mutation in DNA sequence or change in protein conformation (Jay T. Goodwin et al., 2012).

1.3 Emergence of the ribosome through mutualism between nucleic acid and amino acid polymers

The ribosome is the only organelle that exists in both procaryotes and eucaryotes, suggesting its evolutionary role in the last universal common ancestor (LUCA) (Theobald, 2010; C.R. Woese, Kandler, & Wheelis, 1990). The emergence of ribosome then serves as a Darwinian Threshold for extant cellular life (C. R. Woese, 2002). It translates the genetic information encrypted in mRNA sequences to make peptide chains, serving as a digital to analog converter (**Figure 1-2**). This converter itself is a multicomponent supramolecular assembly that generally contains 3-4 RNA subunits and more than 70 ribosomal proteins (Schmeing et al., 2009), and reflects a mutually beneficial co-evolutionary history that directly connects the behaviors of two distinct biopolymer species in a mutual symbiosis (Harish & Caetano-Anolles, 2012).

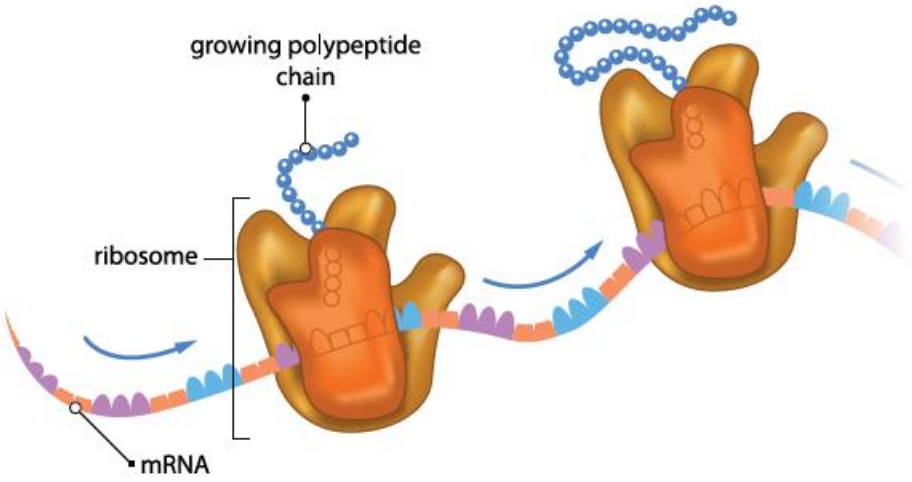


Figure 1-2 The ribosome represents the invention that allows two biopolymer classes, nucleic and amino acid biopolymers, to cooperatively transfer information (J.T. Goodwin et al., 2014).

This synergistic collaboration of nucleic acid-like and amino acid-like polymers for molecular information translation extends almost a billion years before the emergence of the ribosome (Cavalier-Smith, 2006; Schopf, 2006) and that spontaneous emergence of chemical networks which crossed the chemical threshold to give rise to the nucleic acid and amino acid biopolymers, so called **chemical evolution**, could lead to the molecular mutualism seen in the **Figure 1-3**.

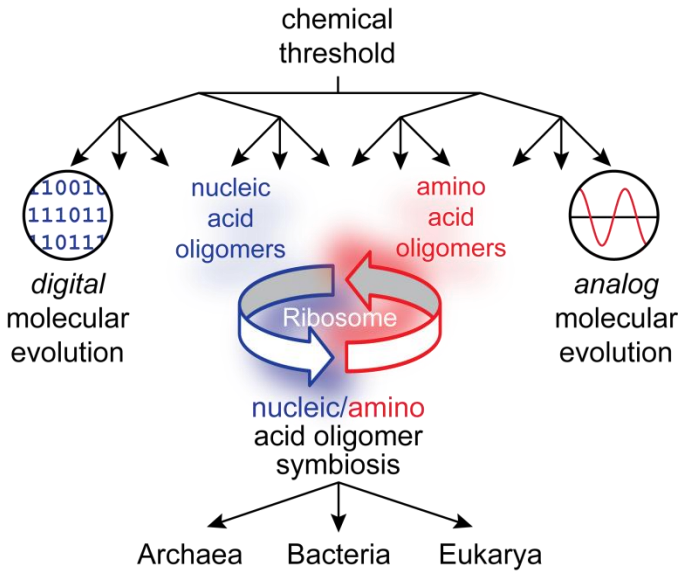


Figure 1-3 Polymer mutualistic symbiosis as the threshold for the three domains of life.The mutuality synergy of nucleic and amino acids to create the ribosome requires a diverse polymer network capable of chemical evolution. The network is shown as bounded by both chemical and Darwinian thresholds (Jay T. Goodwin et al., 2012).

Therefore, it is necessary to create a **dynamic chemical network (DCN)** based on a simple chemical inventory that is able to produce biopolymers. The question is whether the DCN can

serve as the chemical threshold for the creation of a diverse polymer network that allows for dynamic mutualistic co-evolution to give rise to multicomponent supramolecular functional assemblies.

1.4 Dynamic Combinatorial Chemistry and Dynamic Chemical Network

Last two decades witness the booming of research in Dynamic combinatorial chemistry (DCC) (Corbett et al., 2006; Cougnon & Sanders, 2011; de Bruin, Hauwert, & Reek, 2006; Lehn, 2007; Otto, Furlan, & Sanders, 2002), which is a powerful tool for creating chemical complexity in dynamic combinatorial library (DCL) from simple building blocks under thermodynamic control, where the members are generated in a combinatorial way by linking building blocks together through reversible chemical bonds (**Figure 1-4**). Since these members interconvert continuously by exchanging building blocks with each other through breaking/ reforming bonds and the distribution of the members respond to environmental input, "**network**" instead of "**library**" is used in this thesis to better reflect its dynamic property.

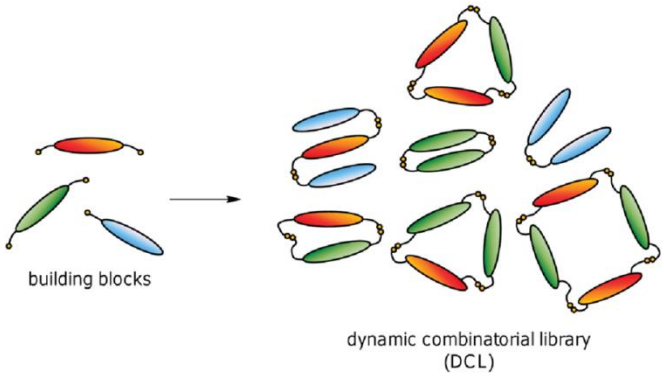
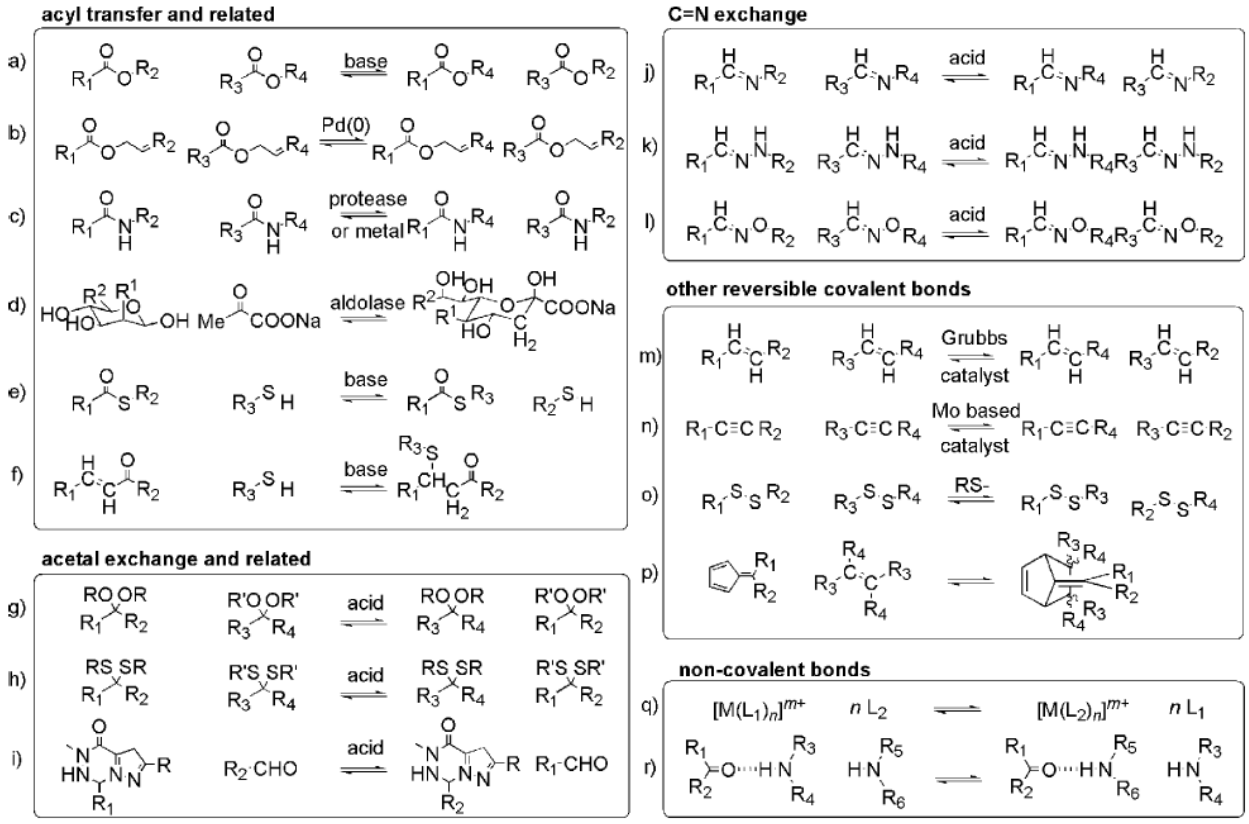


Figure 1-4 Representation of a dynamic combinatorial library built from simple building blocks.

Figure adapted from (Cougnon & Sanders, 2011).

1.4.1 Reversible Reactions in DCN

Many reversible reactions have been utilized in DCN (Corbett et al., 2006) and most of them are summarized in **Scheme 1-1**, including imine exchange, hydrazone exchange, disulfide exchange (Carnall et al., 2010; Otto et al., 2002), transesterification, transamidation, acetal exchange, aldol exchange, transthioesterification, Michael/retro-Michael reactions, etc. Only imine and acetal reactions are discussed below in detail as these two involve in the work described in this thesis.



Scheme 1-1 Reversible reactions used for DCC to date: (a) transesterification; (b) transallylesterification; (c) transamidation; (d) aldol exchange; (e) transthioesterification; (f) Michael/retro-Michael reactions; (g) acetal exchange; (h) thioacetal exchange; (i) pyrazolotriazone metathesis; (j) transimination; (k) hydrazone exchange; (l) oxime exchange; (m) alkene metathesis; (n) alkyne metathesis; (o) disulfide exchange; (p) Diels-Alder/retro-Diels-

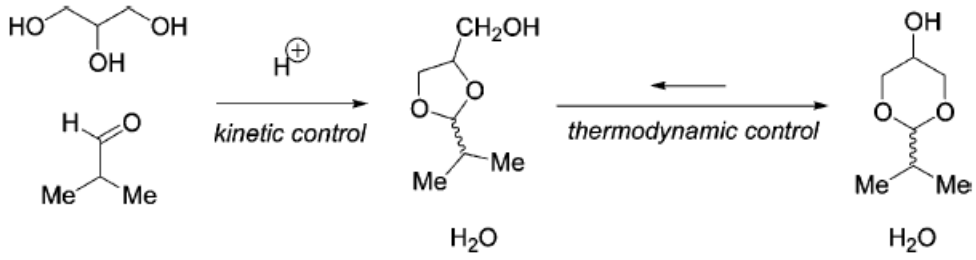
Alder reactions; (q) metal-ligand exchange; (r) hydrogen-bond exchange. This scheme is modified from (Corbett et al., 2006).

1.4.1.1 Imine Exchange in DCN

The most common method for preparing imines is the acid-catalyzed reaction of aldehydes or ketones with amine derivatives. The first dynamic chemical network (DCN) of imine reversible linkage is discussed in 1992 by Goodwin et al. in our group (Jay T. Goodwin & Lynn, 1992), since then DCNs constructed with imine exchange have been extensively studied (Giuseppone & Lehn, 2006; Godoy-Alcántar, Yatsimirsky, & Lehn, 2005; Huc & Lehn, 1997; Wessjohann, Rivera, & Leo´n, 2007; Zameo, Vauzeilles, & Beau, 2006). The direct analysis of imine networks is hindered due to the inherent lability of the imine linkage, making chromatography difficult. However, adjusting pH can slow down the imine hydrolysis rate and allow for direct composition analysis in aqueous medium (Zameo et al., 2006). Alternatively, kinetic tapping of imine-based networks can be achieved through reduction to amines using reducing agents, such as sodium borohydride or tetrabutylammonium cyanoborohydride.

1.4.1.2 Acetal Exchange in DCN

The acid-catalyzed formation of cyclic acetals from diols and aldehydes provides another thermodynamically controlled reversible reaction (Stoddart, 1971). A classic example is seen in the acetal condensation of glycerol and isobutyraldehyde under acidic conditions, where five-membered ring acetals form faster, however, the six-membered ring acetals are thermodynamically more stable and become dominant products at the cost of five-membered ring acetals (**Scheme 1-2**) (Eliel & Wilen, 1994).



Scheme 1-2 Formation of five- and six-membered cyclic acetals from glycerol and isobutyraldehyde under thermodynamic control in the DCN (Eliel & Wilen, 1994).

1.4.2 External Templating Shifts Equilibrium in DCNs

Currently the most explored aspect of DCNs is for new ligands for biomolecules or synthetic receptors for small molecules may be relevant in drug discovery (Shi, Stevenson, Campopiano, & Greaney, 2006). The nature of DCNs allows equilibrium shift towards a favored member with highest affinity when target molecules are added to DCNs as external templates. Molecular recognition between the template and network species shifts the initial equilibrium and leads to the amplification and selection of the product that binds the template with highest affinity. The principle of "survival of the fittest" in biological species is also reflected in the molecular level. The strategies for selection can be divided into two categories: (**a**) selection of a host from DCNs by an introduced guest; and (b) selection of a guest from DCNs by an introduced host (**Figure 1-5**). This effect has been utilized for the discovery of synthetic receptor (Besenius, Cormack, Ludlow, Otto, & Sherrington, 2008; Bru, Alfonso, Bolte, Burguete, & Luis, 2011; Bru, Alfonso, Burguete, & Luis, 2006; Custelcean, 2012) and ligands for biomacromolecules (Azema, Bathany, & Rayner, 2010; Demetriades et al., 2012; McNaughton, Gareiss, & Miller, 2007; Scott, Dawes, Ando, Abell, & Ciulli, 2009; Valade, Urban, & Beau, 2006).

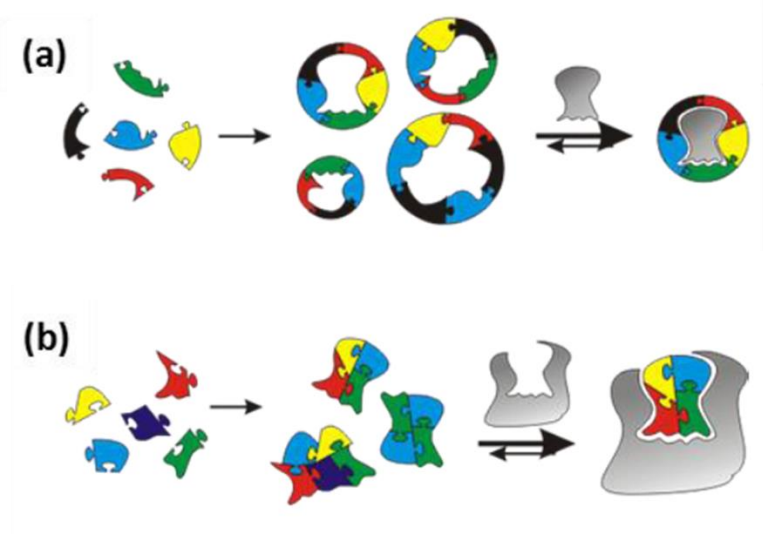


Figure 1-5 Strategies of selecting specific member of DCNs by external templates: (**a**) selection of a host by an introduced guest; and (**b**) selection of a guest by an introduced host. Figure is modified from (J. Li, 2014; Otto et al., 2002).

1.4.3 Internal Templating Shifts Equilibrium and Drives Self-assembly in DCNs

In contrast to external templating, where recognition takes place between exogenous templates and network members, the molecular recognition occurs within network members in internal templating. Network species bind intermolecularly to copy themselves, leading to self-assembly and driving the equilibrium to favor the very species that self-assemble (**Figure 1-6**). Recently internal templating has demonstrate fascinating results in selection of specific species in DCNs.

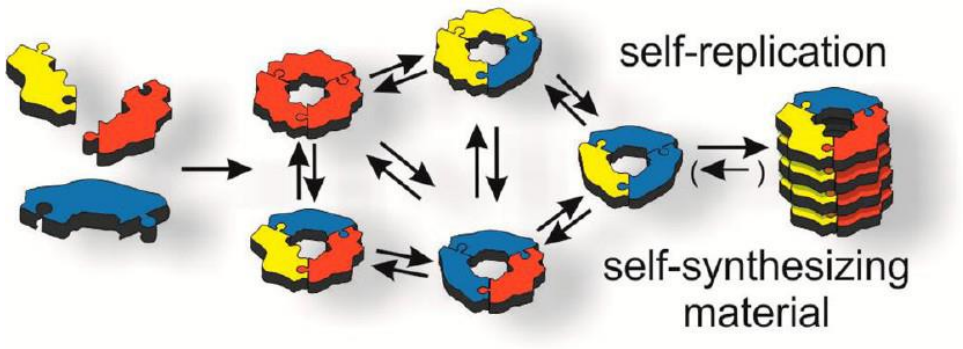


Figure 1-6 Internal templating in a DCN. A molecule that can bind to copy itself is able to promote its own formation by shifting the equilibrium in a DCN (Malakoutikhah et al., 2013).

Otto and coworkers demonstrate a self-replicating network constructed by dithiol-functionalized peptide-derived building blocks (Carnall et al., 2010), where a mixture of macrocycles is generated through disulfide exchange (**Figure 1-7 a**). Hexamer and heptamer can form fibers due to proposed stacking of peptide chains (**Figure 1-7 b**). When the fibers grow long enough, they become susceptible to shear stresses and break, duplicating the number of catalytically active fiber ends and further driving equilibrium away from solution. Different mode of mechanic force, shaking or stirring, induces competition between those two replicators and leads to selection of hexamer or heptamer (Carnall et al., 2010). This suggests that kinetic product from an assembly process can dominate at the cost of thermodynamic product. The transition from thermodynamic to kinetic control in a DCN implies the potential for studying the emergence of biopolymers and chemical evolution en route to the origin of life, as life is far from equilibrium.

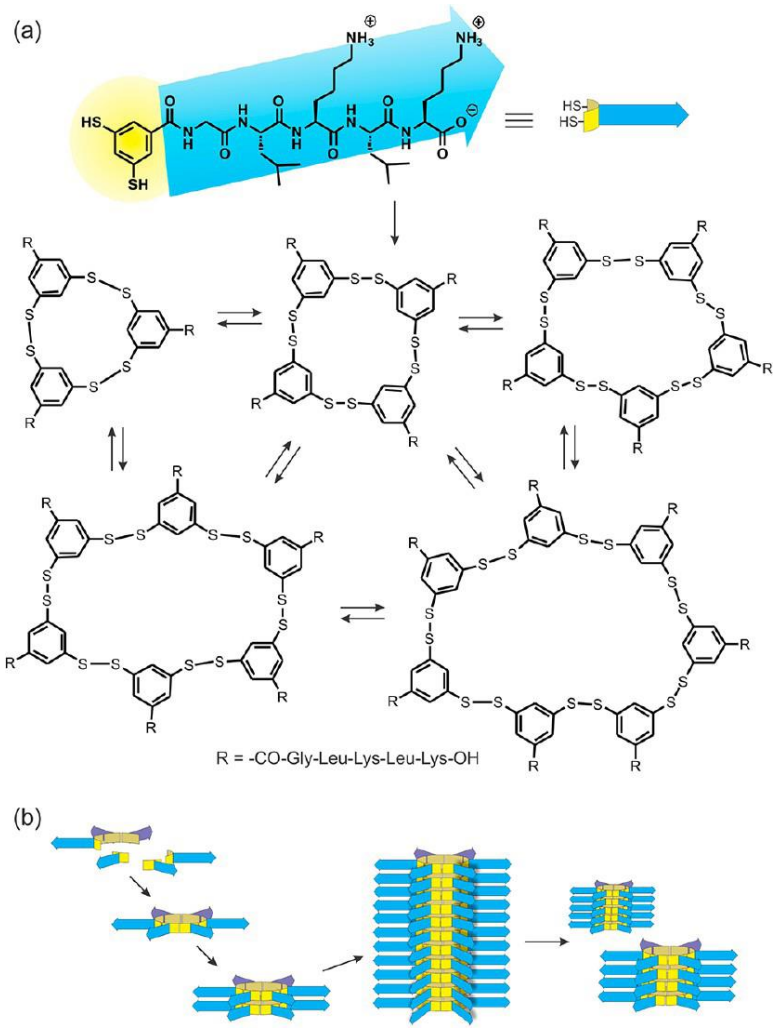


Figure 1-7 A self-replicating network constructed by dithiol-functionalized peptide-derived building blocks. (a) A dithiol-functionalized peptide building blocks are oxidized and form a mixture of macrocycles through disulfide exchange. (b) Mechanical breakage of the fibers duplicates the number of replicators. This figure is modified from (Carnall et al., 2010; J. Li, Nowak, & Otto, 2013).

1.4.4 Self-replicating DCN may serve as the chemical threshold for biopolymer emergence and evolution

Replication of biopolymers is fundamental and essential to all living systems, and the replication behavior has likely played a crucial role in competition in the "primordial soup" for the winning molecules, paving the way for the origin of life. Recently, the studies in the realm of DCNs have extended from selection of particular network members by introducing external templates to exploring the self-replicators that grow by making copies of themselves, similarly to the way that living organisms operate under far-from-equilibrium conditions by constant replication. Besides replication, the responsiveness of chemical networks to environmental inputs is also fundamental for understanding the adaptation of biological network to the environment.

Feedback events where network members respond through self-assembly (Carnall et al., 2010) allow network members to replicate themselves and have the potential for the progressive growth. In these cases, distributions of DCNs are no longer in homogenous solutions but affected by appearance of a different phase. The self-recognition and the phase change in DCNs can also provide transitions between equilibrium thermodynamics and processes under kinetic control. Combining the principles of DCC with self-replicating molecules appears to be a very promising way to study the chemical evolution. These dynamic chemical networks may serve as the chemical threshold for biopolymer emergence and evolution, like the "warm little pond" that Darwin proposed for the origin of life.

1.5 Construction of Dynamic Chemical Network in terms of three dimensions

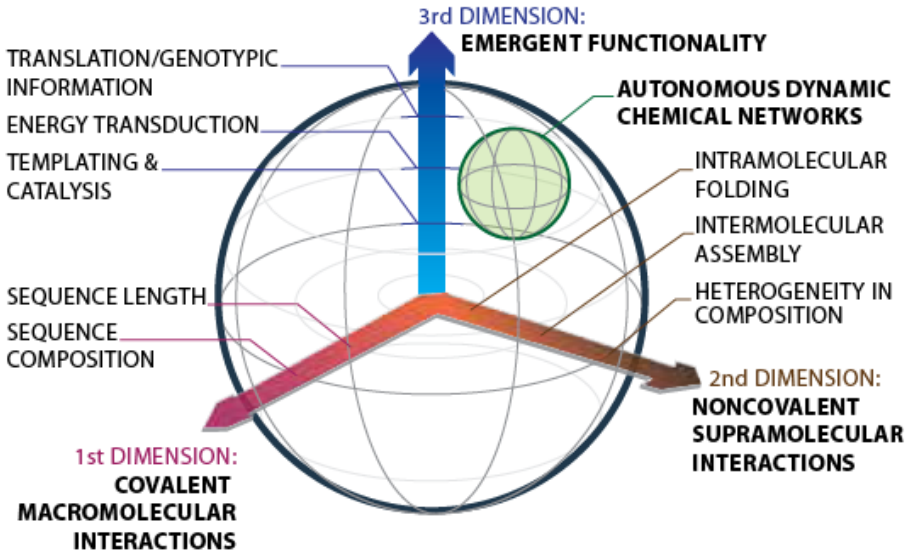


Figure 1-8 Construction of dynamic chemical network through three dimensions. (J.T. Goodwin et al., 2014)

In this dissertation, I demonstrate the *de novo* construction of peptide surrogate polymers linked with reversible acetals in dynamic chemical networks in three separate yet mutually dependent dimensions. The first dimension constitutes the covalent, macromolecular synthesis from simple monomers as building blocks to access chain-length and sequence diversity, where complexity is reflected in the composition of network members with varied length and sequence. The second dimension includes the non-covalent associations and supramolecular folding among the macromolecular species synthesized from the first dimension, extending the structural diversity in the network. The third dimension encompasses the chemical and physical functions (such as templating & catalysis, energy transduction and translation of encoded information, etc.)

emerging from the diverse structural networks built on the first two dimensions. The construction of the DCNs exhibits three dimensions in a progressive way to achieve hierarchical molecular order.

Chapter 2: Exploring Reversible Acetal Linkages for Constructing Dynamic Chemical Networks

2.1 Introduction

2.1.1 Reversible Imine and Acetal linkages in Nucleic Acid Dynamic Chemical Networks

The first dynamic chemical network (DCN) of imine reversible linkage is achieved in our lab in 1992 (Jay T. Goodwin & Lynn, 1992). In the presence of the DNA hexamer template dGCAACG, a complementary 5'-amine trimer 5'-$^{+}H_3N$-dTGC, competes with non-complementary tetramer sequence of 5'-$^{+}H_3N$-dTTTT, for the coupling of the complementary 3'-aldehye trimer (**Figure 2-1**). The nature of imine lability provides equilibrium between coupled and uncoupled substrates, allowing for thermodynamics of substrate-template association to direct product formation. The DNA template biased the production of the complementary hexamer 30-fold over the non-complementary heptamer at 0°C.

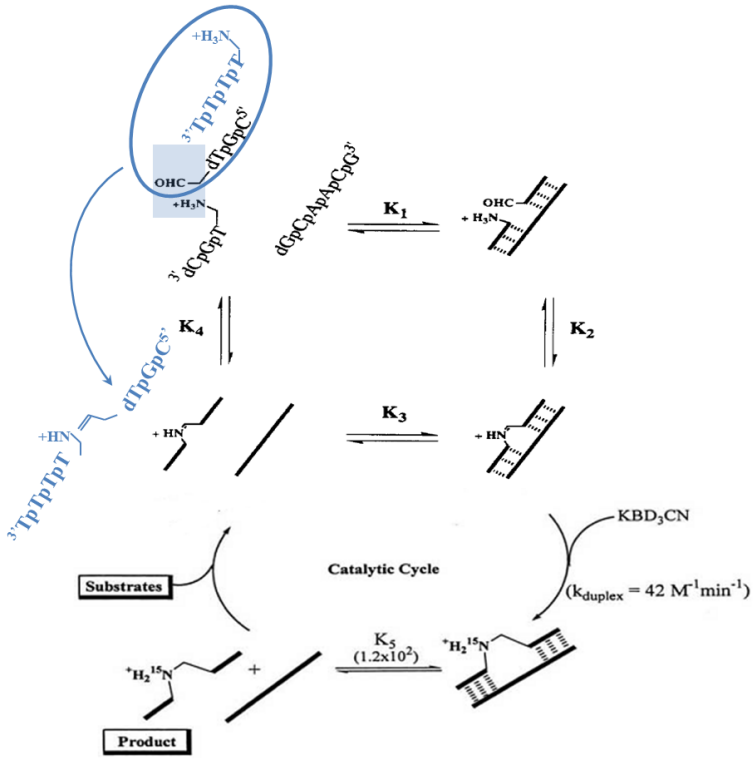


Figure 2-1 DNA template-directed synthesis in imine-linked dynamic network from mono-functionalized substrates (figure modified from (Jay T. Goodwin & Lynn, 1992) (Z. J. Zhan & D. G. Lynn, 1997)).

In 2002, the imine linked network is extended from mono-functionalized substrates with amine or aldehyde to bis-functionalized substrates 5'-H_2N-dT-3'-CH_2CHO, where the equilibrium of imine formation is driven by the octameric DNA template $(dAp)_8$ under thermodynamic control (Xiaoyu Li, Zhan, Knipe, & Lynn, 2002). In this DCN, the imine coupling among monomers in aqueous solution is surprisingly robust and achieved the first chain-length-specific and sequence specific reading of a DNA template through monomer polymerization. The reaction demonstrate a new step-growth kinetics mechanism, and is extended to copy information from a 32-mer DNA sequence (**Figure 2-2**) (X. Li, Hernandez, Grover, Hud, & Lynn, 2011). After genotypic

information transferred to imine linked polymer, the polymers are reduced by sodium borohydride to amine nucleoside polymers (ANP), which are more stable. While reductive amination is synthetically useful at generating ANP, this reduction method fell short of maintaining a dynamic network as reduction of imine to amine is irreversible.

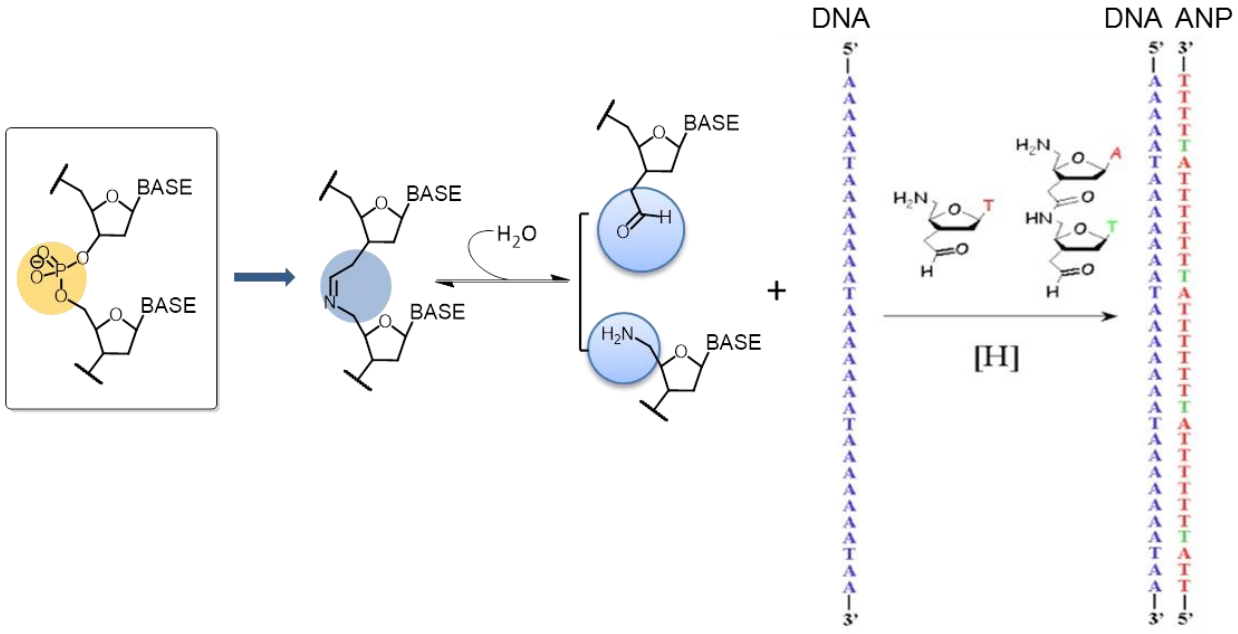


Figure 2-2 DNA template-directed synthesis in imine-linked dynamic network from bis-functionalized substrates. (figure modified from reference (X. Li et al., 2011))

If the imine can be stabilized through an additional reversible reaction instead of irreversible reduction, then the dynamics in the network will be maintained for responsiveness towards external stimuli. Incorporating this imine condensation reaction in an RNA backbone positions the 2'-OH as an intramolecular trap for the intermediate imine, and unlike the ANP, the resulting N,O-acetal would maintain reversibility through a condensation/hydrolysis manifold.

2.1.2 Extending Reversible Acetal linkages in Peptide Dynamic Chemical Networks

I seek to extend this reversible acetal linkage from nucleic acid biopolymers to peptide, reasoning that not only nucleic acid polymers can be connected through kinetically accessible linkages for replacement of phosphodiester bond, but peptides should also be engineered with similar reversible linkages. The hypothesis is that the peptide dynamic network linked by acetal linkage will be able to respond towards external stimuli and provide opportunity for selection of desired oligomer and progressively generated assembly, taking one step closer to an evolvable system.

Simple change in the oxidation state of the C-terminal amino acid to an aldehyde allows coupling of another amino acid's N-terminal amine functional group as imine linkage for generating dynamic networks. Taking advantage of the greater inventory of amino acids with various side chains, the imine linkage can be further stabilized through intramolecular trapping by residue side chains. Side chain of hydroxyl functional group including serine (Ser), threonine (Thr) and homoserine (Hse) residues are positioned to trap imine intermediate as five-membered or six-membered ring N,O-acetal (oxazolidine) respectively, cysteine (Cys) side chain thiol group as a five-membered ring N,S-acetal (thiazolidine) and Asparagine (Asn) as a six-membered ring N,N-acetal (tetrahydropyrimidone) (**Figure 2-3**). The coupling reaction can be further diversified via other amino acids such as aspartic acid (Asp), histidine (His), etc.

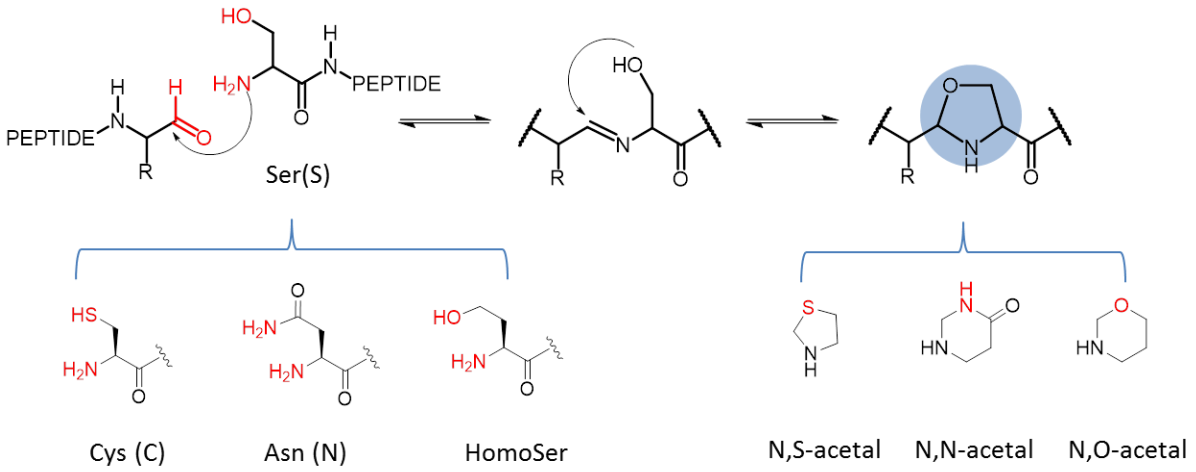


Figure 2-3 Structural illustration of reversible acetal linkages formed by trapping of imine intermediates via amino acid side chains.

To explore the generality of these linkages and examine if these reactions are robust to construct dynamic network, I use mono-functionalized substrates of Ser, Cys, Asn residues (providing amine and additional traps for imine intermediates) and phenylalanine (Phe) residue with modified aldehyde group (-CHO), *N*-Boc-Phe-CHO, to study N,O-acetal, N,S-acetal and N,N-acetal, respectively. Although any amino acid's C-terminal can be modified with aldehyde functional group theoretically, the choice of Phe is due to the great potential of the phenyl ring for promoting higher ordered structures in peptides (Adler-Abramovich et al., 2006; Frederix, Ulijn, Hunt, & Tuttle, 2011; Mehta et al., 2008; Reches & Gazit, 2006; Scanlon & Aggeli, 2008).

2.1.3 Understanding Stereochemistry in Acetal Chiral Centers

These acetal condensation reactions create new chiral centers at the acetal carbons with two possible isomers: (S,S) *trans* and (R,S) *cis* isomers (**Figure 2-4**).

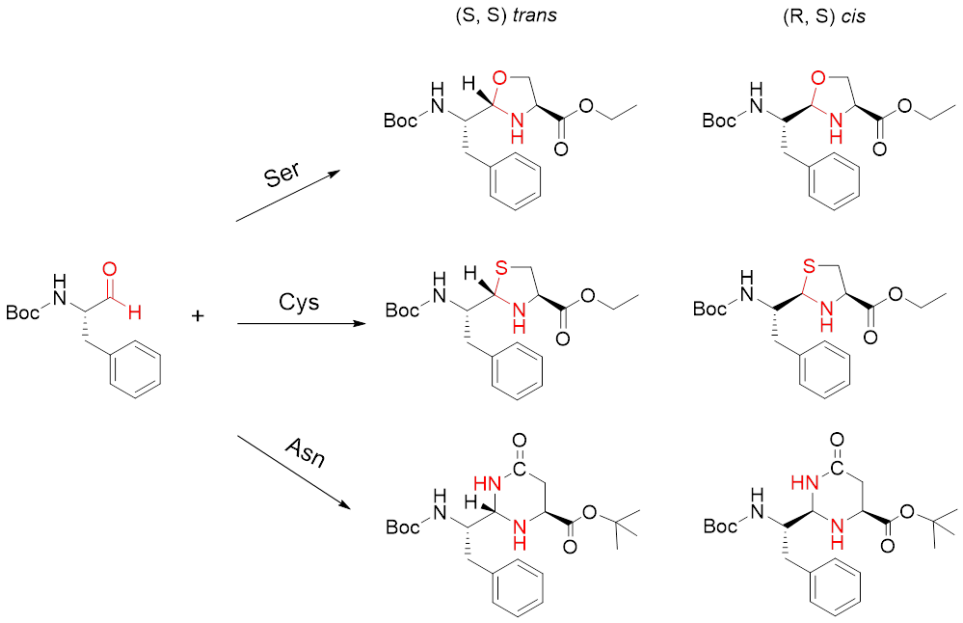


Figure 2-4 (S,S) trans isomer and (R,S) cis isomer of N,O; N,S; N,N-acetal products generated from N-Boc-Phe-CHO with L-Ser ethyl ester, L-Cys ethyl ester and L-Asn tert-butyl ester, respectively

Previous researches (D. Seebach, Aebi, Gander-Coquoz, & Naef, 1987; D. Seebach, Sommerfeld, Jiang, & Venanzi, 1994; Szilagyi & Gyorgydeak, 1979) have shown that acid-catalyzed acetalization of serine, threonine and cysteine esters gives rise to *ca*. 1:1 mixtures of *cis*- and *trans*- substituted heterocycles in various solvents, e.g. acetone, acetic acid, DMSO and pyridine, etc. However when acetals are acylated, pure *cis*- diastereomer result (**Figure 2-5**). The difference between the heterocycles with and without acylation on the ring is due to the well-

known allylic 1,3- strain exerted by an amide group (Hoffmann, 1989) (D. Seebach, Lamatsch, B., Amstutz, R., Beck, A. K., Dobler, M., Martin Egli,, Robert Fitzi, Miguel Maestro, & Christof Schickli, 1992). The energy required to rotate an amide bond is 15-20 kcal/ mol, much larger than the conformational strain (e.g. switching substituents in the axial or equatorial positions).

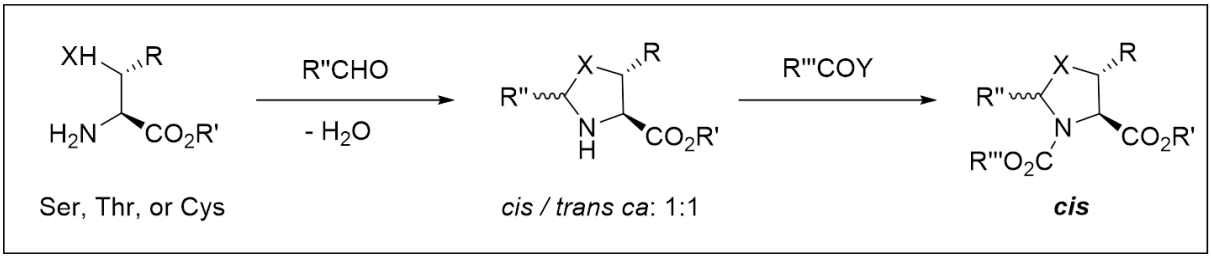


Figure 2-5 Chemical structure of acetals derived from amino acids Ser, Thr and Cys (D. Seebach, Lamatsch, B., Amstutz, R., Beck, A. K., Dobler, M., Martin Egli, et al., 1992). Without acylation on the acetal ring, Ser, Thr and Cys esters gives rise to *ca*. 1:1 mixtures of *cis*- and *trans*- substituted heterocycles. Pure *cis*- diastereomer results from acylation of acetals.

2.2 Materials and Methods

2.2.1 Reagents

Substrates of *N*-Boc-L-phenylalaninal (*N*-Boc-Phe-CHO), L-Serine ethyl ester hydrochloride, L-Cysteine ethyl ester hydrochloride and L-Asparagine *tert*-butyl ester hydrochloride for dimerization reactions are purchased from Sigma-Aldrich. The solvents of benzene-d_6, MeCN-d_3 and DMSO-d_6 are purchased from Cambridge Isotope Laboratories, Inc.

2.2.2 Preparation of Free Amine Substrates by Neutralizing Ser, Cys and Asn Ester Hydrochloride Salts

Before mixing aldehyde substrate with the amine substrate for acetal condensation, the Ser, Cys and Asn ester hydrochloride salts (50 mg) are neutralized by 30% NH_4OH solution (1 ml) and the resulting water and excess NH_4OH solution is removed in vacuum at room temperature. The resulting residue is then dissolved in organic solvent such as acetonitrile (MeCN) (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 down to yield the free amine of Ser, Cys and Asn residues. The ^{1}H NMR spectra of the free amine and the aldehyde substrates are recorded on an INOVA 600 NMR spectrometer or INOVA 400 NMR spectrometer as follows:

L-Ser ethyl ester: ^{1}H NMR (400 MHz, Benzene-d_6) δ 7.16 (d, J = 1.3 Hz, 1H), 3.99 – 3.77 (m, 2H), 3.79 – 3.62 (m, 2H), 3.27 – 3.18 (m, 1H), 2.43 (s, 4H), 0.98 – 0.77 (m, 3H).

L-Cys ethyl ester: ^{1}H NMR (600 MHz, Acetonitrile-d_3) δ 4.41 (s, 3H), 4.25 – 4.13 (q, J = 9.65, 2H), 3.98 (dd, J = 5.5, 4.4 Hz, 1H), 3.07 – 2.96 (dd, J = 9.7, 4.4 Hz, 1H), 2.92 (dd, J = 14.2, 5.5 Hz, 1H), 1.63 (d, J = 5.5 Hz, 1H), 1.46 (s, 1H), 1.34 – 1.18 (t, J = 7.1, 3H).

L-Asn tert-butyl ester: ^{1}H NMR (400 MHz, Acetonitrile-d_3) δ 6.74 (s, 1H), 5.92 (s, 1H), 3.53 (dd, J = 8.0, 4.5 Hz, 1H), 2.46 (dd, J = 15.4, 4.5 Hz, 1H), 2.33 (dd, J = 15.4, 8.0 Hz, 1H), 1.84 (s, 2H), 1.42 (s, 9H).

N-Boc-Phe-CHO: ^{1}H NMR (400 MHz, Acetonitrile-d_3) δ 9.54 (s, 1H), 7.39 – 7.05 (m, 6H), 5.55 (s, 1H), 4.15 (ddd, J = 9.5, 7.7, 5.0 Hz, 1H), 3.14 (dd, J = 14.1, 5.0 Hz, 1H), 2.79 (dd, J = 14.1, 9.4 Hz, 1H), 1.34 (s, 9H).

2.2.3 Mixing of Ser/Cys/Asn with the aldehyde for N, O-/ N, S-/ N, N-acetal product, respectively

Free amine substrates of Ser, Cys and Asn residues and the aldehyde substrate *N*-Boc-Phe-CHO are dissolved individually in solvents. And then the amine substrate is added into the aldehyde solution closely to 0.99: 1 with 45 mM concentration for each substrate. Specifically, for example, *N*-Boc-Phe-CHO (14.58 mg) is dissolved in 650 µl acetonitrile-d_3 for 90 mM solution and adding 321 µl of the Ser substrate solution with the same concentration to 325 µl of the aldehyde solution gave rise to the final mixture. Note that 1) the sequence of adding substrate is important and the amine should be added into the aldehyde, not vice versa due to that side reactions might be caused by nucleophilic attack to the imine intermediate from additional primary amines of Ser, Cys or Asn; 2) keeping the aldehyde substrate slightly in excess is due to the same reason mentioned above. The reaction progress is followed by ^{1}H NMR spectroscopy and recorded on an INOVA 600 NMR spectrometer or INOVA 400 NMR spectrometer.

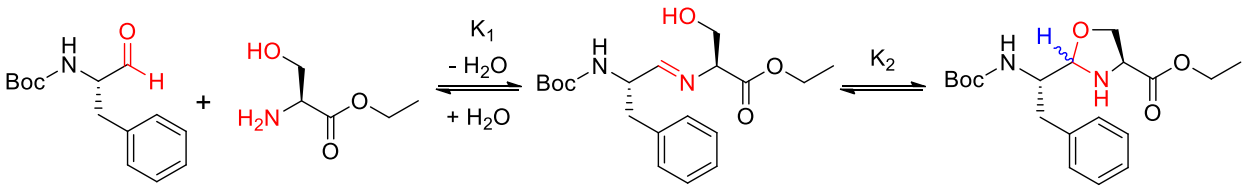
2.3 Results and Discussions

2.3.1 N, O-acetal Linkage

2.3.1.1 Diastereoselectivity of N,O-acetal in benzene

Mono-functionalized two substrates *N*-Boc-Phe-CHO and L-Ser ethyl ester are selected for studying N, O-acetal linkage. As shown in **Scheme 2-1**, this condensation reaction includes two steps: the first step is intermolecular nucleophilic attack of amine functional group from L-Ser ethyl ester to aldehyde, which forms imine intermediate and releases a water molecule; the

second step is intramolecular nucleophilic attack involving hydroxyl group from Ser side chain to form N, O-acetal product. The overall equilibrium constant can be calculated by the equations shown in **Scheme 2-1, bottom**. To define this ligation reaction and follow its kinetics, substrates are mixed *ca* 1:1 ratio in benzene-d_6 and followed by NMR.



$$K_1 = \frac{[\text{Imine}][H_2O]}{[\text{Phe}][\text{Ser}]}; K_2 = \frac{[\text{Oxazolidine}]}{[\text{Imine}]}; K_{overall} = K_1 \times K_2 = \frac{[\text{Oxazolidine}][H_2O]}{[\text{Phe}][\text{Ser}]}$$

Scheme 2-1 (**top**) Imine/ acetal condensation products generated from reactants *N*-Boc-Phe-CHO and L-Ser ethyl ester (reactive functional groups are labeled as red). (**bottom**) equilibrium constants for step reactions and overall reaction.

Once the two substrates *N*-Boc-Phe-CHO (**Figure 2-6, A**) and L-Ser ethyl ester (**Figure 2-6, B**) are mixed in benzene-d_6 closely to 1:1 ratio as described in the Method session, the course of condensation reactions is followed by the CH signal of the imine bond in the ^{1}H NMR spectra appeared downfield at 7.5 ppm (data not shown) relative to the peak of the parent aldehyde at 9.17 ppm (**Figure 2-6, A**). Imine condensation product then quickly converts to N,O-acetal product within minutes. The increase of N,O-acetal proton resonance at 4.25 ppm (**Figure 2-6, C**) is concomitant with the decrease and eventual disappearance of imine resonance. The yields of

imine and acetal are determined by the integration of imine and acetal resonances to the internal standard of phenylalanine ring. Within two hours, reactions reaches equilibrium and most of the reactants convert to the acetal product (**Figure 2-7**). Detailed assignments of the acetal product have been shown in 2D COSY spectrum (**Figure 2-8**).

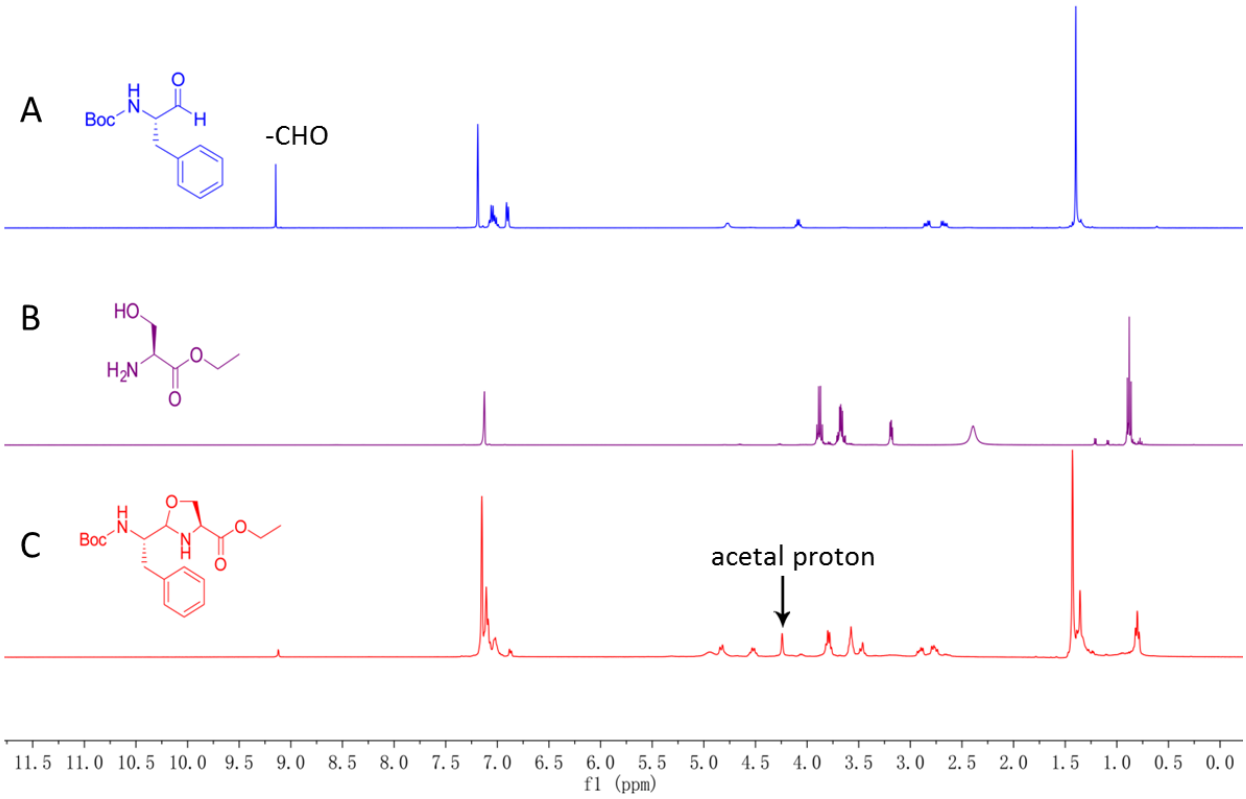


Figure 2-6 NMR spectrum of substrate *N*-Boc-Phe-CHO (**A**), L-Ser ethyl ester (**B**) and the N,O-acetal condensation product at 100 minutes (**C**) in benzene-d_6

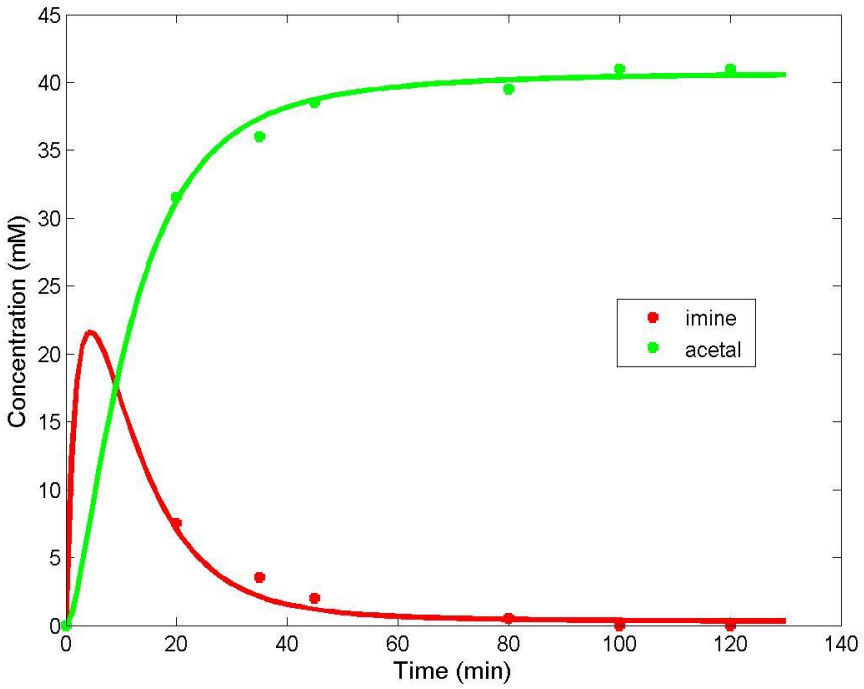


Figure 2-7 Model fits of imine intermediate and N,O-acetal kinetics in benzene. Model fits of imine intermediate and N,O-acetal kinetics in benzene. Forward rate constant and backward rate constant for imine condensation are 4.14E-01 and 1.00E-01; forward rate constant and backward rate constant for acetal formation are 1.09E-01and 2.12E-03.

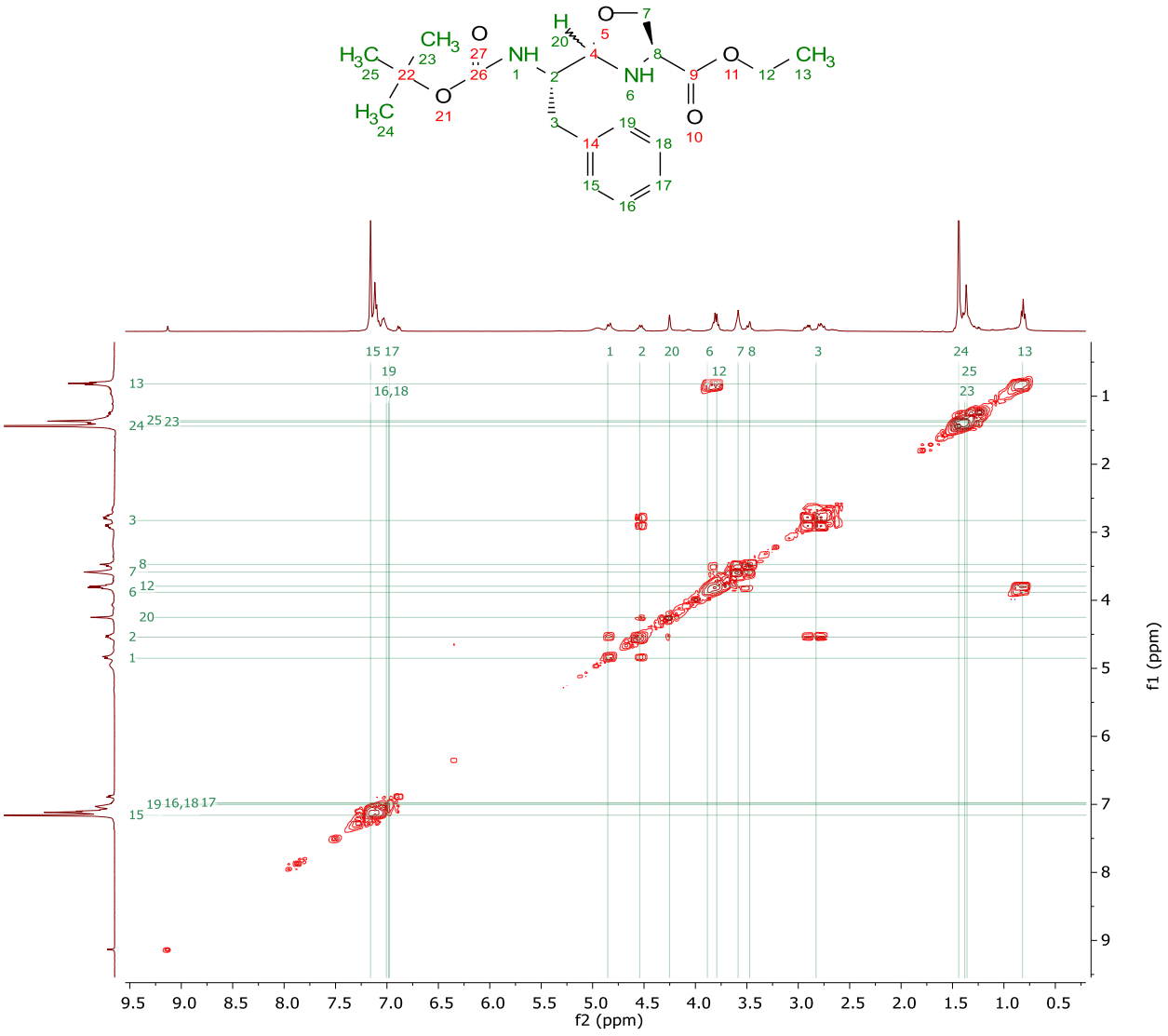
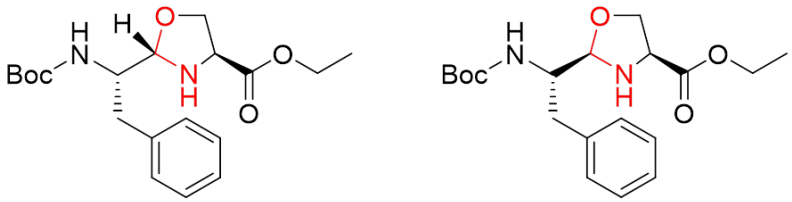


Figure 2-8 2-D COSY NMR spectrum of the N,O-acetal condensation product in benzene-d6 with assignments.

The two possible isomers for N,O-acetal products are: (S,S) *trans* and (R,S) *cis* isomers (**Figure 2-9**). However, only a single CH resonance is observed in benzene at 25 °C (**Figure 2-6**), suggesting one of the diastereomers is favored in the two possible conformations of N,O-acetal products. No Nuclear Overhauser Effect (NOE) observed between acetal proton and the proton

connects to –C-N suggests the conformer might be (S,S) *trans* configuration. It might be due to the anomeric effect, which stabilize the axial conformer, where large substituent on the acetal carbon resides at axial position rather than less hindered equatorial orientation expected from steric considerations.



(S, S) *trans* (R, S) *cis*

Figure 2-9 Two possible conformations of N,O-acetal products: (S,S) trans isomer and (R,S) cis isomer.

2.3.1.2 Anomeric effect observed by experiments in polar solvents

Dimethyl sulfoxide (DMSO) and acetonitrile (MeCN) are chosen as more polar solvents for the N,O-acetal condensation reaction. As anticipated, two CH resonance peaks of acetal protons are observed in both MeCN-d_3 and DMSO-d_6 instead of only one CH resonance peak observed in benzene-d_6. This suggests that the increased dielectric constants of MeCN (37.5) and DMSO (46.7) help stabilize the more polar equatorial conformation, in agreement with the expectation that the polar species is stabilized more in polar solvents. Thus, the equilibrium is driven to shift in favor of the equatorial conformer from the dominant axial conformer in non-polar solvent like benzene. Integrations of two conformers' acetal protons are calculated based on internal standard phenylalanine ring, and their ratio gave rise to diastereomeric ratio (dr), which is 1.7:1 in MeCN and 1.2:1 in DMSO (**Table 2-1**). The decrease of dr with increased solvent dielectric constant

indicates the axial conformational preference of (S,S) *trans* N,O-acetal conformer in benzene reduces as the solvent polarity increases, consistent with previously reported other anomeric systems.

solvent	dielectric constant	dr
Benzene	2.3	>20:1
MeCN	37.5	1.7:1
DMSO	46.7	1.2:1

Table 2-1 Solvent dependence of conformational preference of N,O-acetal dimerization product. Diastereomeric ratio (dr) increases with dielectric constant of solvents.

2.3.1.3 Kinetics and thermodynamics of N,O-acetal in different solvents

The kinetics and thermodynamics of the reaction is modulated by the concentration of the substrates and the nature of the solvents. To compare the solvent effect of N,O-acetal condensation in benzene, MeCN and DMSO, concentration of each substrate remains the same as 45 mM in the three samples. As shown in **Figure 2-10**, the condensation reaction demonstrates the largest kinetic rate and equilibrium constant in benzene. Initial water concentrations in solvents confirmed by water resonance integration in benzene, MeCN and DMSO are 0.9 mM, 23.4 mM and 60.3 mM, respectively, which explains the difference seen in equilibrium constants. The equilibrium constants of N,O-acetal condensation estimated in benzene, MeCN and DMSO are 20.4, 16.1, 2.0, respectively (calculation is based on the **Scheme 2-1, bottom**).

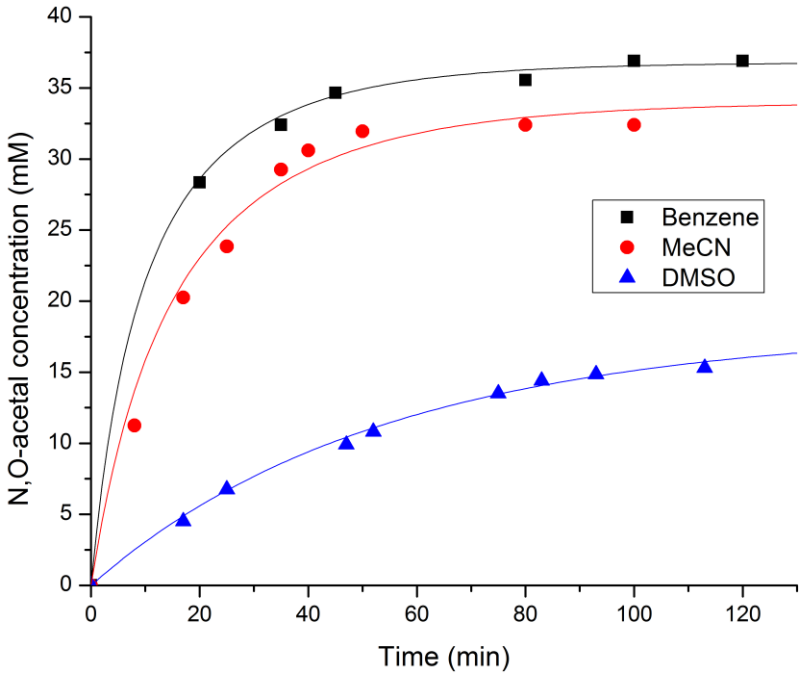


Figure 2-10 N,O-acetal condensation in benzene, MeCN and DMSO with different reaction rate and equilibrium. Equilibrium constants are calculated based on the **Scheme 2-1, bottom**.

The high yield (over 82% conversion) and great diastereoselectivity of this N,O-acetal condensation in benzene suggests a useful method for synthesizing enantiomerically pure compounds with stereogenic acetal centers. N,O-acetal linkage demonstrates greater stability than imine, providing itself as a potential linkage for constructing dynamic chemical networks. Although the reduced diastereoselectivity in more polar solvents (MeCN) might cause complexity of analysis in DCNs with N,O-acetal linked oligomers, it renders opportunity for selecting homochiral polymers based on self-assemblies. Next, the N,N-acetal condensation is examined for its potential as reversible linkage to construct DCNs.

2.3.2 N,N-acetal linkage

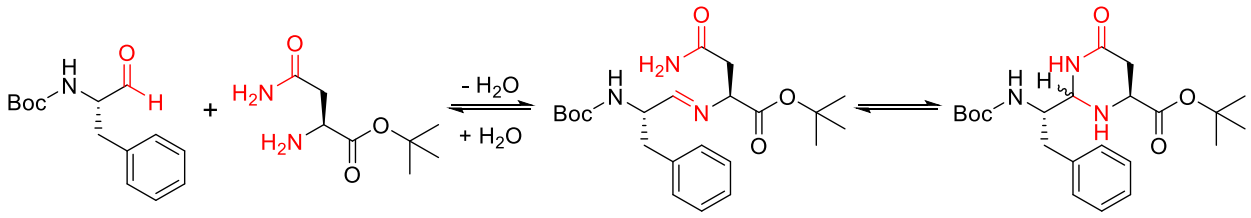


Figure 2-11 Imine and N,N-acetal condensation products generated from reactants N-Boc-Phe-CHO and L-Asn-tert-butyl ester.

Mono-functionalized monomers *N*-Boc-Phe-CHO and L-Asn-tert-butyl ester are selected for studying N,N-acetal linkage in MeCN. As shown in **Figure 2-11**, imine intermediate is formed by nucleophilic attack of amine to aldehyde, and then is trapped by Asn side chain to generate the N,N-acetal linked product. The assignments of protons in N,N-acetal are shown in ^{1}H NMR spectrum (**Figure 2-12**), which are confirmed by 2D COSY spectrum in **Figure 2-13**. The acetal proton labeled as '25' is at δ 5.54 (d, 1H), which shows coupling with α-proton derived from Phe residue (proton '2' at 4.37 ppm).

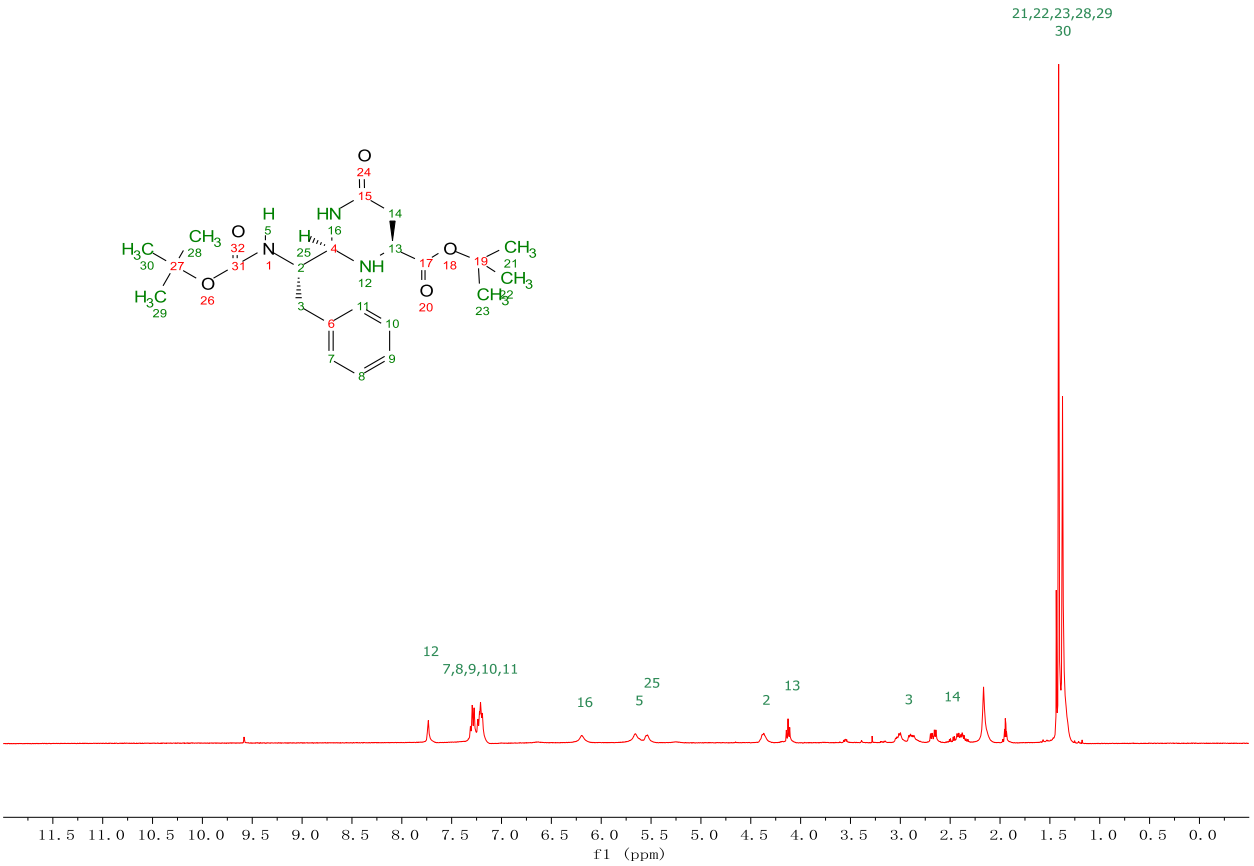


Figure 2-12 1H NMR spectrum of N,N-acetal product in CD3CN-d3 with assignments , confirmed in COSY spectrum in **Figure 2-18**. The acetal proton is at δ 5.54 (d, 1H), labeled as '25' in the spectrum.

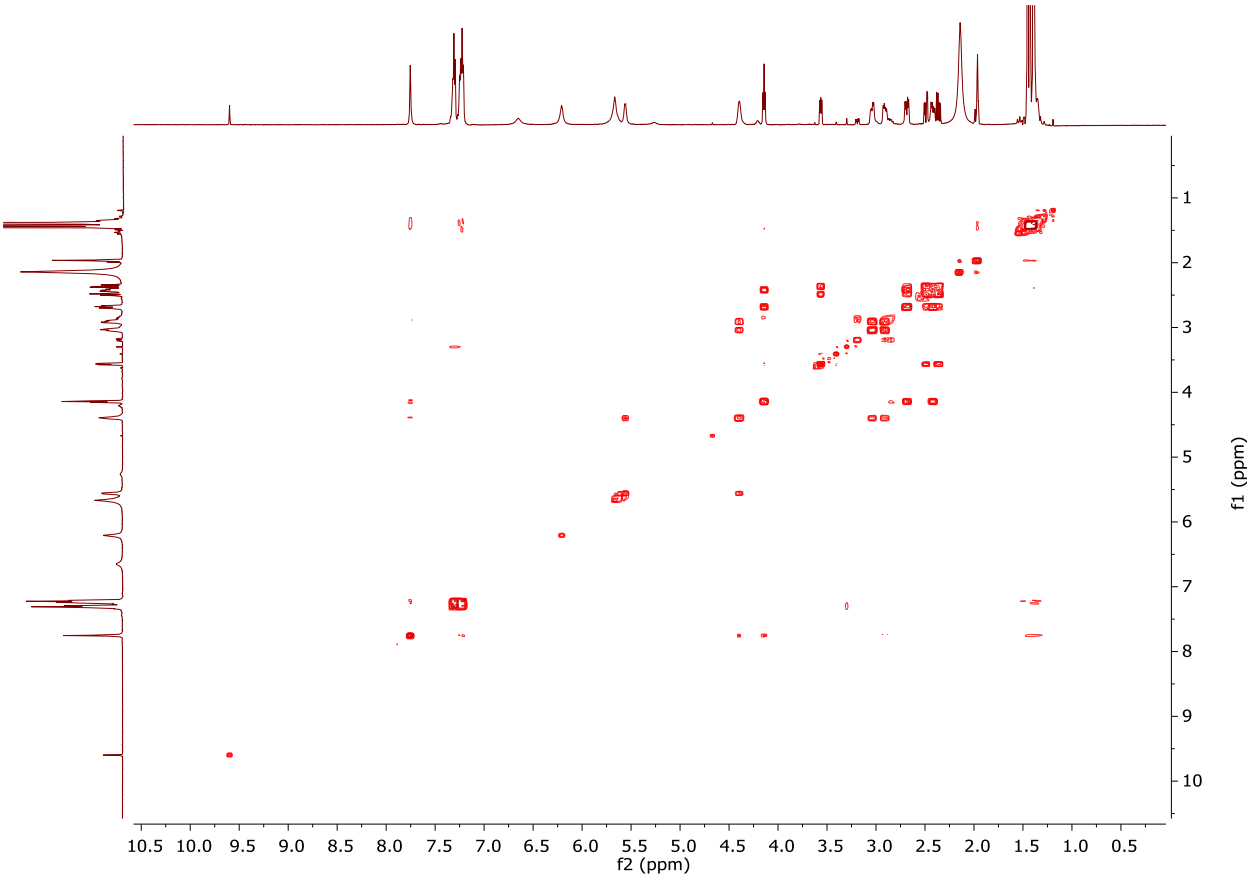


Figure 2-13 N,N-acetal 2D COSY spectrum clearly indicates the coupling between acetal proton (5.54 ppm) and α-proton (4.37 ppm) of Phe residue

To define this ligation reaction and follow its kinetics, substrates are titrated *ca* 1:1 ratio in MeCN-d_3 and followed by ^{1}H NMR. The choice of MeCN as the solvent is to keep consistence of future analysis of dynamic network system of bis-functionalized monomer. Assignments for the starting materials and the acetal product are shown in the top spectrum (red) taken at 10 minutes (**Figure 2-14**). Proton '1' is the aldehyde proton from the starting material *N*-Boc-Phe-CHO; proton '2' and '3'are the α-protons from *N*-Boc-Phe-CHO and L-Asn-tert-butyl ester.

Proton ‘4’ is the intermediate imine proton, and protons ‘5’ to ‘10’ are from the N,N-acetal product with ‘6’ being the acetal proton.

At 10 minutes (red spectrum), both imine intermediate and acetal product are observed. Gradual changes of the resonances are followed in a series of spectra monitored from 10 minutes to 56 minutes in **Figure 2-14**. Up or down arrows indicate the increase or decrease of the specific proton integration, respectively. The curved arrows (from proton ‘2’ to ‘5’ and proton ‘3’ to ‘8’) demonstrate that protons ‘5’ and ‘8’ in the product are converted from the protons ‘2’ and ‘3’ in starting reagents, respectively. During the time course, the increase of N,N-acetal proton doublet resonance at 5.54 ppm (proton ‘6’) is concomitant with the decrease of aldehyde resonance at 9.58 ppm (proton ‘1’) and imine proton ‘4’ and the increase of protons ‘5’ and ‘8’ is concomitant with the decrease of protons ‘2’ and ‘3’. At 56 minutes, only trace amount of proton ‘1’, ‘2’, ‘3’and ‘4’ is detected.

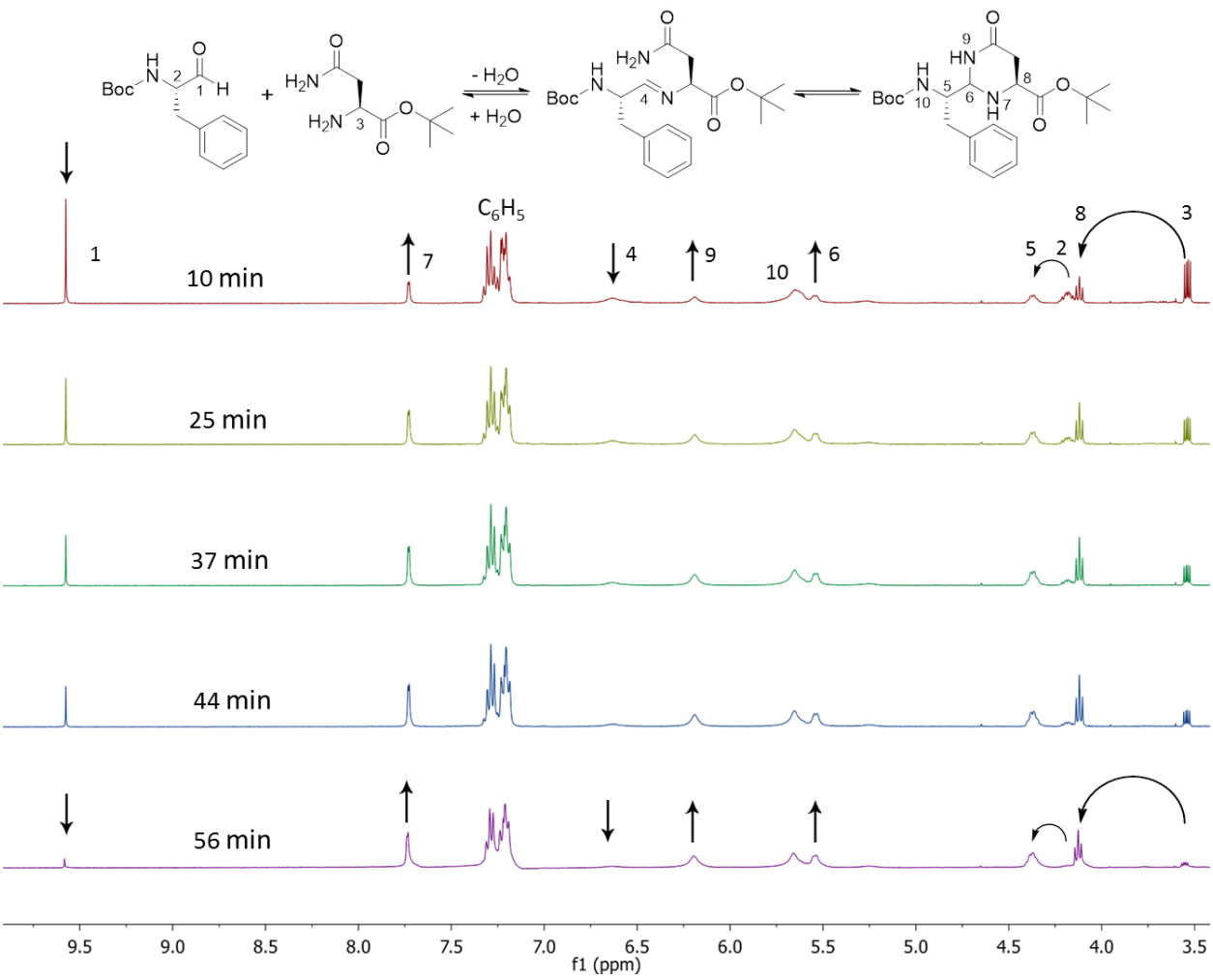


Figure 2-14 1H NMR spectra for monitoring the N, N-acetal condensation in MeCN. Resonance assignments for the starting materials, imine intermediate and the acetal product are shown in the top spectrum (red) taken at 10 minutes. Up or down arrows indicate the increase or decrease of the specific proton integration in the time course, respectively. The curved arrows demonstrate protons in product converted from the starting reagents.

The integrations of imine and acetal protons at each time point are then converted to concentrations and fitted by the kinetic model (**Figure 2-15**). Imine (red line) as the transient

intermediate reaches its highest concentration before 10 minutes and then gradually converts to the acetal product (green line). Within an hour, the condensation reaction reaches equilibrium and most of the reactants convert to acetal. The forward rate constant k_1 and backward rate constant k_2 for imine condensation are 1.52E-02 and 3.53E-04, respectively, and the forward rate constant k_3 and backward rate constant k_4 for acetal formation are 6.65E-02 and 5.70E-04, respectively. The equilibrium constants for each step are 43.1 and 116.7, indicating the favored condensation in this condition.

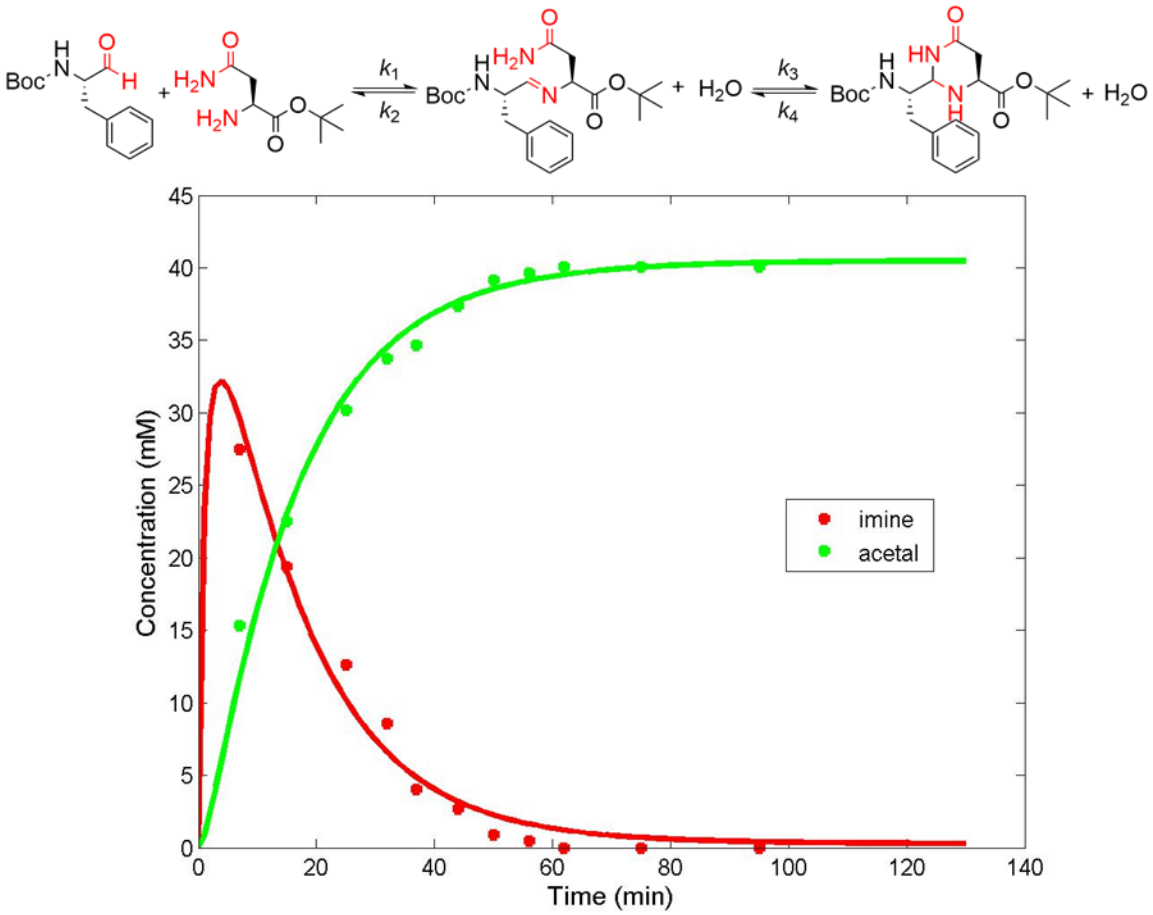


Figure 2-15 Kinetics fit of imine intermediate and N,N-acetal. Forward rate constant k_1 and backward rate constant k_2 for imine condensation are 1.52E-02 and 3.53E-04, respectively;

forward rate constant k_3 and backward rate constant k_4 for acetal formation are 6.65E-02 and 5.70E-04, respectively.

2.3.3 Diastereoselectivity of N,N-acetal in MeCN

Only a single acetal proton of N,N-acetal (labeled as proton '25') is detected in MeCN (**Figure 2-12, 13, 14**), indicating the great diastereoselectivity of this reaction in MeCN. Nuclear Overhauser Effect (NOE), the transfer of nuclear spin polarization from one nuclear spin population to another via cross-relaxation, is used for identifying the preferred conformation of N,N-acetal. Protons that are in close proximity in space (usually within 5 Å) can give a NOE signal. As the α-proton from L-Asn substrate has the absolute 'S' configuration, the stereochemistry of the new chiral center can be determined by comparing the intensity of NOE signal between the acetal proton and the α-proton from Asn substrate located on the ring. Upon irradiation of acetal proton (labeled as '25' in **Figure 2-16**), over 3% NOE enhancement of α-proton (labeled as '13') is observed. Therefore, the *cis* diastereomer of N,N-acetal is confirmed as (2R,6S)-4-pryrimidinone with the two substituents on the same face of the ring.

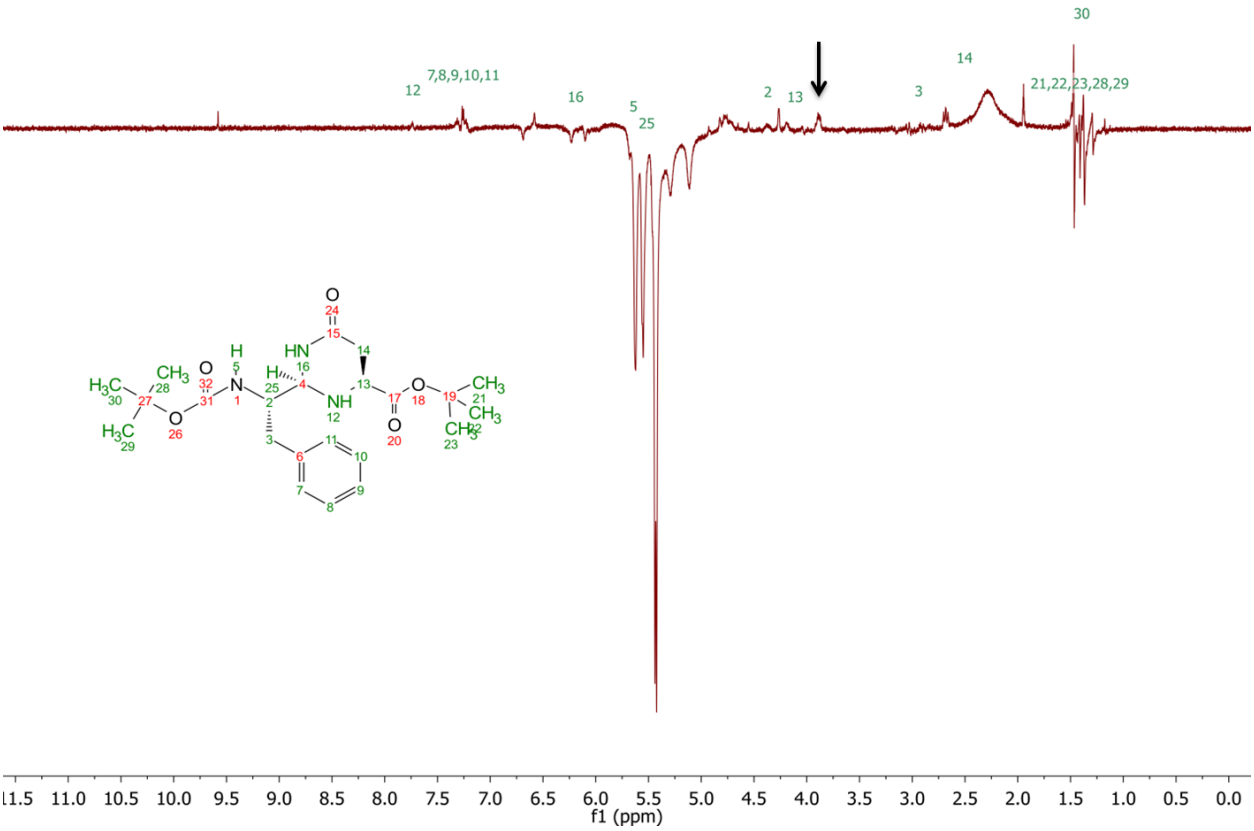


Figure 2-16 NOE for N,N-acetal product. *cis* (2R,6S)-4-pyrimidinone with the two substituents on the same face of the ring is confirmed by NOE observed between the acetal proton (labeled as '25') and α-proton of Asn substrate (labeled as '13') with 3% enhancement. (Enhancement for proton 2 is less than 1%; enhancement for proton 3 is 3%. The broad resonance at 2.25ppm is from water residue in the sample.)

The steric effect induced by amide group in the N,N-acetal ring structure would probably control the stereochemistry as *cis* conformation (**Figure 2-17, middle**) to reduce the interference between the two large substituents, which has been seen in similar O,O-acetal (2,6-disubstituted

1,3-dioxin-4-ones) in $CHCl_3$ (**Figure 2-17, right**) (D Seebach, Zimmermann, Gysel, Ziegler, & Ha, 1988).

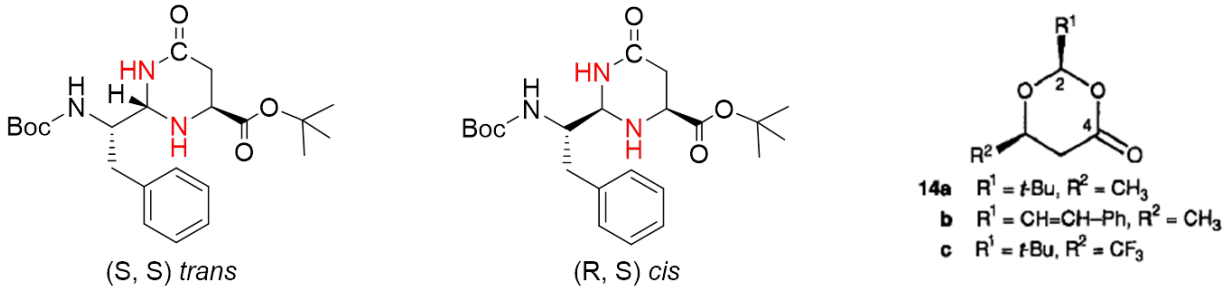


Figure 2-17 Two conformations of N,N-acetal products: (S,S) *trans* isomer (**left**) and (R,S) *cis* isomer (**middle**). O,O-acetal *cis* conformation of varied sets of substituents (**right**) (D Seebach et al., 1988).

2.3.4 Kinetics and Thermodynamic Comparison of N,O-; N,S-; N,N-acetal Condensations in MeCN

Besides N,O- and N,N-acetal condensation, the N,S-acetal is also examined and the three condensations are compared by plotting their concentrations calculated from resonance integration vs. time in **Figure 2-18**. N,S- and N,N- acetal condensations are kinetically faster compared to N,O-acetal condensation due to stronger nucleophilicity of thiol group in Cys and nitrogen in Asn side chain than hydroxyl group in Ser side chain. Larger equilibrium constants of N,S- and N,N- acetal condensations compared to N,O-acetal can be explained by the increased stability of sulfur containing heterocycles due to reduced steric strain (Valters & Flitsch, 1985); and more stable N,N-acetal six-membered ring with amide group than the N,O five-membered ring.

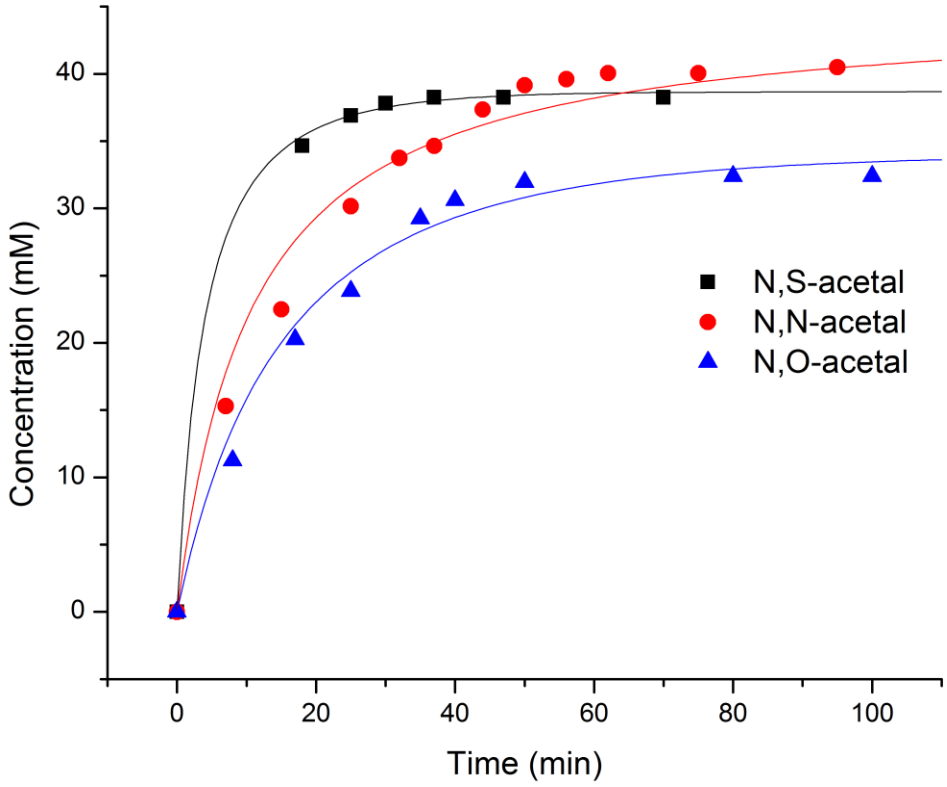


Figure 2-18 N,O-acetal, N,S-acetal and N,N-acetal condensation in MeCN with different reaction rate and equilibrium.

2.3.5 Temperature effect on diastereoselectivity of N,O-; N,S-; N,N-acetals

These condensation reactions also respond to other environmental stimuli, such as temperature change. All of the diastereomeric ratios (dr) for N,O-; N,S-; N,N-acetals are the largest at relatively low temperature (25°C) with one diastereomer being dominant. However, at elevated temperatures, the population of the minor conformers increases which is confirmed by the appearance of the new acetal proton. Higher temperatures help the less favored conformers overcome the energy barriers, which results in the decrease of diastereoselectivity (**Figure 2-19, 20, 21**). At 55°C, the dr of N,O-acetal conformers is 1.16:1. At 65°C, the dr of N,S-acetal

conformers and N,N-acetal conformers are 1.67 and 1.52, respectively. N,S-acetal and N,N-acetal demonstrate the larger diastereomeric ratios than N,O-acetal may be due to the more rigid ring structures of the former two causing greater free energy difference of two diastereomers.

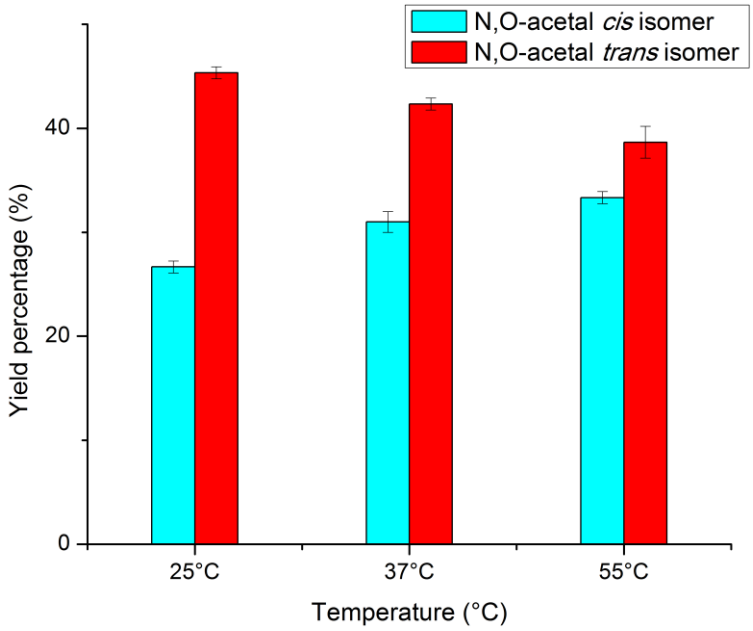


Figure 2-19 Temperature dependence of N,O-acetal in MeCN

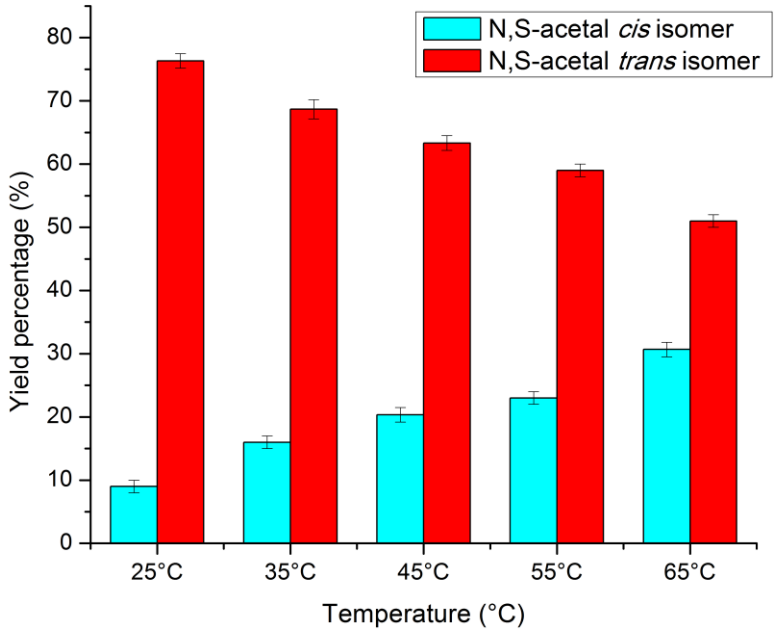


Figure 2-20 Temperature dependence of N,S-acetal in MeCN

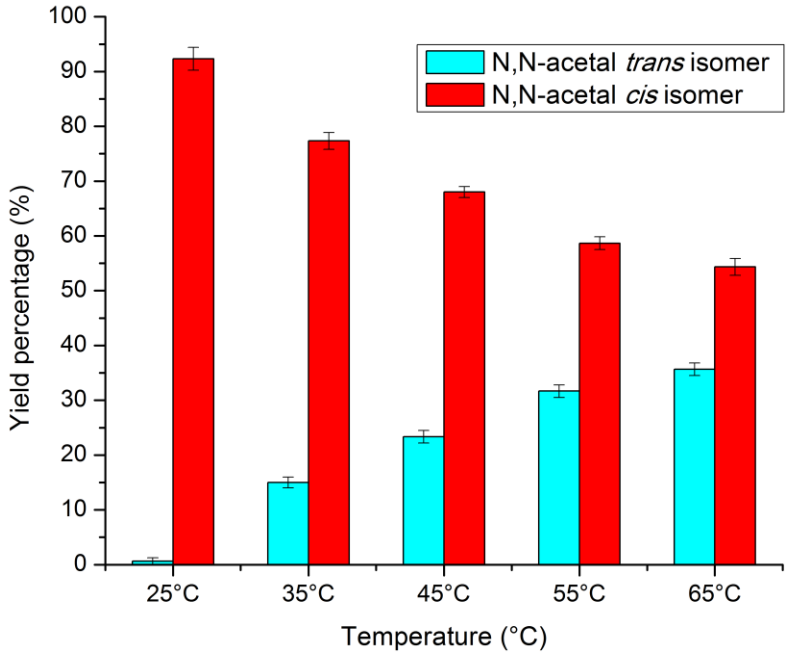


Figure 2-21 Temperature dependence of N,N-acetal in MeCN

2.4 Discussion

The three acetal condensations (N,O-; N,N- and N,S-acetals) in a series of solvents and at different temperatures have been demonstrate in this chapter. Not only do the condensation reactions respond to varied solvents, they also respond to other environmental stimuli, such as temperature change.

The high yield (over 82% conversion) and great diastereoselectivity for this N,O-acetal condensation in benzene suggests a useful method for synthesizing enantiomerically pure compounds with stereogenic acetal centers. N,O-acetal demonstrates greater stability than imine, providing itself as a potential linkage for constructing dynamic chemical networks. The (S,S) *trans* conformer of N,O-acetal is dominant in benzene due to the anomeric effect and switching to more polar solvents such as MeCN results in the appearance and enhancement of the other (R,S) *cis* conformer. Although the reduced diastereoselectivity in more polar solvents (MeCN) might cause complexity of analysis in DCNs with N,O-acetal linked oligomers, it renders opportunity for selecting homochiral higher order structures through self-assembly.

Polarity change of the reaction medium also changes the kinetics of condensation reactions, which might be due to a change in the microscopic solute-solvent interactions. The polar solvents stabilize the more polar *trans* conformer with equatorial substituent, the major *cis* conformer with less polar solvated transition state would not be favored, resulting in lower kinetics in polar solvents.

The structural investigation of the acetals in this chapter extends our understating of reversible linkages and set foundations for utilizing these linkages for constructing novel dynamic chemical

networks in aqueous media. Due to the great diastereoselectivity of N,N-acetal in acetonitrile, this linkage is utilized to build NF-CHO and NFF-CHO networks.

Chapter 3: NF-CHO Dynamic Chemical Network

3.1 Introduction

The ability to produce and replicate with high fidelity is fundamental for biopolymer evolution, yet how production and replication of biopolymers emerged before the cellular machinery necessary for life remains unknown. Extant biopolymers rely on complex catalysts to construct specific materials through template-directed polymerization. However, in the absence of sophisticated catalysts, the task of constructing biopolymers seems impossible. This chapter describes a novel dynamic chemical network (DCN) capable of producing specific N, N-acetal linked peptide oligomers, which satisfy the self-organizing/ self-replication requirements necessary for chemical evolution and demonstrate the selection of specific assemblies through prion-like selection processes.

Among the extant biopolymers, acetal linkages are not present in proteins, but sugars are polymerized through acetal condensation and the nucleic acid bases are affixed to sugars through acetals. Simple oxidation state change of an α-amino acid to an α-amino aldehyde framework would allow for reversible covalent coupling of monomers through imine condensation requiring only protons for catalysis. This approach is inspired by the simple change of nucleoside termini to amino aldehydes, lowering the activation barrier for ligation and making polymerization reactions energetically accessible to nucleoside backbones (Jay T. Goodwin & Lynn, 1992; X. Li et al., 2011; Xiaoyu Li et al., 2002; Z.-Y. J. Zhan & D. G. Lynn, 1997).

In Chapter 2, I analyzed different acetal linkages, including N, N-, (oxazolidine) N, O- and N, S-acetals as reversible linkages for polymerization in a peptide scaffold. NMR analyses of simple model dimerization reactions confirmed diastereoselective control by the asparagine α-carbon of

the 6-membered ring N, N-acetal, which gave rise to the cis arrangement of the (2R, 6S)-4-pryrimidinone ring system forming at ambient temperature (**Figure 2- 21**). As phenylalanine (Phe/ F) has the highest propensity among any of the natural amino acids to promote self-assembly (Frederix et al., 2011), I focused on the bis-functional monomer Asn-Phe-CHO (NF-CHO) for both of the potential to polymerize and assemble into higher-order structures.

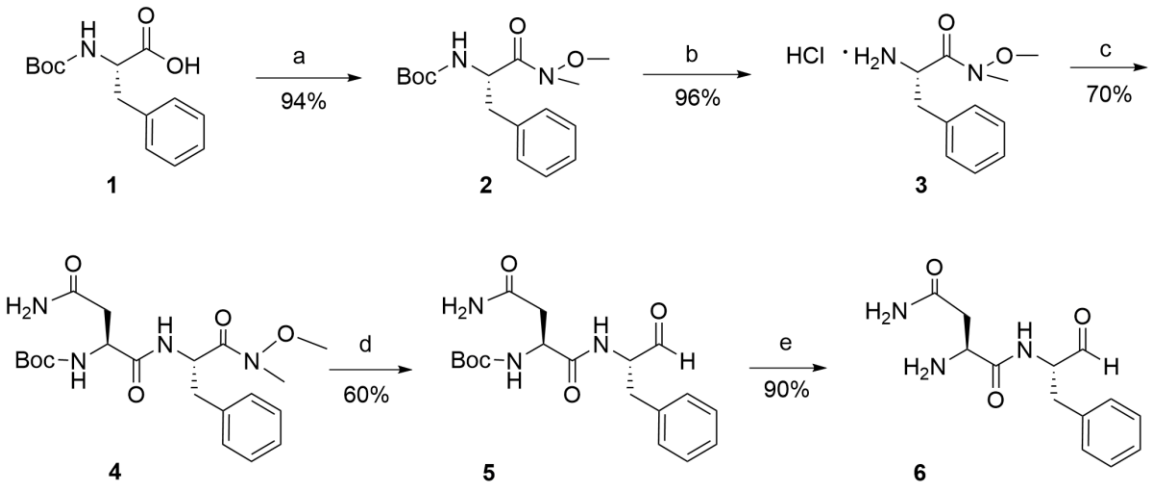
3.2 Materials and Methods

3.2.1 Materials

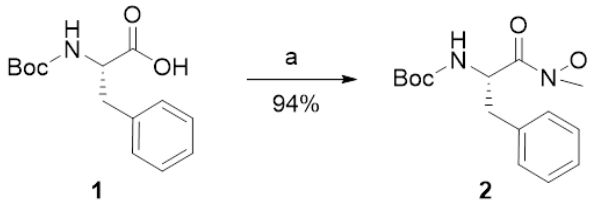
All commercially available chemicals are purchased from Sigma-Aldrich, AnaSpec and Nova Biochem. Anhydrous solvents are either dried over molecular sieve (4 Å) that had been pre-treated overnight at 300 ºC or purchased from EMD or Acros organics. HPLC grade acetonitrile and water are obtained from Sigma-Aldrich and/or Fisher Scientific. TLC plates are purchased from EMD (silica gel 60 F_{254}). Fmoc-amino acids, resins and solid phase peptide synthesizer reagents are purchased from AnaSpec. Distilled deionized water for sample preparation is obtained from EMD chemicals Inc.

Synthesis of Network Building Blocks

Synthesis of NF-CHO



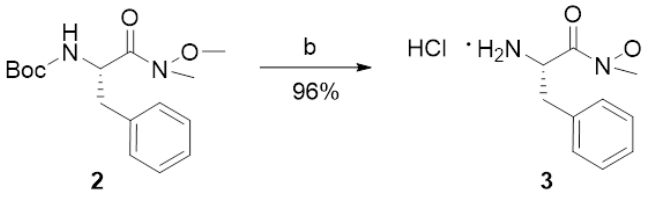
a. Preparation of the Boc-Phe-weinreb amide **2**



N-Boc-L-Phe (2.65 g, 10 mmol) is dissolved in anhydrous dichloromethane (DCM), and then 1,1-carbonyldiimidazole (1.78 g, 11 mmol) is added. The resultant mixture is stirred for an hour at room temperature. Subsequently, N,O-dimethylhydroyxlamine hydrochloride (1.07 g, 11 mmol) is added and the reaction mixture (light yellow suspension) is stirred and allowed to proceed overnight. Then the solvent is removed by vacuum at 25 °C and the resulting residue is

purified by extraction, which is dissolved in ethyl acetate (100ml) and ished successively with 2×100ml 1M HCl, 2×100ml saturated aqueous $NaHCO_3$ and 2×100ml brine. Finally the organic layer is dried with anhydrous Na_2SO_4 and evaporation of the solvent yielded product **2**. Product is visualized by UV on TLC plate (hexane/ ethyl acetate (1/1) as solvent system, $R_f \sim 0.46$). ^1H NMR (600 MHz, Chloroform-d) δ 7.30 –7.14 (m, 5H, C_6H_5), 5.17 (d, J = 8.8 Hz, 1H, CH (α)), 4.94 (q, J = 7.1 Hz, 1H, NH), 3.65 (s, 3H, O-CH_3), 3.16 (s, 3H, N-CH_3), 3.05/2.87 (dd, 2H, CH_2(β)), 1.38 (s, 9H, $C(CH_3)_3$). MS (ESI) m/z $(M+H)^+$:309.1803 Calculated $(M+H)^+$: 308.1808 ($C_{16}H_{25}N_2O_4$)

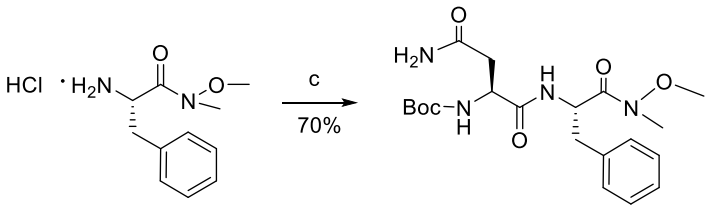
b. Removal of the t-butoxycarbonyl protecting group of **2**



2 **3**

The white solid **2** (2.90 g, 9.4 mmol) prepared above is dissolved in 10 ml ~ 20 ml 4M HCl in dioxane (purchased from Aldrich), and then the resultant mixture is stirred at room temperature for 0.5hr to 1hr at room temperature. Solvent is removed with vacuum at room temperature, leaving the product (**3**) of the hydrochloride salt as thick waxy oil. ^1H NMR (600 MHz, Chloroform-*d*) δ 8.83 – 8.29 (s, 3H, NH_3^+), 7.47 – 7.05 (m, 5H, C_6H_5), 4.74 (q, *J* = 5.9 Hz, 1H, CH (α)), 3.62 (s, 3H, O-CH_3), 3.40/3.28 (m, 2H, CH_2(β)), 3.10 – 3.07 (s, 3H, N-CH_3). MS (ESI) m/z $(M+H)^+$: 209.1284 Calculated $(M+H)^+$: 209.1285 $C_{11}H_{17}N_2O_2$

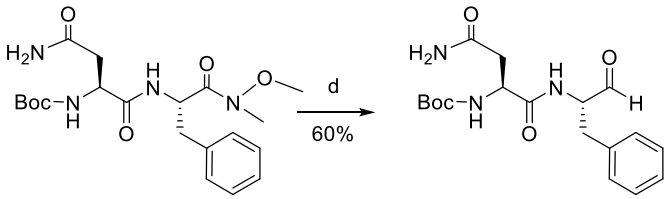
c. Coupling Boc-Asn to **3**



3 **4**

The waxy oil **3** (2.20 g, 9 mmol) is dissolved in anhydrous DCM (50-80 ml) before Boc-L-Asn (2.30 g, 9.9 mmol) and a two-fold excess of trimethylamine (2.5 ml, 18 mmol) are added. The reaction mixture is cooled to 0°C in an ice bath prior to the addition of 1-Ethyl-3(3-dimethylaminopropyl) carbodiimide (EDC) (1.90 g, 9.9 mmol), and the reaction mixture is stirred at 0°C under N_2 for 2hrs and then stirred at room temperature for overnight. The solvent is removed in vacuo at room temperature and the resulting residue dissolved in ~100ml ethyl acetate and ished successively with 2×100ml of 1M HCl, 2× 100ml of 4% $NaHCO_3$ and 2×100ml of brine. The excess of reagents of Boc-L-Asn and EDC are ished away by aqueous layer.The organic layer is dried with anhydrous sodium sulfate, decanted, and the solvent removed in vacuo to give white solid 4. UV visualization on florescent TLC plates (ethyl acetate eluent, R_f ~ 0.19) is possible. ^{1}H NMR (600 MHz, Chloroform-d) δ 7.42 (d, J = 8.3 Hz, 1H), 7.32 – 7.16 (m, 5H), 5.96 (d, J = 9.4 Hz, 2H, NH_2), 5.49 – 5.45 (m, 1H, NH), 5.16 (q, J = 7.2 Hz, 1H, CH(α)), 4.47 (d, J = 8.5 Hz, 1H, CH(α)), 3.65 (s, 3H, O-CH_3), 3.15 (s, 3H, N-CH_3), 3.07 (dd, J = 13.6, 6.0 Hz, 1H), 2.97 – 2.83 (m, 2H, CH_2(β)), 2.55 (dd, J = 15.5, 6.1 Hz, 1H), 1.44 (s, 9H, $C(CH_3)_3$). MS (ESI) m/z $(M+Na)^+$: 445.2054, Calculated $(M+Na)^+$: 445.2057 ($C_{11}H_{16}N_2O_2Na$)

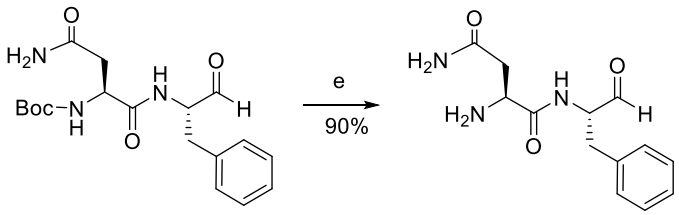
d. Reduction of the Boc-Asn-Phe-weinreb amide **4**



4 **5**

White solid **4** (2.66 g, 6.3 mmol) is dissolved in 60~80 ml anhydrous THF under N_2 and cooled to -78 °C in a dry ice and acetone bath before 1.5 equivalent of lithium aluminum anhydride (1.0 M) in THF (9.5 ml, 9.5 mmol) is added dropwise via syringe. The temperature is raised to 0 °C with an ice bath for 30min and then cooled again to -78°C before the reaction is quenched with 10 ml of 1M aqueous solution of potassium bisulfate ($KHSO_4$). The mixture is allowed to warm to room temperature and the aluminate residue extracted several times with ethyl acetate. These combined extracts are ished with 2×50ml $NaHCO_3$ and 2×50ml brine before the organic phase is dried with anhydrous Na_2SO_4, decanted, and taken to dryness in vacuo at <20 °C to provide waxy solid 5. ^{1}H NMR (600 MHz, THF-d8) δ 9.46 (s, 1H, CHO), 7.94 (d, J = 7.2 Hz, 1H, NH), 7.28 – 7.09 (m, 5H, C_6H_5), 6.83/ 6.31 (s, s, 2H, NH_2), 6.45 (m, 1H, NH), 4.37 (q, J = 6.7 Hz, 1H, CH(α)), 4.20 (q, J = 6.8 Hz, 1H, CH(α)), 3.14/2.64 (dd, dd, 2H, CH_2(β)), 2.95 (ddd, 2H, CH_2(β)), 1.39 (s, 9H, $C(CH_3)_3$). MS (ESI) m/z $(M+Na)^+$:386.1698, Calculated $(M+Na)^+$: 386.1670 ($C_{18}H_{25}N_3O_5Na$)

e. Deprotection of Boc group of **5** to give rise to the free monomer **6**



5 **6**

The waxy solid **5** (1.37 g, 3.78 mmol) is dissolved in 10 ml trifluoroacetic acid under N_2, stirred at room temperature for 1hr, and the solvent removed in vacuo to give thick waxy solid **6**. MS (ESI) m/z $(M+H)^+$: 264.1344 Calculated $(M+H)^+$: 264.1343 ($C_{13}H_{18}N_3O_3$)

3.2.2 NMR Analysis

^{1}H NMR data are recorded on an INOVA 600 or an INOVA 400 NMR spectrometer (equipped with Bore Oxford super conducting magnet). The solvents Chloroform-d and THF-d8 are purchased from Cambridge Isotope Laboratories, Inc. The choice of the solvent is determined by the solubility of the compounds.

3.2.3 Dynamic Peptide Network Preparation

After final deprotection of the precursor of the network, Boc-NF-CHO by trifluoroacetic acid, resulting network building blocks of NH_2-NF-CHO are dried under vacuum at room temperature. Dried NH_2-NF-CHO are stored with Argon protection at -20°C. To prepare the dynamic network, NH_2-NF-CHO are dissolved to a concentration of 8mM in water/acetonitrile (3/2, v/v) at ambient temperature under N_2 protection. All the solvents are flushed with N_2 gas for at least 15 min before usage. Dissolution is assisted by ~2 minutes of continuous vortexing, followed by ~15 minutes of bath sonication until solution became clear. The pH of the solution is adjusted by

titrating aliquots of 10mM NaOH and the final pH value is determined by pH meter (Fisher Scientific Accumet Basic AB15 pH meter).

3.2.4 HPLC and LC-MS Analyses

HPLC analyses are performed on Waters Delta 600 equipped with a photodiode array UV/Vis detector at room temperature using a reversed-phase HPLC column (Kromasil 100-5C18, 4.6 × 250mm). Solvent A: water (0.1 vol % trifluoroacetic acid). Solvent B: acetonitrile (0.1 vol % trifluoroacetic acid). UV absorbance is monitored at 258nm (for Phenyl ring side chain absorption) and 222nm (for amide bond absorption). Flow rate is 1.0 mL/min. Gradient is from 10% acetonitrile to 90% acetonitrile, 2% acetonitrile/ min. LC-MS analyses are performed on Waters Synapt G2 MS/Acquity UPLC system. Positive-ion mass spectra are obtained using electrospray ionization.

Time (mins)	Solvent A	Solvent B	Note
0	90%	10%	Gradient starts
40	10%	90%	Gradient ends
41	0%	100%	Cleaning starts
48	0%	100%	Cleaning ends
49	10%	90%	Re-equilibrium starts
56	10%	90%	Re-equilibrium ends

3.2.5 IMS-MS

Direct total ion current analysis is performed on Synapt G2 High Definition Mass Spectrometry system (Waters Corporation, Manchester, UK), which is a hybrid quadrupole-ion mobility-

orthogonal acceleration time-of-flight instrument, with typical resolving power of 20,000 m/Δm (FWHM) and mass accuracy of 9 ppm at m/z 556.2771. The instrument is operated in positive ion mode with a probe capillary voltage of 3 kV, and a sampling cone voltage of 45 V. The source and desolvation temperatures are set to 120 °C and 250 °C, respectively; and the nitrogen desolvation flow rate is set to 650 L h^{-1}. The mass spectrometer is calibrated across the 50-1200 m/z range using a 0.5 mM sodium formate solution prepared in 90:10 2-propanol:water v/v. Data are mass corrected during acquisition using a leucine encephalin (m/z 556.2728) reference spray (LockSpray) infused at 2 µL min^{-1}. The scan time is set to 1 s. Data acquisition and processing is carried out using MassLynx v4.1 and Drift Scope v2.1 (Waters Corp.) Column: Waters Acquity UPLC BEH C18 column (1.7 µm, 2.1 x 50 mm)

3.2.6 Transmission Electron Microscopy and Electron Diffraction

Aliquots (10 µl) of sample solutions are dropped on TEM grids (200 mesh copper grid covered with a thin carbon film, purchased from Electron Microscopy Sciences) for 3 minutes before excessive solution is blotted with filter paper. Uranyl acetate (10 µl of 5% solution) is added for 3 minutes for negative staining. Extra fluid is blotted with filter paper. The grids are analyzed on a Hitachi H-7500 transmission electron microscope with a LaB6 emission filament at an accelerating voltage of 75 kV.

3.2.7 Microwave Assisted Solid-Phase Peptide Synthesis

Peptides H-NFNF-NH_2 and H-NFNFNF-NH_2 are synthesized on a Liberty CEM Microwave Automated Peptide Synthesizer utilizing FMOC-Rink Amide MBHA Resins purchased from AnaSpec. All FMOC protected amino acids are from Anaspec, and other chemicals from Sigma-Aldrich or Fisher Scientific. Each peptide synthesis is performed at 0.1 mmol using a 45 mL reaction vessel at a scale of 0.1mmol. FMOC-Rink Amide MBHA Resin is initially swollen

using ~7 ml dimethylformamide (DMF) for 15 minutes. FMOC deprotection is achieved by addition of 20% piperdine 0.1M N-Hydroxybenzotriazole (HOBt) in DMF with microwave power set to maintain temperature between 45-55°C for 180 sec, followed by 3X flushing with DMF. Each coupling step is performed using 0.1M FMOC protected amino acid, and activated with 0.1 M 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate (HBTU), and 0.2M N,N- Diisopropylethylamine (DIEA) in DMF. Coupling temperatures are maintained between 75-82°C 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°C) diethyl ether. Precipitated product is centrifuged at 3000 rpm for 15 min, and the pellet is subjected to 3 additional rounds of ishing 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.

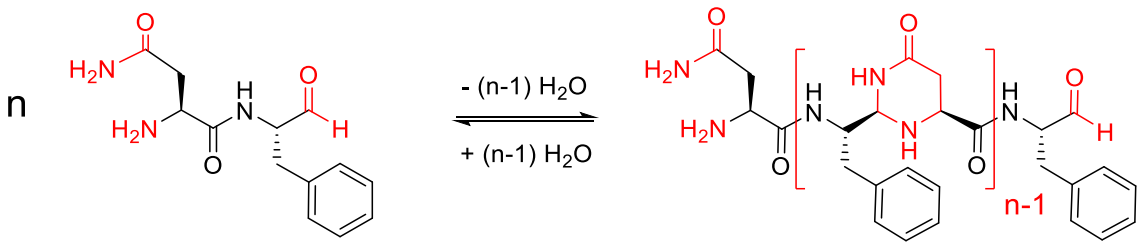
Purified peptides are dissolved in water/acetonitrile (3/2, v/v) with 0.1% TFA for nanofiber assemblies. Dissolution is assisted by ~2 minutes of continuous vortexing, followed by ~15 minutes of sonication until solution became clear.

3.3 Results and Discussions

3.3.1 Construction of NF-CHO Dynamic Chemical Networks (NF-DCN)

The dipeptide aldehyde (NF-CHO) is prepared with standard Boc-peptide synthesis and reduced with $LiAlH_4$ (detailed procedures are shown in Method session). The *tert*-Butyloxycarbonyl protecting (Boc) group of the precursor Boc-NF-CHO is removed by trifluoroacetic acid (TFA) at room temperature. The choice of solvents for the network incubation is determined by the solubility of network members, minimal interference with condensation reactions for oligomerization and potential for promoting self-assemblies in the network. Solvents like methanol and DMSO which can dissolve the network members well are excluded as they will mask the functional groups of network building blocks. 40% acetonitrile in water is selected to incubate NF network, which has shown to increase the homogeneity of peptide assemblies (Lu, Jacob, Thiyagarajan, Conticello, & Lynn, 2003).

The bis-functional monomer NF-CHO (8 mM) is incubated in water/acetonitrile (3/2, v/v) with a resulting pH measured at 4 for generating a network of oligomers with N,N-acetal linkages. As shown in **Scheme 3-1**, the monomers go through condensation reactions in aqueous solution and the intermediate imine linkages are trapped and stabilized as N, N-acetals in oligomers.



Scheme 3-1 Chemical structures of the *NF*-CHO dipeptide building blocks and resulting oligomers in the dynamic chemical network

3.3.2 pH Dependence of NF-DCN

Based on pH dependence studies on the model system discussed in Chapter 2, it is known that both the acetal ring opening and imine hydrolysis rates are the lowest near neutral pHs (Fife & Hutchins, 1980; Giuseppone & Lehn, 2006), therefore, mobile solvents CH_3CN/ H_2O with a quasi-neutral pH range from 5.5 to 6 in reverse-phase HPLC analysis is selected for the NF network analysis to slow down the hydrolysis and exchanging rate and trap the equilibrating mixtures. As shown in **Figure 3-1 A**, the network members are well resolved in chromatogram with good Gaussian peak shape, indicating the sufficient slowing down of the hydrolysis and exchanging rate among those members. This NF-DCN mixture of oligomers with N, N-acetal linkages demonstrates a distribution of monomer, linear dimer, linear trimer, linear tetramer and their cyclic counterparts confirmed by mass spectrometry. Each network member isolated at quasi neutral pH is stable for more than 48 hours. **Figure 3-1 B** demonstrates the isolation of cyclic dimer which remains stable at 2 days. To further test if the NF network is reversible and truly dynamic, the isolated cyclic dimer is incubated back to water/acetonitrile (3/2, v/v), pH 4 for 2 days, the resulting chromatograph **Figure 3-1 C** shows a new network regenerated at 2 days. No epimerization of the aldehyde α-carbon detected under these conditions.

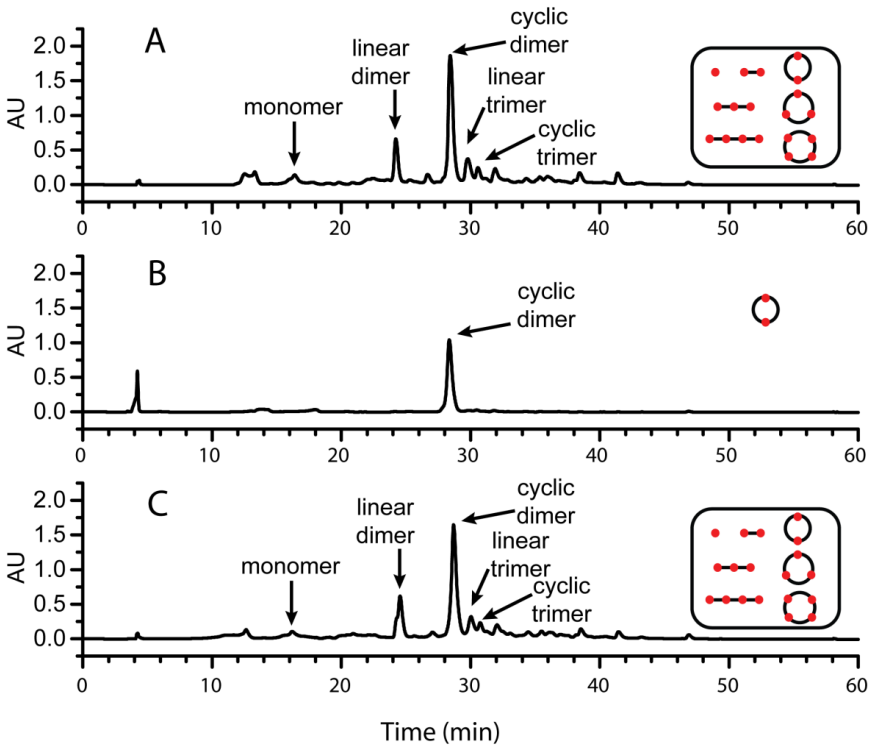


Figure 3-1 High performance liquid chromatography (HPLC) analysis of the (**A**) dynamic peptide network after 2 days in water/acetonitrile (3/2, v/v) at pH 4, (**B**) the cyclic dimer collected under quasi-neutral pH is stable for > 2 days, and (**C**) the new network generated from the cyclic dimer at pH 4. The mobile phase for all HPLC analysis is water/ acetonitrile, flow rate is 1.0 ml/ min with gradient from 10% acetonitrile to 90% acetonitrile for 40 minutes monitored at 222 nm.

3.3.3 Temperature Dependence of NF-DCN

As introduced in Chapter 1, thermodynamic control in dynamic networks implies that changing the experimental conditions can induce changes in network composition; that is, network will respond to external influences. Besides NF-DCN's responsiveness to pH shown in the last section, this section demonstrates the network members' dependence on temperature changes.

When the NF network is heated at 40°C water bath for 3 hours and analyzed by HPLC, the splitting of peaks for each network member suggests the isomerization of the species (**Figure 3-3 B**). When the network is cooled back at room temperature for 2 hours, analysis of the network demonstrate that previous splitted peaks of network members resulted from 40°C changed back to single peaks at room temperature (**Figure 3-3 C**), indicating the restoring of diastereoselectivity in the network to the original one (**Figure 3-3 A**).

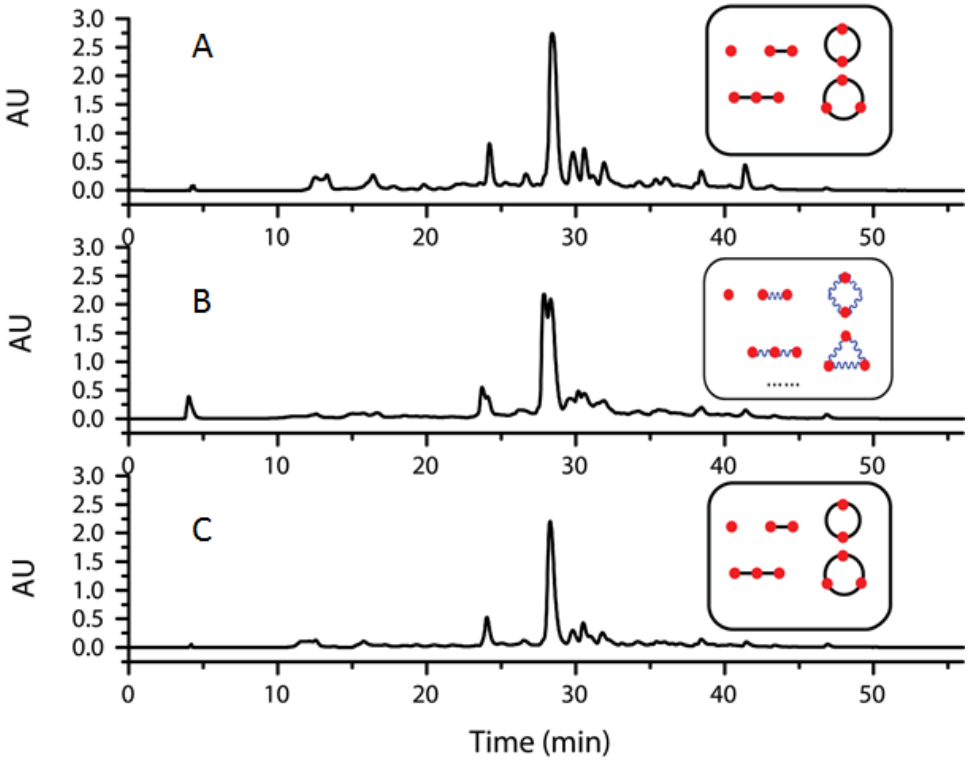


Figure 3-2 Temperature dependence of NF-DCN . (A) distribution of NF- DCN at 2 days; (B) heating to 40°C results in isomerization of network members; (C) cooling back to room temperature drives the network to restore its diastereoselectivity.

The temperature dependence of diastereoselectivity in the NF network is in accordance with N, N-acetal model system described in Chapter 2. In both cases, only one kind of diastereomers is observed at room temperature but two diastereomers at increased temperatures. The stereochemistry of the newly generated chiral center in N, N-acetal ring is greatly influenced by temperature confirmed by NMR experiment shown in Chapter 2 (**Figure 2-25**). At ambient temperature the *cis* isomer with the (2R, 6S)-4-pryrimidinone ring is the only stereoisomer observed (confirmed by NOE experiment), while the appearance of the other *trans* isomer with (2S, 6S)-4-pryrimidinone ring is detected at raised temperature, which is confirmed by the appearance of the new acetal proton. No new α-proton resonance observed indicates no epimerization occurred at the α-carbon position.

Based on the NMR experimental data of N, N-acetal model system and HPLC analyses of NF network, we reason that the network members with *cis* arrangement of acetal ring are the only observed isomers at room temperature, as the thermodynamic products. At higher temperature, the population of the network members with *trans* acetal ring increases as elevated temperature help the less favored conformer overcome the activation barrier. When the temperature drops to room temperature, the thermodynamic control restores the dominance of *cis* diastereomers.

3.3.4 NF-DCN Aging and Kinetic Modeling

The aging of NF-CHO dynamic network is followed by HPLC analysis. To better understand the kinetics in the NF network, the concentrations of network members quantified by peak integration on the HPLC chromatogram are fitted to the simplified kinetic model shown in **Figure 3-4**, including acetal formation rate constant *k*, cyclization propensity parameter *B*, and

equilibrium constants *K* assumed to be equivalent for the progressive addition of each monomer to the growing macromolecular species. Oligomerization and cyclization reactions in the network are outlined in **Eq 1**and **Eq 2**, respectively.

$$L_i + L_j \underset{k/K}{\overset{k}{\rightleftharpoons}} L_{i+j} \qquad \textbf{(Eq 1)}$$

$$L_i \underset{k_{bi}}{\overset{k_{ci}}{\rightleftharpoons}} C_i \qquad \textbf{(Eq 2)}$$

In the equations, *L* is linear members, *C* is cyclic members and subscripts *i* and *j* are the number of monomers that constitute the network members. In equation 1, the linear members linked through acetal bond with rate constant *k*, while the bond breaking rate constant is presented as *k/K*, the ratio of bond formation rate constant *k* to equilibrium constant *K*. In the equation 2, k_{ci} and k_{bi} are the cyclization and hydrolysis rate constant respectively. As a starting point, k_{ci} is estimated as the product of *k* and the cyclization factor, $Bi^{-3/2}$ (Eichinger, 2000), where *B* is the propensity of cyclization and k_{bi} is approximately *k/K*.

$$k_{ci} = kBi^{-3/2}$$
$$k_{bi} = kK^{-1}$$

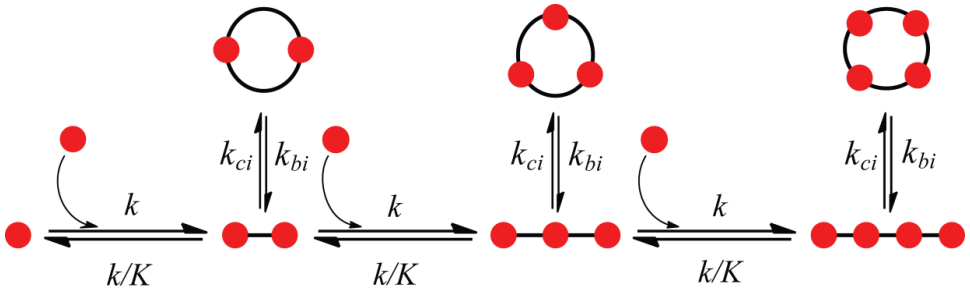


Figure 3-3 Kinetic model for the NF dynamic peptide network from monomer up to tetramer . In this model, the simplifying assumptions include i) the equilibrium constant (*K*) is identical for

each addition to monomer, dimer, trimer, etc. and ii) the condensation of two linear dimers do not contribute significantly to linear tetramer concentration.

Parameter values *k, K* and *B* are estimated by minimizing the sum of square error (SSE) between the calculated from model and the experimental HPLC data (**Table 3-1, 3-2**) (Miller & Rawlings, 1994). A Latin Hypercube sampling (McKay, Beckman, & Conover, 1979) provides the suitable initial guesses from the parameter space and the optimization procedure is carried out with the fmincon function in MATLAB. The confidence intervals of best fit parameters are estimated based on the singular value decomposition (Press, 1988) for their covariance matrices (Miller & Rawlings, 1994).

k	$8.79 \pm 1.34 \times 10^{-6}$ $mM^{-1}sec^{-1}$
K	$1.64 \pm 0.74 \times 10^{-1}$ mM^{-1}
B	$3.02 \pm 1.56 \times 10^{2}$ mM

Table 3-1 Fit parameters for the NF network model regressed to HPLC data up to 15 hours.

k	$8.73 \pm 1.34 \times 10^{-6}$ $mM^{-1}sec^{-1}$
K	$2.75 \pm 1.40 \times 10^{-1}$ mM^{-1}
B	$1.82 \pm 0.88 \times 10^{2}$ mM

Table 3-2 Fit parameters for the NF network model regressed to HPLC data up to 50 hours.

The model fits the early approach to equilibrium within 15 hours (**Figure 3-5, left**), however, after 15 hours, a drop in the monomer concentration and a concomitant increase in other oligomers are observed, resulting in deviation from the initial fitting of this model (**Figure 3-5, right**). Transmission electron microscopy (TEM) analyses provided evidence for the emergence of a new particle phase at 17 hours with particle size ~10 nm (**Figure 3-6 A, b**), which is in contrast to 4 hours when no particles are apparent (**Figure 3-6 A, a)**. The emergence of the particle phase correlated roughly in time with the drop in monomer concentration and the increase in the other oligomers. Examples of peptide assembly as particle phase are known previously in the amyloid peptide supramolecular assembly (Childers, Anthony, Mehta, Berland, & Lynn, 2012; Yan Liang, Lynn, & Berland, 2010) (**Figure 3-6 B**), but particles in those cases do not involve with the condensation reactions. Here, the emergence of particles phase appears to catalyze the reactions.

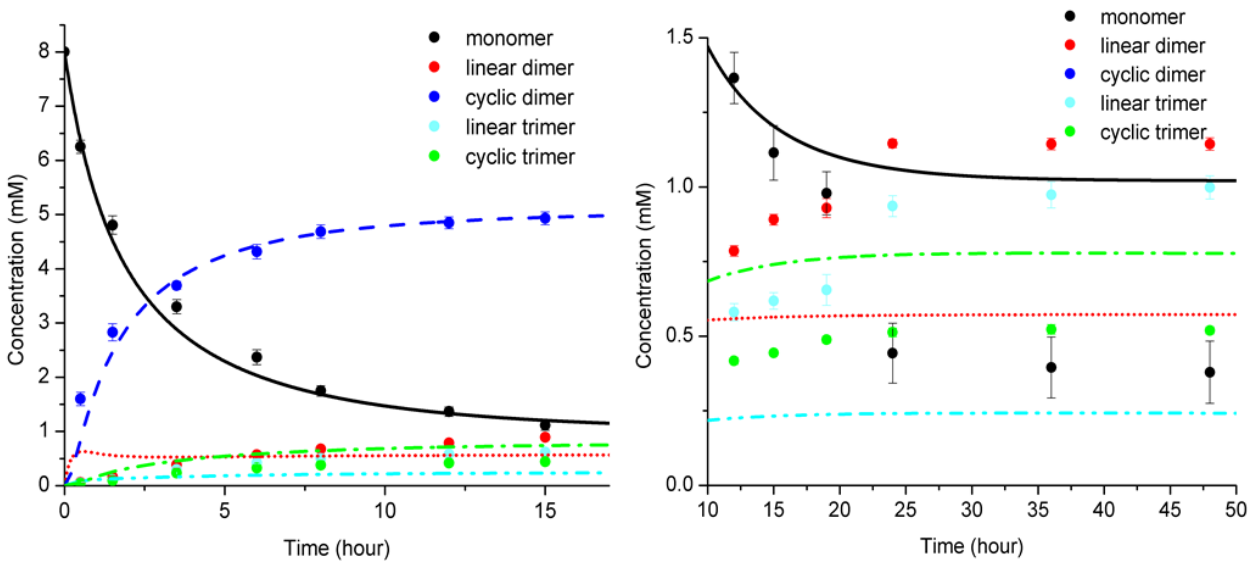


Figure 3-4 Kinetic model fits to the HPLC quantified network member concentrations. Kinetic model fits to the HPLC quantified network member concentrations. (**Left**) fits to early time points within 17 hours. (**Right**) model fits to 48 hours.

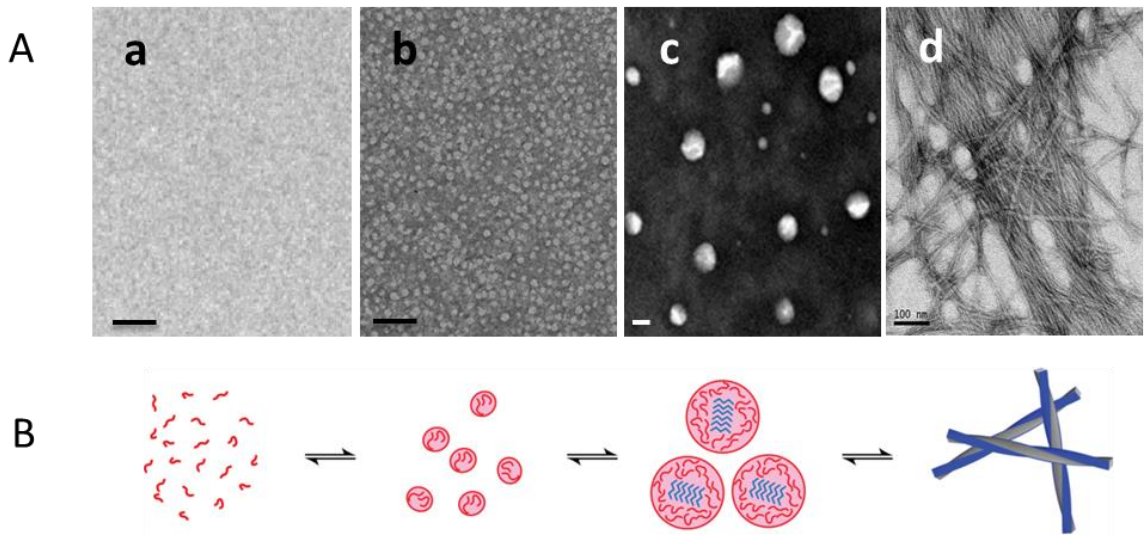


Figure 3-5 Transmission electron micrographs (**A**) of the NF-DCN reaction network taken at (**a**) 4 hrs, (**b**) 17 hrs, (**c**) 48 hrs, and (**d**) 96 hrs. Scale bar: 100 nm (**B**) Models of phases based on amyloid peptide analyses from reference (Childers et al., 2012).

To adjust the reaction schemes to these emergent phase transitions, the updated model (**Scheme 3-2**) employed two sets of kinetic values (**Table 3-3**) to fit the stages before (Stage 1) and after (Stage 2) the phase transition. *M*, *D*, D_C, *T* and T_C represent monomer, linear dimer, cyclic dimer, linear trimer and cyclic trimer, respectively. The rate constants are again estimated by minimizing the sum of square error (SSE) between the calculated and the experimental HPLC data. Some rate constants have large confidence interval due to the reason that limited data sets cannot accurately predict those eight rate constants. In the future, more sets of kinetics data need to be collected to narrow down the confidence interval of rate constants.

$$M + M \underset{k_2}{\overset{k_1}{\rightleftharpoons}} D$$

$$D \underset{k_4}{\overset{k_3}{\rightleftharpoons}} D_c$$

$$M + D \underset{k_6}{\overset{k_5}{\rightleftharpoons}} T$$

$$T \underset{k_8}{\overset{k_7}{\rightleftharpoons}} T_c$$

Scheme 3-2 Kinetic model for NF network. *M*, *D*, D_C, *T* and T_C represent monomer, linear dimer, cyclic dimer, linear trimer and cyclic trimer, respectively.

	Stage 1	**Stage 2**
$\mathbf{k_1}$	$7.98 \pm 20.4 \times 10^{-6}$ $mM^{-1}sec^{-}$	$8.66 \pm 473 \times 10^{-4}$ $mM^{-1}sec^{-}$
$\mathbf{k_2}$	$1.44 \pm 27.1 \times 10^{-5}$ sec^{-1}	$1.82 \pm 99.4 \times 10^{-4}$ sec^{-1}
$\mathbf{k_3}$	$2.50 \pm 32.2 \times 10^{-1}$ sec^{-1}	$2.40 \pm 813 \times 10^{-7}$ sec^{-1}
$\mathbf{k_4}$	$3.47 \pm 45.9 \times 10^{-2}$ sec^{-1}	$5.64 \pm 1870 \times 10^{-9}$ sec^{-1}
$\mathbf{k_5}$	$1.59 \pm 6.38 \times 10^{-5}$ $mM^{-1}sec^{-}$	$1.84 \pm 465 \times 10^{-2}$ $mM^{-1}sec^{-}$
$\mathbf{k_6}$	$1.17 \pm 21.9 \times 10^{-5}$ sec^{-1}	$1.14 \pm 287 \times 10^{-2}$ sec^{-1}
$\mathbf{k_7}$	$3.06 \pm 21.2 \times 10^{1}$ sec^{-1}	$2.69 \pm 14.4 \times 10^{-5}$ sec^{-1}
$\mathbf{K_8}$	$4.15 \pm 28.6 \times 10^{1}$ sec^{-1}	$5.20 \pm 27.8 \times 10^{-5}$ sec^{-1}

Table 3-3 Fits of rate constants for the NF network model including two stages.

Since the substantial drop of the monomer concentration occurs between 19 and 24 hours, a boundary at the middle point, 21.5 hour is set to separate the two stages as an initial approach

(**Figure 3-7**). The updated model is able to fit the NF network as two stages. In the first stage the network approaches equilibrium around 20 hours but then begins to change. At the onset of the particle stage, there is a drop in the monomer concentration and an increase in the linear dimer and linear trimer. However, there is no significant change in either the cyclic dimer or cyclic trimer. After the transition in the beginning of the second stage, the network members reach equilibrium and remain constant for over 25 hours and during this stage less than 5% of monomer remains.

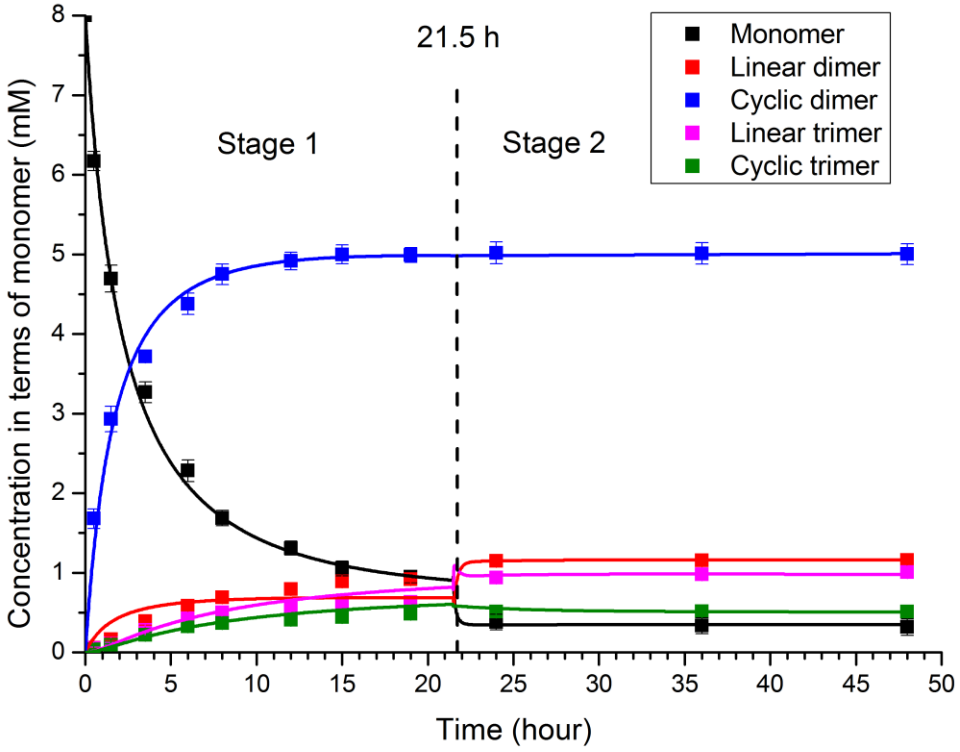


Figure 3-6 NF network members' kinetics are fitted into two stages . Stage 1 from 0 hour to 21.5 hours and Stage 2 from 21.5 to 48 hours.

During the equilibrium period, there is a global increase in the particle size, similar to Ostwald ripening (Baldan, 2002; Ostwald, 1901; Wang, Richards, Shields, & Buhro, 2014). Small particles gradually fuse into big particles and/ or are progressively dissolved and deposit into bigger ones. Large particles grow at the expense of smaller ones due to the driving force to lower total surface area and minimize the energy in the system. Meanwhile, the appearance of noticeable para-crystalline structures (as white area) inside big particles with average diameters of 120 nm at 48 hours (**Figure 3-6A, c**) are observed towards the end of the second stage. After 2 days, a significant increase in linear trimers at the expense of all the other species is observed (**Figure 3-8**), and TEM analyses reveals the transition from particles to paracrystalline fibril phase assemblies (**Figure 3-6A, d**). The increase of fibers is concomitant of notable decrease in the particle size and eventual disappearance of almost all the particles when the fibers mature at 2 weeks.

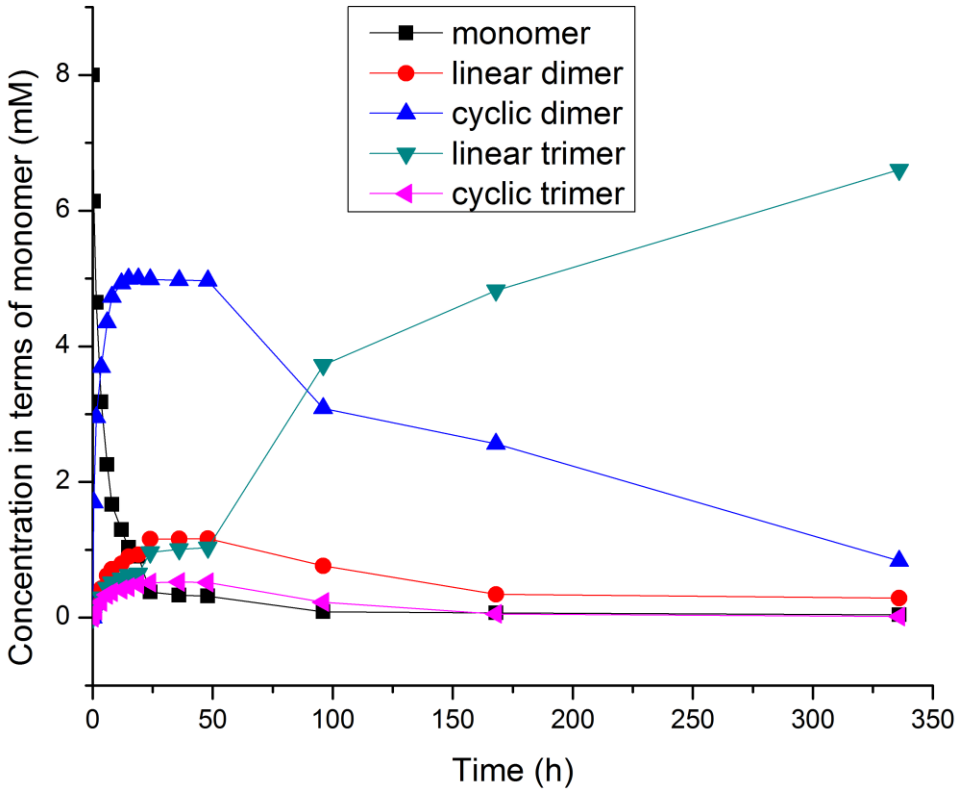


Figure 3-7 Kinetics of the NF dynamic network in two weeks.

3.3.5 Identification of Assemblies in the NF-DCN

As shown above, the phase transition from particles to fibers is accompanied by the increase of linear trimers, which is suggested to be the component of the fiber assemblies. To confirm this, the assemblies in the NF network at 7 days are enriched through centrifugation at 16,000 × g for 30 minutes, re-suspended in 40% CH_3CN/ 60% water, and analyzed by HPLC. As particle assemblies cannot be spun down at the above centrifugation condition, the pellet spun down contains only the fiber assemblies. The chromatogram confirms that fiber assemblies are composed of linear trimers (**Figure 3-9, bottom**). After 2 weeks, the pellet enriched again from the network shows that linear trimers are the components of fiber assemblies.

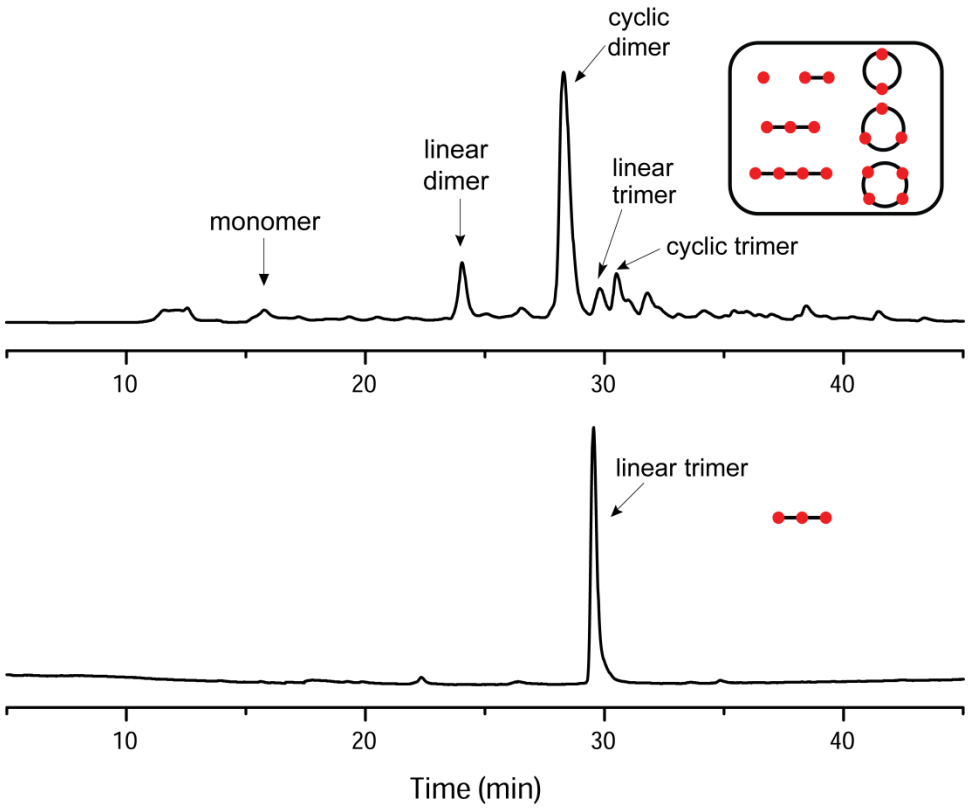


Figure 3-8 HPLC analyses of NF-DCN prepared in 60% water/40% CH_3CN after 2 days at 25°C. The linear dimer, cyclic dimer, linear trimer and cyclic trimer are assigned by mass spectral analysis (**top**). Following sedimentation at 16,000 x g, the pellet is re-suspended in 60% water/40% CH_3CN and analyzed under the same HPLC elution conditions (**bottom**). HPLC UV absorbance is monitored at 222 nm with a flow rate of 1.0 mL/min following a gradient from 10% acetonitrile to 90% acetonitrile in 40 minutes.

3.3.6 Probing the Critical Concentrations for Network members to Self-assemble

Even though the concentration of the cyclic dimer is 7 times larger than the linear trimer's at the second stage and linear dimer population is also more than the one fold of linear trimer, the fiber assemblies are formed by linear trimers instead of other network members. It is probably due to

higher propensity to self-assemble of linear trimers than dimers. Once the linear trimers reach the critical concentration, the exponential growth rate takes place due to the self-templating effect (Yan Liang et al., 2010). To test this hypothesis, control peptides H-NFNF-NH_2 (dimer counterpart in NF-DCN) and H-NFNFNF-NH_2 (trimer counterpart in NF-DCN) are synthesized and compared for their critical micelle concentration. Indeed, the trimer counterpart (six-residue peptide) assembled at much lower concentration than the dimer counterpart (four-residue peptide). The peptide H-NFNFNF-NH_2 self assembles at 0.3 mM (**Figure 3-10**), similar to the critical concentration for the linear trimer in NF-DCN to self-assemble. However, the critical concentration for peptide H-NFNF-NH_2 (linear dimer counterpart) to self-assemble to homogenous fibers is above 10 mM, which is ~ 30 times of the linear trimer counterpart.

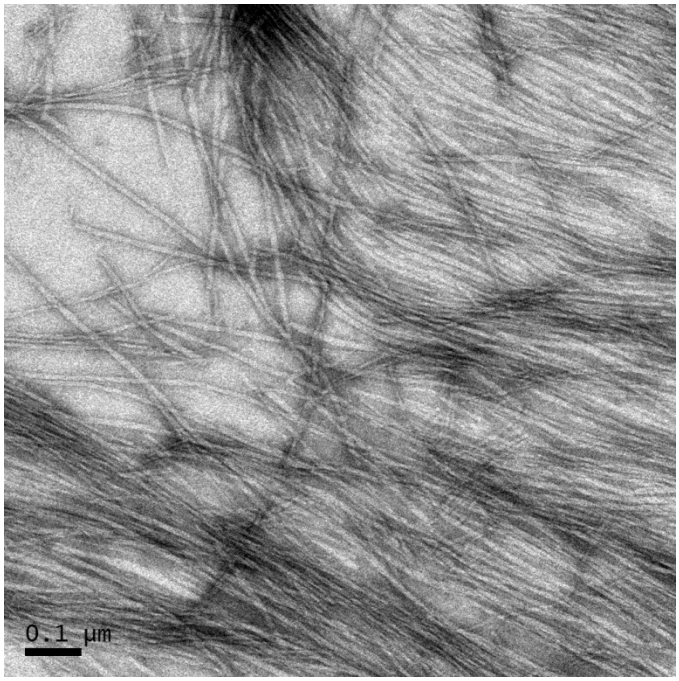


Figure 3-9 Transmission electron micrographs of NFNFNF-NH2 assemblies in water/acetonitrile (3/2, v/v). Scale bar: 0.1µm

3.3.7 Rapid Analysis of DCN

Although HPLC analysis has been established for probing the dynamic chemical networks, this method takes relative long time (~ 1 hour) for completing a single analysis. For more rapid network analysis, the ion-mobility spectrometry-mass spectrometry (IMS-MS) is optimized for analyzing the NF-DCN. When gaseous ions pass through a drift-tube filled with a buffer gas (e.g. helium), they are accelerated by an electric field and the mobility is governed by the mass to charge ratio (m/z), size and shape which results in an ion's characteristic collision cross-section (CCS) (**Figure 3-11, top**) (Creaser & Thomas, 2004; Kanu, Dwivedi, Tam, Matz, & Hill, 2008; Metz et al., 2008; Thiel et al., 2011; Woods, Radford, & Ashcroft, 2013). Large ions with expanded CCS have a longer drift time as they undergo more collisions with the buffer gas in the drift tube, whereas the smaller, more compact ions of the same molecular mass experience fewer collisions and traverse the drift tube in a shorter time. In comparison with HPLC, this rapid analysis of IMS-MS detection for each DCN sample proved to be 15-60 times faster for real-time analysis and demonstrate the distribution of NF-DCN clearly shown in **Figure 3-11 bottom**, where the monomer, linear dimer, cyclic dimer, linear trimer and cyclic trimer are separated according to their m/z and mobility. This method provides great potential for analysis of more complex networks in the future.

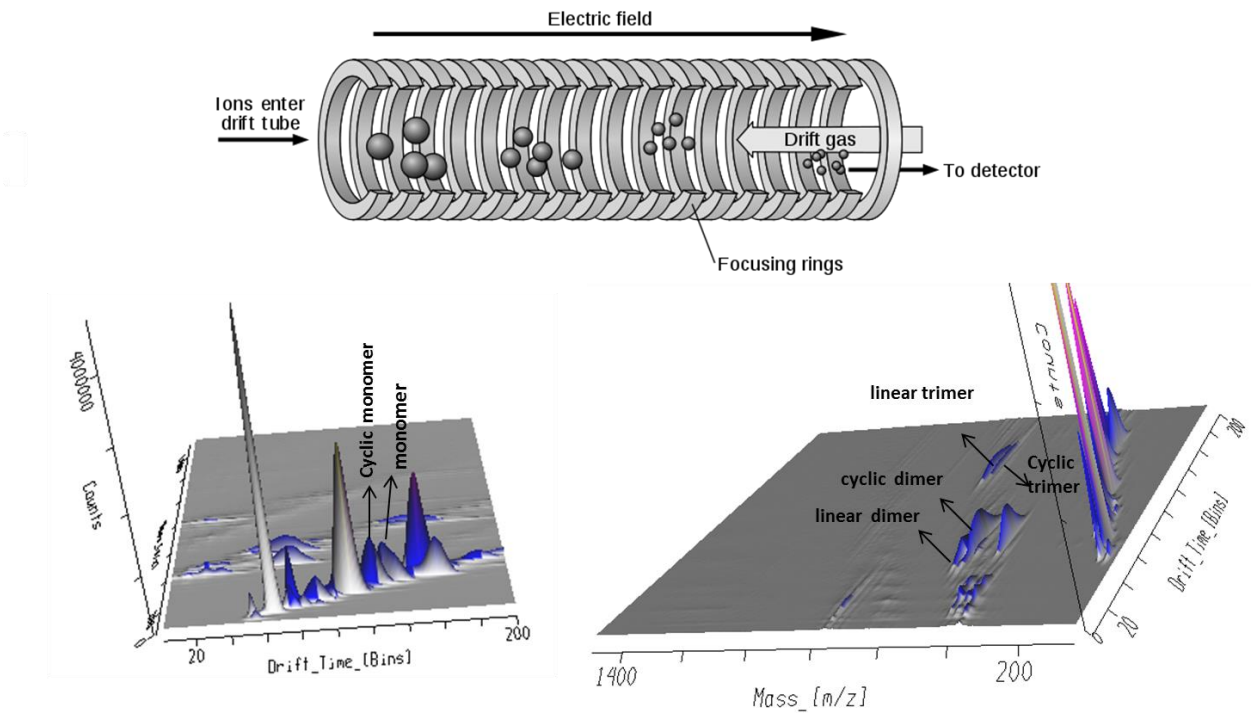


Figure 3-10 Ion-mobility spectrometry—mass spectrometry (IMS-MS) analysis of NF-DCN. (**Top**) illustration of the IMS unit for separation of analytes with the help of drift gas; (**Bottom**) Separation of network members through size, shape and m/z.

3.4 Discussions

This chapter has demonstrate the construction of the NF dynamic chemical network built from the simple dipeptide aldehyde NF-CHO. The strategy of reducing the oxidation state of the amino acids allows for making biomolecule-like oligomers in the absence of complex catalysts in aqueous solutions. The oligomers linked via reversible acetal bonds in this NF network can respond to environmental inputs (pH, temperature, solvents, etc.) and linear trimers are selected into assembled phase (**Figure 3-12**).

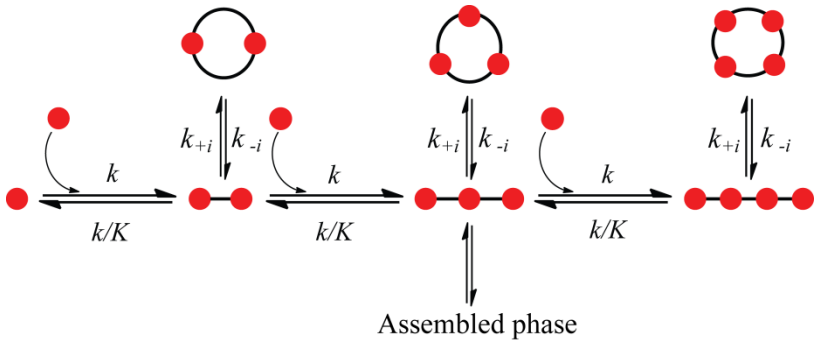


Figure 3-11 Illustration of thermodynamic equilibrium in NF dynamic network and the selection of linear trimers into assembled phase

The aging of NF network demonstrates progressive growth marked by two phase transitions. In the first stage, the cyclic dimer is the dominant species. In the second and third stage, population of linear trimers increases at the expense of other species, driven by the phase transitions from the solution to particle phase and to fiber phase. The emergence of particles at early stage in the NF network is similar to molten globule intermediates that occur in protein folding (Dinner, Sali, Smith, Dobson, & Karplus, 2000; Dobson, 1994; Muschol & Rosenberger, 1997; Pace, Shirley, Mcnutt, & Gajiwala, 1996), where backbone desolvation driven by hydrophobic collapse is thought to be critical for nucleation and assembly of native protein (Rose, Fleming, Banavar, & Maritan, 2006). The resulting aggregated state as molten globules remain disordered, but creates the desolvated interior necessary for structural element nucleation (Rose et al., 2006). This phenomenon is also directly observed in the pathway for amyloid cross-β peptide paracrystalline assemblies by fluorescence imaging and fluorescence correlation spectroscopy experiments (Yan Liang et al., 2010) and TEM analysis (Childers et al., 2012; Dong, Lu, Lakdawala, Mehta, & Lynn, 2006) (Y. Liang, Pingali, et al., 2008). The free peptides in solution are desolvated within

particle aggregates and display properties similar to the dehydrated molten globule intermediates in protein folding (Fernandez, Kardos, & Goto, 2003; Fernandez, Sosnick, & Colubri, 2002; Kim & Hecht, 2006; Meijer et al., 2007), where the early desolvation events through hydrophobic collapse are accountable for the nucleic formation (Kayed et al., 2007; Necula, Kayed, Milton, & Glabe, 2007). Computational analyses (Cheon et al., 2007; Cheon, Favrin, Chang, Dobson, & Vendruscolo, 2008) with Aβ (16-22) peptides suggest intermolecular hydrophobic collapse drive peptides to disordered molten aggregates initially, predisposing the peptides for interchain hydrogen bonding and reorganization to ordered paracrystalline phases, which is critical for the emergence of more ordered fibril phase assemblies.

As shown above, the NF network with acetal-linked members demonstrates the similar phase transition phenomenon as wild type peptides and proteins. But this is the first observation of particles impacting the growth of oligomer chain length. The particles and condensation reactions promote each other during the aging of network. Through oligomerization in solution, the increase in linear trimers drive the formation of particles by hydrophobic collapse, and in turn the desolvated interior environment provided by the particles facilitates the condensation reactions, resulting in the increase of linear trimers at the second stage.

Both linear dimer and linear trimer species increase at the onset of Stage 2, however, neither cyclic dimer nor cyclic trimer has obvious change in the second stage. The hypothesis for this is that since the partition coefficient is different for cyclic and linear species dictated by the nature of each network member, cyclic species might not be sufficiently incorporated into the particles as their structures are not typical amphiphilic as linear species with positive N-termini. The change in populations of monomer and linear species is probably due to re-distribution of those

members in particles phase. Clearly, further analysis of the particle phase including probing the inside environment and chemistry will be required.

Towards the end of Stage 2, some para-crystalline filaments appeared within the spherical particles (**Figure 3-6A, c**), suggesting conformational conversion from disordered structures in molten globules to ordered structures, which may be a consequence of the slow formation of highly directional interchain hydrogen bonds (Cheon et al., 2007). Both experimental and theoretical data in a range of peptide systems (Bader, Bamford, Zurdo, Luisi, & Dobson, 2006; Nguyen, Li, Stock, Straub, & Thirumalai, 2007; Pellarin & Caflisch, 2006; Petty & Decatur, 2005; Plakoutsi, Bemporad, Calamai, Taddei, & Dobson, 2005; Serio, Cashikar, Kowal, Sawicki, & Moslehi, 2000) have strongly suggest that this process as a generic one. Bigger particles provide greater statistical chances to allow for developing ordered structures, which serve the initial nuclei.

Phase transitions from building blocks to disordered oligomers and an ordered para-crystalline state are illustrated in **Figure 3-13**. In the third stage, para-crystalline fibers in the spherical particles released into the solution where fibers serve as templates for recruiting more building blocks at the fiber ends. Breakage of fibers due to collision, shearing force, etc. produced more fiber templates for autocatalytic growth, which resulted in exponential increase in linear trimers in stage 3.

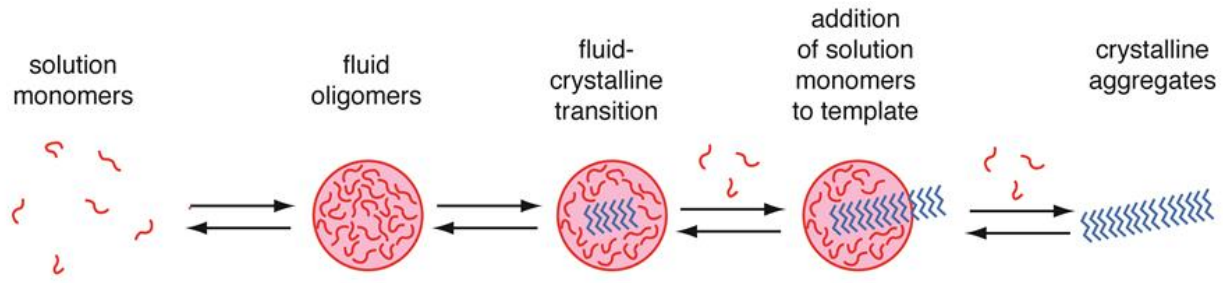


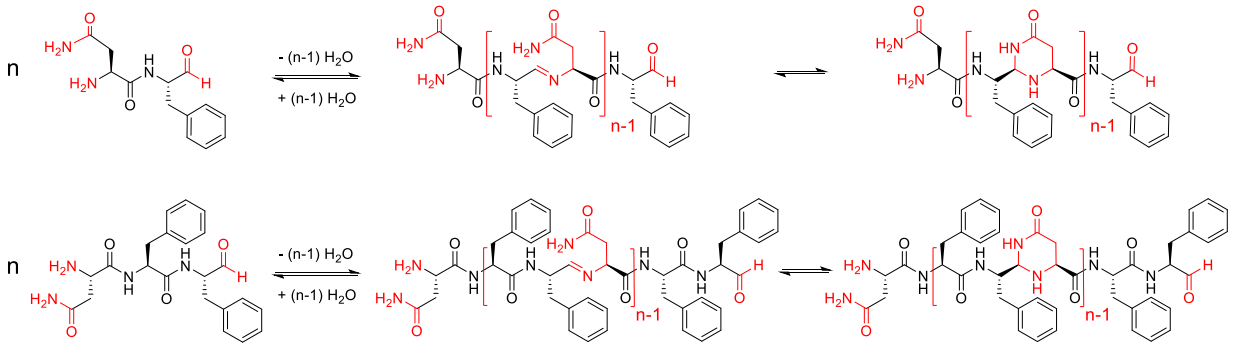
Figure 3-12 Model for the phase transition from monomers to disordered oligomers and an ordered para-crystalline state. Red denotes unstructured peptides and blue represents the transition to the more crystalline structures. (modified from ref **(Childers, 2010)**).

Within two weeks, the fact that the linear trimer dominate the network (**Figure 3-8**) at the cost of other members. We reason that the molecular recognition through hydrogen bonding and Phe-Phe residue stacking of linear trimers drive the conformational change from randomly distributed status to organized para-crystalline filament structures in molten globules. Later the filaments released into solution serve as templates for recruiting free linear trimers, shifting the equilibrium in solution to drive the condensation reactions towards linear trimer. Prior to the assembly phases, cyclic dimer dominates in solution by thermodynamic control, however, the emergence of phase transitions select linear trimers to self- assemble into fibers, which drive the equilibrium away from the initial status in solution. The self-assemblies serve as a kinetic trap to slow re-equilibrium of linear trimers. The fascinating results of self-organization/ self-templating of linear trimers in the NF-DCN imply potential for chemical evolution where the production of self-replicating biopolymers is dictated by the tension between equilibrium thermodynamics and kinetic control.

Chapter 4: NFF-CHO Dynamic Chemical Network

4.1 Introduction

In Chapter 3, I developed a dynamic chemical network (NF-DCN) built from the modified dipeptide monomer Asn-Phe-CHO (NF-CHO) (**Scheme 4-1, top**) and demonstrate its responsiveness to environmental changes, such as pH change, temperature change, solvent change and the kinetic selection of linear trimer through autocatalysis and self-templating effect. In this chapter, I generalize the process to construct a monomer with stronger potential for self-assembling as the tripeptide Asn-Phe-Phe-CHO (NFF-CHO). The strategy is to use the same N, N-acetal but with additional phenylalanine residue to stabilize inter-sheet associations (Lu et al., 2003; Mehta et al., 2008) by placing the phenylalanine sidechains arrayed on both faces of the sheet (**Scheme 4-1, bottom**).



Scheme 4-1 Chemical structures of *NF*-CHO (**top**) and *NFF*-CHO (**bottom**) dipeptide building block and resulting oligomers in dynamic chemical networks.

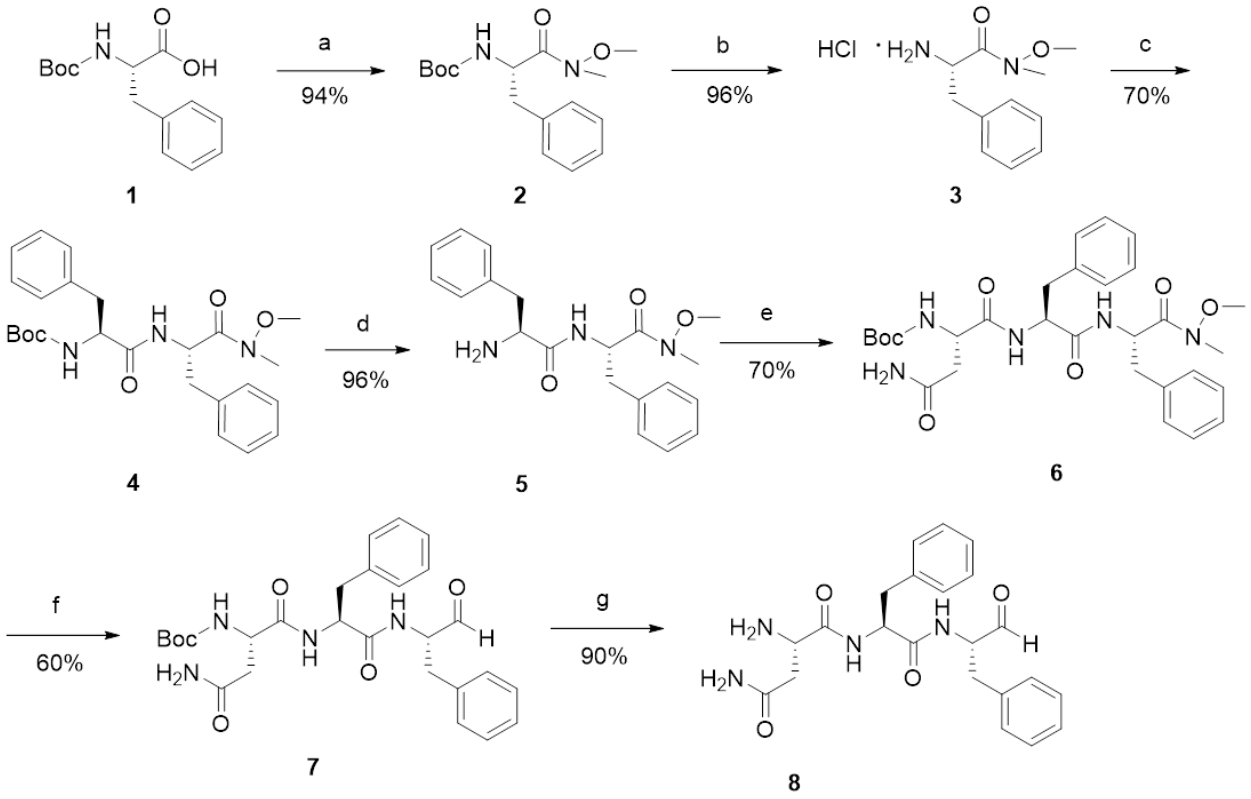
4.2 Materials and Methods

4.2.1 Materials

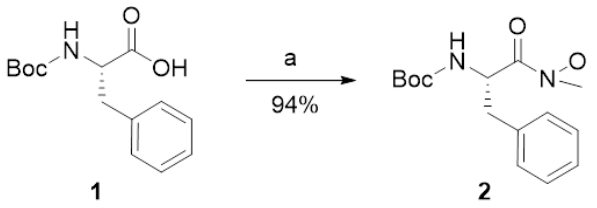
All commercially available chemicals are purchased from Sigma-Aldrich, AnaSpec and Nova Biochem. Anhydrous solvents are either dried over molecular sieve (4 Å) that had been pre-treated overnight at 300 ºC or purchased from EMD or Acros organics. HPLC grade acetonitrile and water are obtained from Sigma-Aldrich and/or Fisher Scientific. TLC plates are purchased from EMD (silica gel 60 F_{254}). Fmoc-amino acids, resins and solid phase peptide synthesizer reagents are purchased from AnaSpec. Distilled deionized water for sample preparation is obtained from EMD chemicals Inc.

4.2.2 Synthesis of Building Blocks

Synthesis of NFF-CHO



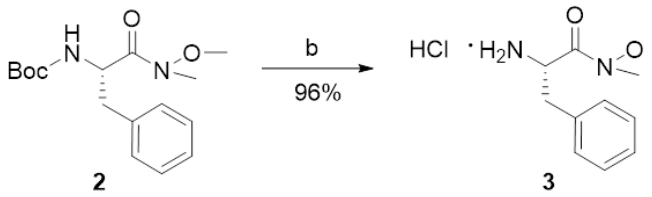
f. Preparation of the Boc-Phe-weinreb amide **2**



N-Boc-L-Phe (2.65 g, 10 mmol) is dissolved in anhydrous dichloromethane (DCM), and then 1,1-carbonyldiimidazole (1.78 g, 11 mmol) is added. The resultant mixture is stirred for an hour at room temperature. Subsequently, N,O-dimethylhydroyxlamine hydrochloride (1.07 g, 11 mmol) is added and the reaction mixture (light yellow suspension) is stirred and allowed to proceed overnight. Then the solvent is removed by vacuum at 25 °C and the resulting residue is

purified by extraction, which is dissolved in ethyl acetate (100ml) and ished successively with 2×100ml 1M HCl, 2×100ml saturated aqueous $NaHCO_3$ and 2×100ml brine. Finally the organic layer is dried with anhydrous Na_2SO_4 and evaporation of the solvent yielded product **2**. Product is visualized by UV on TLC plate (hexane/ ethyl acetate (1/1) as solvent system, $R_f \sim 0.46$). ^1H NMR (600 MHz, Chloroform-d) δ 7.30 –7.14 (m, 5H, C_6H_5), 5.17 (d, J = 8.8 Hz, 1H, CH (α)), 4.94 (q, J = 7.1 Hz, 1H, NH), 3.65 (s, 3H, O-CH_3), 3.16 (s, 3H, N-CH_3), 3.05/2.87 (dd, 2H, CH_2(β)), 1.38 (s, 9H, $C(CH_3)_3$). MS (ESI) m/z $(M+H)^+$:309.1803 Calculated $(M+H)^+$: 308.1808 ($C_{16}H_{25}N_2O_4$)

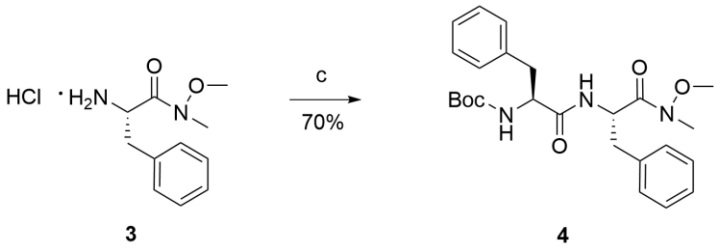
g. Removal of the t-butoxycarbonyl protecting group of **2**



2 **3**

The white solid **2** (2.90 g, 9.4 mmol) prepared above is dissolved in 10 ml ~ 20 ml 4M HCl in dioxane (purchased from Aldrich), and then the resultant mixture is stirred at room temperature for 0.5hr to 1hr at room temperature. Solvent is removed with vacuum at room temperature, leaving the product (**3**) of the hydrochloride salt as thick waxy oil. ^1H NMR (600 MHz, Chloroform-*d*) δ 8.83 – 8.29 (s, 3H, NH_3^+), 7.47 – 7.05 (m, 5H, C_6H_5), 4.74 (q, *J* = 5.9 Hz, 1H, CH (α)), 3.62 (s, 3H, O-CH_3), 3.40/3.28 (m, 2H, CH_2(β)), 3.10 – 3.07 (s, 3H, N-CH_3). MS (ESI) m/z $(M+H)^+$: 209.1284 Calculated $(M+H)^+$: 209.1285 $C_{11}H_{17}N_2O_2$

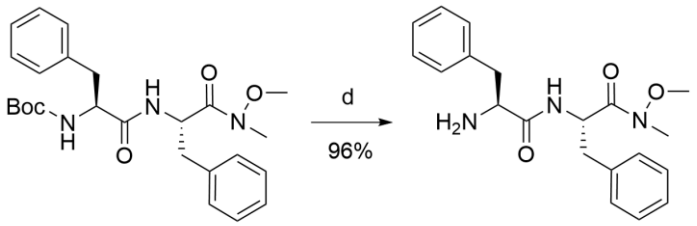
h. Coupling Boc-Phe residue to **3**



3 **4**

The waxy oil **3** (2.20 g, 9 mmol) is dissolved in anhydrous DCM (50-80 ml) before adding Boc-L-Phe (2.63 g, 9.9 mmol) and trimethylamine (2.5 ml, 18 mmol). The reaction mixture is cooled to 0°C in an ice bath before adding 1-Ethyl-3(3-dimethylaminopropyl) carbodiimide (EDC) (1.90 g, 9.9 mmol) stirred for 2hrs at 0°C under N_2 followed by overnight at room temperature. The solvent is evaporated *in vacuo* at room temperature, and the resulting residue is dissolved in ~100ml ethyl acetate and ished successively with 2×100ml of 1M HCl, 2× 100ml of 4% $NaHCO_3$ and 2×100ml of brine. The excess of reagents of Boc-L-Asn and EDC are ished away by aqueous layer. The organic layer is dried with anhydrous sodium sulfate, decanted, and the solvent removed in vacuo to give white solid **4**. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.40 – 7.07 (m, 10H, 2 C_6H_5), 6.55 (d, J = 8.3 Hz, 1H, NH), 5.19 (s, 1H, NH), 4.94 (m, 1H, CH (α)), 4.36 (m, 1H, CH (α)), 3.61 (s, 3H, O-CH_3), 3.13 (s, 3H, N-CH_3), 3.08 – 2.78 (m, 4H, 2 CH_2(β)), 1.39 (s, 9H, C(CH_3) $_3$). MS (ESI) m/z $(M+H)^+$: 456.2497 Calculated $(M+H)^+$: 455.2496 ($C_{25}H_{34}N_3O_5$)

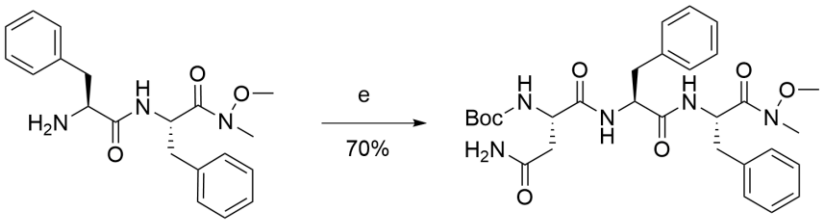
i. Removal of the t-butoxycarbonyl protecting group of **4**



4 **5**

The white solid **4** (2.87 g, 6.3 mmol) is dissolved in 10~20 ml 4M HCl in dioxane (Aldrich), and the resultant mixture stirred at room temperature for 0.5hr to 1hr. The solvent is removed directly *in vacuo* to give **5** hydrochloride as a thick waxy solid. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.64 (s, 2H, NH_2), 7.84 (s, 1H, NH), 7.40 – 7.12 (m, 10H, 2 C_6H_5), 5.21 – 4.95 (m, 1H, CH (α)), 4.35 (s, 1H, CH (α)), 3.50 (s, 3H, O-CH_3), 3.41 (m, 2H, CH_2(β)), 3.19 (m, 2H, CH_2(β)), 3.03 (s, 3H, N-CH_3). MS (ESI) m/z $(M+H)^+$: 356.1976 Calculated $(M+H)^+$: 356.1969 ($C_{20}H_{26}N_3O_3$)

j. Coupling of Asn to **5**

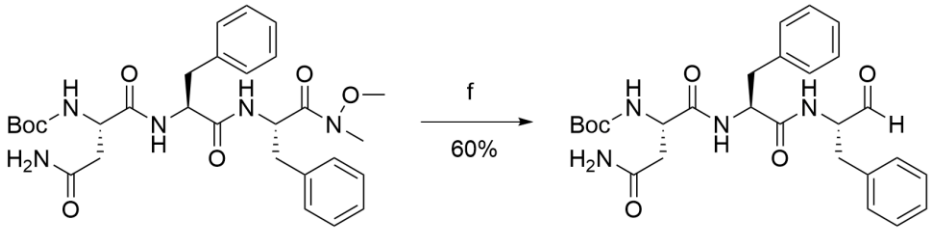


5 **6**

The thick waxy solid **5** (2.13 g, 6.0 mmol) is dissolved in anhydrous DCM (50-80 ml) before Boc-L-Asn (1.53 g, 6.6 mmol) and trimethylamine (1.7 ml, 12 mmol) are added. The reaction mixture is cooled to 0°C in an ice bath before EDC (1.27 g, 6.6 mmol) is added and the mixture stirred for 2hrs at 0°C and overnight at room temperature. When the solvent is removed *in vacuo*,

the resulting residue can be dissolved in ~100ml ethyl acetate and ished successively with 2×100ml of 1M HCl, 2× 100ml of 4% $NaHCO_3$ and 2×100ml of brine. The organic layer is dried with anhydrous sodium sulfate, decanted, and the solvent removed in vacuo to give white solid 6. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.30 – 7.13 (m, 10H, 2 C_6H_5), 7.11 – 6.98 (m, 2H, NH_2), 6.23 (s, 1H, NH), 5.92 – 5.87 (m, 1H, NH), 5.59 (s, 1H, NH), 5.19 (q, *J* = 7.2 Hz, 1H, CH (α)), 4.70 (p, *J* = 8.6, 7.2 Hz, 1H, CH (α)), 4.47 (q, *J* = 6.6 Hz, 1H, CH (α)), 3.64 (s, 3H, O-CH_3), 3.11 (s, 3H, N-CH_3), 3.09 – 2.88 (m, 4H, 2 CH_2(β)), 2.81 (dd, *J* = 16.5, 4.0 Hz, 1H), 2.51 (dd, *J* = 15.3, 7.6 Hz, 1H), 1.42 (s, 9H, $C(CH_3)_3$). MS (ESI) m/z $(M+H)^+$: 570.2932, Calculated $(M+H)^+$: 570.2922 ($C_{29}H_{40}N_5O_7$)

k. Reduction of the Boc-Asn-Phe-Phe-weinreb amide

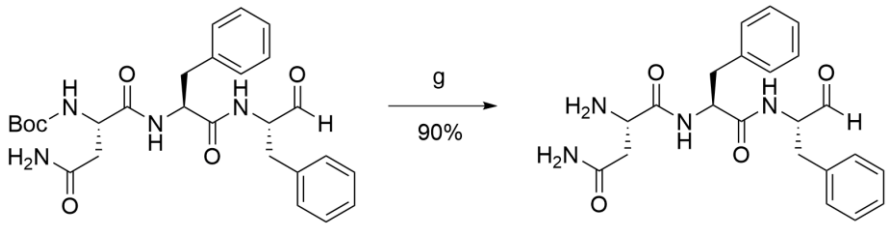


6 **7**

White solid **6** (4.2 mmol) is dissolved in 60 ml anhydrous THF under N_2 and cooled to -78°C in a dry ice / acetone bath before LAH (1.0 M) in THF (6.3 ml, 6.3 mmol) is added dropwise via syringe. When addition is complete the acetone bath is replaced with a 0 °C ice bath for 30min and the mixture is cooled again to -78°C and quenched quickly with 10 ml of 1M aqueous potassium bisulfate ($KHSO_4$). This mixture is allowed to warm to room temperature, extracted several times with ethyl acetate, and the combined extracts ished with 2×50ml $NaHCO_3$, 2×50ml brine, dried with anhydrous Na_2SO_4. The solvents are evacuated by vacuum under 20 °C,

providing products 7 as waxy solid. ^{1}H NMR (600 MHz, DMSO-d_6) δ 9.32 (s, 1H, CHO), 8.55 (d, J = 7.5 Hz, 1H, NH), 7.95 (d, J = 8.0 Hz, 1H, NH), 7.32 (s, 1H, NH), 7.30 – 7.13 (m, 10H, 2 C_6H_5), 6.93 – 6.87 (m, 2H, NH_2), 4.46 (m, 1H, CH (α)), 4.23 (2H, 2 CH (α)), 3.10 (m, 1H, CH_2(β)), 2.94 (dd, J = 13.8, 5.1 Hz, 1H, CH_2(β)), 2.84 – 2.69 (m, 2H, CH_2(β)), 2.39 (m, 1H, CH_2(β)), 2.31 (dd, J = 15.1, 7.9 Hz, 1H, CH_2(β)), 1.35 (s, 9H, $C(CH_3)_3$) MS (ESI) m/z $(M+Na)^+$:533.2376 Calculated $(M+Na)^+$:533.2371 ($C_{27}H_{34}N_4O_6$)

l. Deprotection of Boc group to give rise to the free monomer NFF-CHO **8**



The waxy solid **7** (1.29 g, 2.52 mmol) is dissolved in 10 ml trifluoroacetic acid (bubbled with N_2 gas for at least 15 min before usage) under N_2, stirred at room temperature for 0.5hr to 1hr, and the solvent removed in vacuo to give **8**·hydrochloride as a thick waxy solid. MS (ESI) m/z $(M+H)^+$:411.2031 Calculated $(M+H)^+$: 411.2027 ($C_{22}H_{27}N_4O_4$)

4.2.3 NMR Analysis

^{1}H NMR data are recorded on an INOVA 600 NMR spectrometer or an INOVA 400 NMR spectrometer (equipped with Bore Oxford super conducting magnet). The solvents Chloroform-*d* and DMSO-d_6 are purchased from Cambridge Isotope Laboratories, Inc. The choice of the solvent is determined by the solubility of the compounds.

4.2.4 Dynamic Peptide Network Preparation

After final deprotection of the precursor of the network, Boc-NFF-CHO by trifluoroacetic acid, resulting network building blocks of NH_2-NF-CHO are dried under vacuum at room temperature. Dried NH_2-NFF-CHO is stored with Argon protection at -20°C. To prepare the dynamic network, NH_2-NFF-CHO are dissolved to a concentration of 8mM in water/acetonitrile (3/2, v/v) at ambient temperature under N_2 protection. All the solvents are flushed with N_2 gas for at least 15 min before usage. Dissolution is assisted by ~2 minutes of continuous vortexing, followed by ~15 minutes of bath sonication until solution became clear. The pH of the solution is adjusted by titrating aliquots of 10mM NaOH and the final pH value is determined by pH meter (Fisher Scientific Accumet Basic AB15 pH meter).

4.2.5 HPLC and LC-MS Analyses

HPLC analyses are performed on Waters Delta 600 equipped with a photodiode array UV/Vis detector at room temperature using a reversed-phase HPLC column (Kromasil 100-5C18, 4.6 × 250mm). Solvent A: water (0.1 vol % trifluoroacetic acid). Solvent B: acetonitrile (0.1 vol % trifluoroacetic acid). UV absorbance is monitored at 258nm (for Phenyl ring side chain absorption) and 222nm (for amide bond absorption). Flow rate is 1.0 mL/min. Gradient is from 10% acetonitrile to 90% acetonitrile, 2% acetonitrile/ min. LC-MS analyses are performed on Waters Synapt G2 MS/Acquity UPLC system. Positive-ion mass spectra are obtained using electrospray ionization.

Time (mins)	Solvent A	Solvent B	Note
0	90%	10%	Gradient starts
40	10%	90%	Gradient ends
41	0%	100%	Cleaning starts
48	0%	100%	Cleaning ends
49	10%	90%	Re-equilibrium starts
56	10%	90%	Re-equilibrium ends

4.2.6 IMS-MS

Direct total ion current analysis is performed on Synapt G2 High Definition Mass Spectrometry system (Waters Corporation, Manchester, UK), which is a hybrid quadrupole-ion mobility-orthogonal acceleration time-of-flight instrument, with typical resolving power of 20,000 m/Δm (FWHM) and mass accuracy of 9 ppm at m/z 556.2771. The instrument is operated in positive ion mode with a probe capillary voltage of 3 kV, and a sampling cone voltage of 45 V. The source and desolvation temperatures are set to 120 °C and 250 °C, respectively; and the nitrogen desolvation flow rate is set to 650 L h-1. The mass spectrometer is calibrated across the 50-1200 m/z range using a 0.5 mM sodium formate solution prepared in 90:10 2-propanol:water v/v. Data are mass corrected during acquisition using a leucine encephalin (m/z 556.2728) reference spray (LockSpray) infused at 2 µL min-1. The scan time is set to 1 s. Data acquisition and processing is carried out using MassLynx v4.1 and Drift Scope v2.1 (Waters Corp.) Column: Waters Acquity UPLC BEH C18 column 1.7 µm, 2.1 x 50 mm)

4.2.7 Transmission Electron Microscopy and Electron Diffraction

Aliquots (10 µl) of sample solutions are dropped on TEM grids (200 mesh copper grid covered with a thin carbon film, purchased from Electron Microscopy Sciences) for 3 minutes before

excessive solution is blotted with filter paper. Uranyl acetate (10 µl of 5% solution) is added for 3 minutes for negative staining. Extra fluid is blotted with filter paper. The grids are analyzed on a Hitachi H-7500 transmission electron microscope with a LaB6 emission filament at an accelerating voltage of 75 kV.

4.2.8 Thioflavine T (ThT) Fluorescence

The ThT stock solution is prepared by adding 8mg ThT to 10ml phosphate buffer and filter through a 0.2µm syringe filter. A aliquot of ThT stock solution is then added to the dynamic peptide solution to reach a final concentration of 0.25mM, 2.5mM, respectively (ThT: peptide molar ration is 1:10). The working solution is excited at 440nm and the fluorescence intensity is recorded at emission 482nm.

4.2.9 X-ray diffraction (XRD) Analyses

The samples for X-ray diffraction are dried through lyophilization. The diffraction patterns are measured Bruker APEX-II diffractometer with graphite monochromated Cu radiation K-alpha radiation, lambda= 1.54184 Ang, 40 kV and 35 mA, with a 0.5 pinhole collimator and with exposure times of 900 s per frame. The samples are loaded into mylar capillaries.

4.3 Results and Discussions

4.3.1 Construction of NFF-CHO Dynamic Chemical Network (NFF-DCN)

The tripeptide aldehydes (NFF-CHO) are prepared with standard Boc-peptide synthesis and reduced with $LiAlH_4$ (detailed procedures are shown in Method session). The *tert*-Butyloxycarbonyl protecting (Boc) group of the precursor Boc-NFF-CHO is removed by trifluoroacetic acid (TFA) at room temperature. Then the dried monomer (8 mM) is incubated in

water/acetonitrile (3/2, v/v), yielding a dynamic network of oligomers linked via reversible N, N-acetal (**Scheme 4-1, bottom**). Similarly to the NF-DCN, the acetal stereochemistry and degree of assembly is responsive to pH, temperature and solvent conditions as confirmed by high performance liquid chromatography (HPLC) and mass spectrometry.

4.3.2 Direct Analysis of the Network Composition by HPLC

As developed in Chapter 3, the direct HPLC analysis of N, N-acetal network has been achieved by using CH_3CN/ H_2O at quasi neutral pH to sufficiently slow down the hydrolysis rate of the acetal linkage. This method is applied to NFF-DCN as well. As shown in **Figure 4-1**, the NFF network of oligomers demonstrates a distribution of monomer (H-NFF-CHO), linear dimer (H-NFFaNFF-CHO, 'a' represents the acetal linkage) and linear trimer (H-NFFaNFFaNFF-CHO), confirmed by electrospray ionization (ESI) mass spectrometry shown in **Figure 4-2 A, B, C**, respectively. HPLC chromatogram shows that each network member is well resolved with good Gaussian peak shape, suggesting adequate hydrolysis stability and they can be isolated under these conditions and stay stable for more than 48 hours.

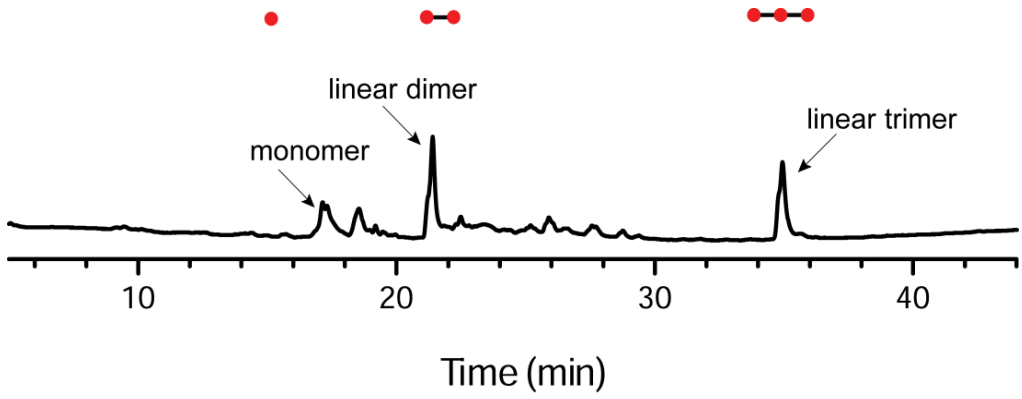


Figure 4-1 High performance liquid chromatography (HPLC) analysis of NFF dynamic peptide network at 2 days in water/acetonitrile (3/2, v/v) at pH 4. HPLC UV absorbance is monitored at

254 nm; flow rate is 1.0 mL/min with gradient from 10% acetonitrile to 90% acetonitrile in 40 minutes.

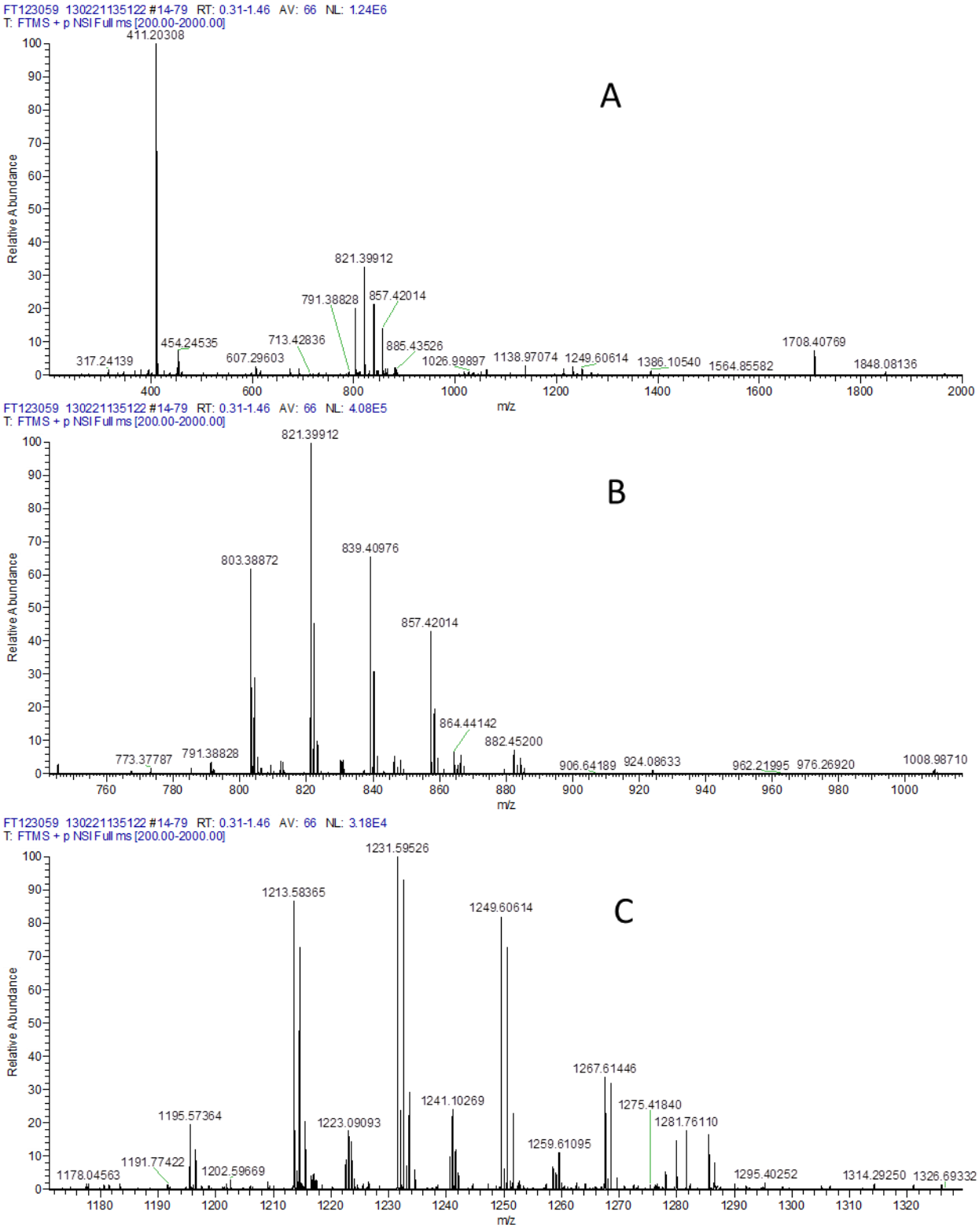

FT123059 130221135122 #14-79 RT: 0.31-1.46 AV: 66 NL: 1.24E6
T: FTMS + p NSI Full ms [200.00-2000.00]
411.20308
821.39912
791.38828
857.42014
713.42836
454.24535
885.43526
317.24139
607.29603
1026.99897
1138.97074
1249.60614
1386.10540
1564.85582
1708.40769
1848.08136
A
Relative Abundance
m/z
FT123059 130221135122 #14-79 RT: 0.31-1.46 AV: 66 NL: 4.08E5
T: FTMS + p NSI Full ms [200.00-2000.00]
821.39912
803.38872
839.40976
857.42014
864.44142
882.45200
773.37787
791.38828
906.64189
924.08633
962.21995
976.26920
1008.98710
B
Relative Abundance
m/z
FT123059 130221135122 #14-79 RT: 0.31-1.46 AV: 66 NL: 3.18E4
T: FTMS + p NSI Full ms [200.00-2000.00]
1231.59526
1213.58365
1249.60614
1267.61446
1275.41840
1241.10269
1281.76110
1195.57364
1223.09093
1259.61095
1191.77422
1178.04563
1202.59669
1295.40252
1314.29250
1326.69332
C
Relative Abundance
m/z

Figure 4-2 Positive mode of ESI mass spectrum of NFF-DCN in acetonitrile/ water. (**A**) Monomer $(M+H)^+$: 411.203; (**B**) linear dimer $(M+H)^+$: 803.388; $(M+H+H_2O)^+$: 821.399, $(M+H+2H_2O)^+$: 839.410; $(M+H+3H_2O)^+$: 857.420 and (**C**) linear trimer $(M+H)^+$: 1195.574, $(M+H+H_2O)^+$: 1213.584, $(M+H+2H_2O)^+$: 1231.595, $(M+H+3H_2O)^+$: 1249.606, $(M+H+3H_2O)^+$: 1267.614

4.3.3 pH Profile of NFF-DCN

The pH dependence of NF dynamic network is shown in the previous chapter. Built on the same reversible linkages, NFF dynamic network should also respond to pH change. Ion-mobility spectrometry—mass spectrometry (IMS-MS) is applied for rapid analysis of 20 NFF-DCNs at a range of pHs. Compared to HPLC, the IMS-MS detection for each DCN proved to be 15-60 times faster for real-time analysis, where the mobility of ions in the IMS is governed by the size, shape and mass to charge ratio (m/z) (Creaser & Thomas, 2004; Kanu et al., 2008; Metz et al., 2008; Thiel et al., 2011). Samples of NFF-DCN are prepared in water/acetonitrile (3/2, v/v) with different pH values adjusted by NaOH and HCl solutions and checked by pH meter (details in Materials and Method session). The samples are purified through a short analytical column to rinse off any excessive salts before injecting samples of DCNs into IMS-mass spectrometer. As indicated in **Figure 4-3**, network members elute at different time periods, similarly to what we observed in HPLC analyses. The monomer (green peaks) eluted first, then linear dimer (purple peak) and linear trimer (data not shown in this figure) subsequently. To quantify the amount of each network member, the elution peaks for monomer, dimer and trimer are extracted by the their exact mass with 4 decimals, which greatly reduce any background noise of each peak, eliminating potential chemicals with similar molecular weight or eluting time. This ensures that

only the correct species with exact m/z are selected for integration of ion current. The percentage of each network member's ion current peak integration is converted to its concentration, by multiplying the percentage and total concentration.

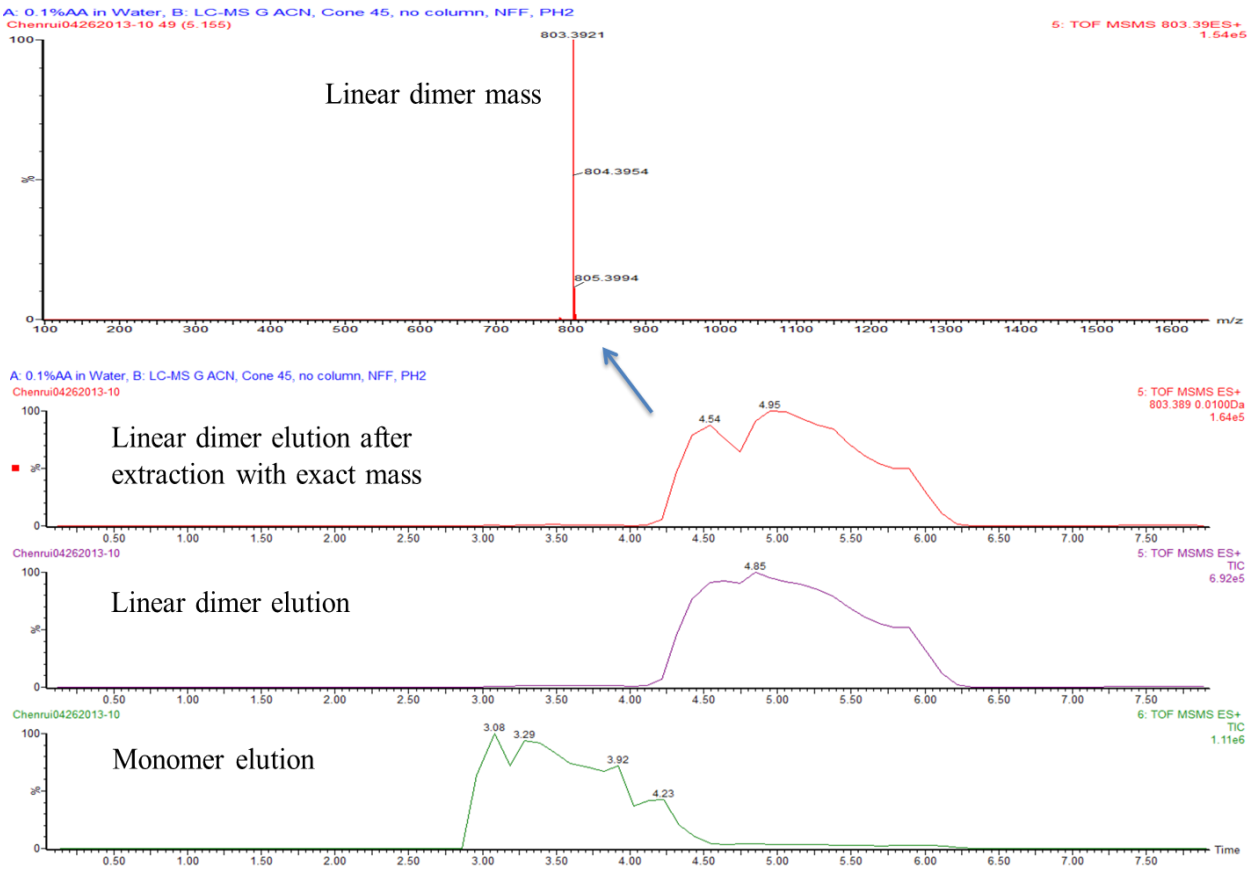


Figure 4-3 IMS-MS Analysis of NFF-DCN at pH 2. Peaks denoted in green is monomer elution ion current peaks. Peak in purple is the linear dimer ion current peak. Red peaks are linear dimer peak after extracted by the exact mass of linear dimer. Linear dimer mass $(M+H)^+$ 803.3921 is shown on the top.

The concentrations of each network members versus twenty pHs are plotted out (**Figure 4-4 bottom**). Interestingly, two pseudo pK_a values can be derived through this plot. One pK_a is

around 7, indicating the protonation of nitrogen atom in the N, N-acetal ring, consistent with the reported pK_a in similar structures (Wuitschik et al., 2010); the other pK_a is around 3, indicating the protonation of oxygen atom that is attached to the N,N-acetal ring (**Figure 4-4 top**). This direct real time compositional analysis assigns pseudo pKa values for N, N-acetal, suggesting the stability of N, N-acetal at various pHs and provides insights for optimization of the pH for network analysis.

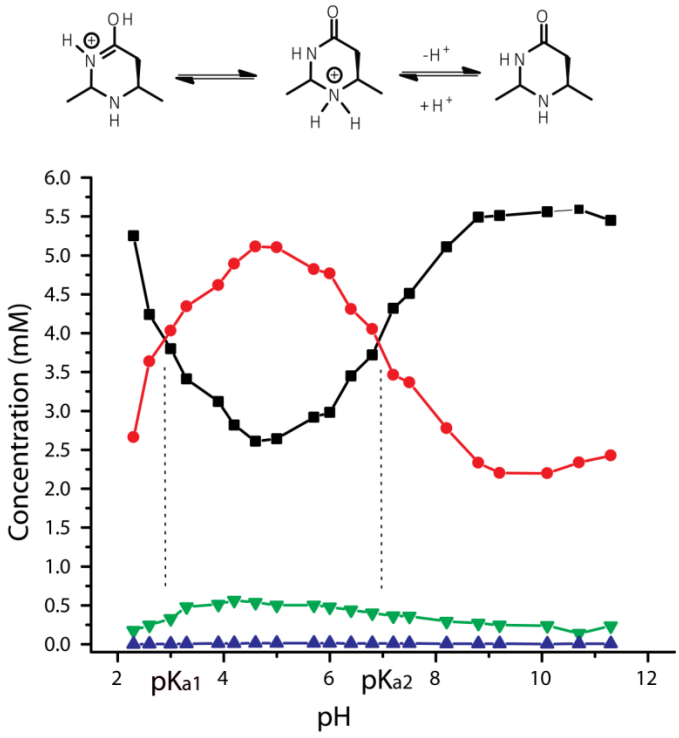


Figure 4-4 Proposed protonation states of the N,N-acetal from (**top**), the total integrated ion current of the network members from equilibrating 8 mM NFF-CHO in water/acetonitrile (3/2, v/v) at the indicated apparent pH values (**bottom**). (Black: monomer; Red: linear dimer; Green: linear trimer; Blue: linear tetramer)

4.3.4 Kinetics Study of NFF-DCN Reveals an Emergent Phase That is Far From Equilibrium

The aging of the NFF dynamic network is followed by HPLC analyses. Concentration of each network member is quantified by peak integration on the chromatograms. Similarly to NF network, the phase transitions from solution to particle and to fiber phases are also observed by TEM (data is presented later in this section). Therefore, a kinetic model is developed to describe the reaction schemes in three stages - solution stage (Stage 1), molten globule stage (Stage 2) and paracrystalline stage (Stage 3).

In this model (**Scheme 4-2**), *M*, *D* and *T* represent monomer, linear dimer and linear trimer, respectively. T_A is the linear trimer that has assembled into mature fiber assemblies. '$T + T_A$' represents the self-templating of unassembled trimer into fibers by preformed fibers. The top two reactions are reversible while the bottom two reactions regarding the self-assembly process are assumed as irreversible, meaning that the linear trimers assembled into paracrystalline structures are not exchanging with other members freely in solution.

$$M + M \underset{k_2}{\overset{k_1}{\rightleftharpoons}} D$$

$$M + D \underset{k_4}{\overset{k_3}{\rightleftharpoons}} T$$

$$T \xrightarrow{k_5} T_A$$

$$T + T_A \xrightarrow{k_6} 2T_A$$

Scheme 4-2 Kinetic model for NFF-CHO network. *M*, *D* and *T* represent monomer, linear dimer and linear trimer respectively. T_A is trimer that has assembled into mature fiber assemblies. $T + T_A$ represents the autocatalysis of unassembled trimer into fibers by preformed fibers.

The rate constants estimated by minimizing the sum of square error between calculated values from the model and experimental data are listed in **Table 4-1**. k_1, k_3, *and* k_6 are second-order reaction rate constants with units as mM^{-1} sec^{-1}, while k_2, k_4, k_5 are first-order reaction rate constants with units as sec^{-1}. Significant increase of k_1 and decrease of k_4 in Stage 2 (molten globule stage) suggest that the transition to a different phase enhanced the condensation reactions in particles. MATLAB regression suggests value of 0 for k_5 in Stage 2 but increased values of k_5 and k_6 in Stage 3 (fiber stage), implying that fiber formation does not occur in Stage 2 but in Stage 3, which is consistent with TEM analyses.

	Stage 1	**Stage 2**	**Stage 3**
k_1 $(mM\ s)^{-1}$	1.78(0.15)E-05	31.2(1364)E-05	0.30(0.18)E-05
k_2 (s^{-1})	6.76(1.67)E-05	177(7721)E-05	0.02(0.06)E-05
k_3 $(mM\ s)^{-1}$	1.32(2.49)E-05	0.52(0.04)E-05	0.42(1.76)E-05
k_4 (s^{-1})	32.1(63.5)E-05	0.75(0.08)E-05	2.37(13.40)E-05
k_5 (s^{-1})	N/A	0(N/A)	43.87(142.95)E-05
k_6 $(mMs)^{-1}$	N/A	N/A	1.05(30.03)E-05

Table 4-1 Kinetic parameters from best fit to NFF network in **Figure** 4-**7.**

In the first stage (**Figure 4-5**), the monomer population dropped quickly within the first 8 hours with increasing number of linear dimers and timers, indicating the fast condensation reaction led to higher degree of oligomers in the aqueous solution. After 3 hours, the linear dimer is the

dominant species in the network at the expense of monomers while linear trimers remained low at this stage. Direct TEM analysis of the network in the time course clearly show the appearance of small particles with diameter size of 5-7 nm (**Figure 4-6B**) at 6 hours, implying that the transition from liquid solution phase to particle phase possibly involved linear dimers whose concentration are highest at this stage.

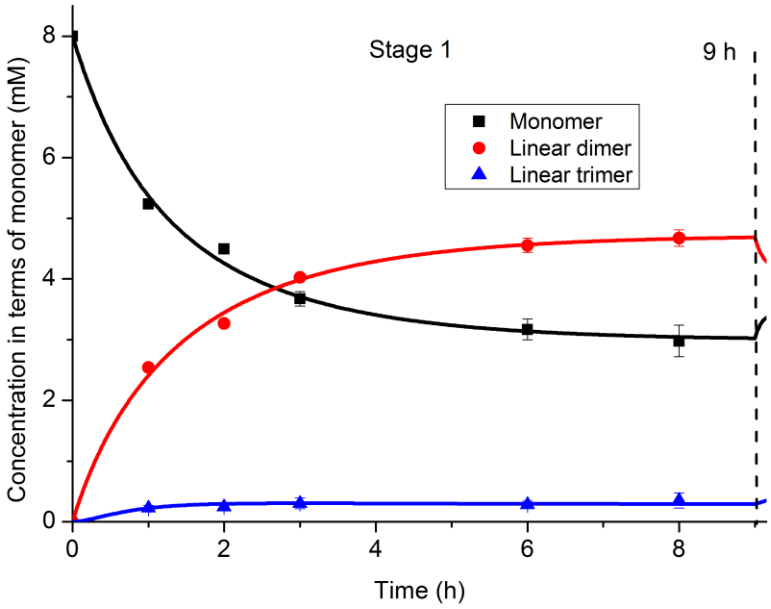


Figure 4-5 Kinetics of NFF network members within the first 9 hours in Stage 1 of network maturation (based on 3 sets of data).

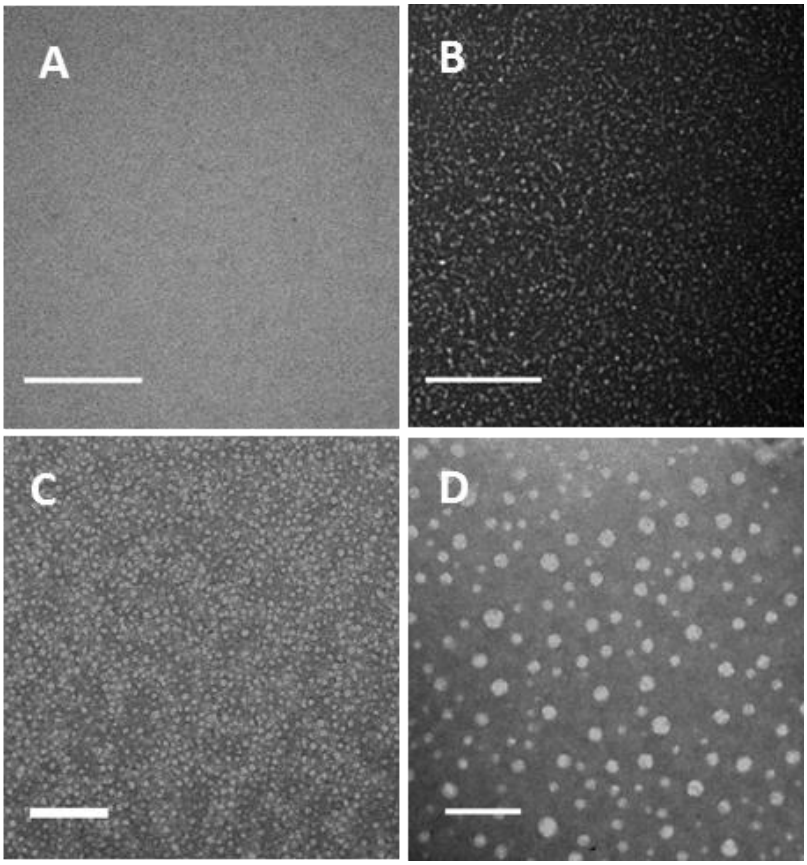


Figure 4-6 Transmission electron micrographs of samples taken directly from the NFF-DCN reaction network at (**A**) 2 hrs, no noticeable particles (**B**) 6 hrs, particles with 5-7 nm diameter (**C**) 15 hrs, particles with 10-15 nm diameter and (**D**) 19 hrs, particles with 20-45 nm diameter. Scale bar: 200 nm.

In Stage 2, a substantial growth of linear trimers is noticed (**Figure 4-7**), which is concomitant with the increase in particle sizes from 5-7 nm diameter at 6 hours to 10-15 nm diameter at 15 hours and 20-45 nm diameter at 19 hours (**Figure 4-6 C, D**). In the latter half of the second stage from 36 hours to 60 hours, although there is no obvious change in the distribution of network members, the size of the particles in the network increased from the range of 44-120 nm

diameter at 24 hours (**Figure 4-8 A**) to 100-165 nm diameter at 48 hours (**Figure 4-8 B**), meanwhile the amount of particles decreased, similarly to Ostwald ripening (Baldan, 2002; Ostwald, 1901; Wang et al., 2014). Large particles grow at the expense of smaller ones due to the driving force to lower total surface area and minimize the energy in the system. The change of mean particle size in NFF-DCN is in a rough linear relationship with time (**Figure 4-9**), which is commonly seen in Ostwald ripening (Lifshitz & Slyozov, 1961; Wagner, 1961; Wang et al., 2014).

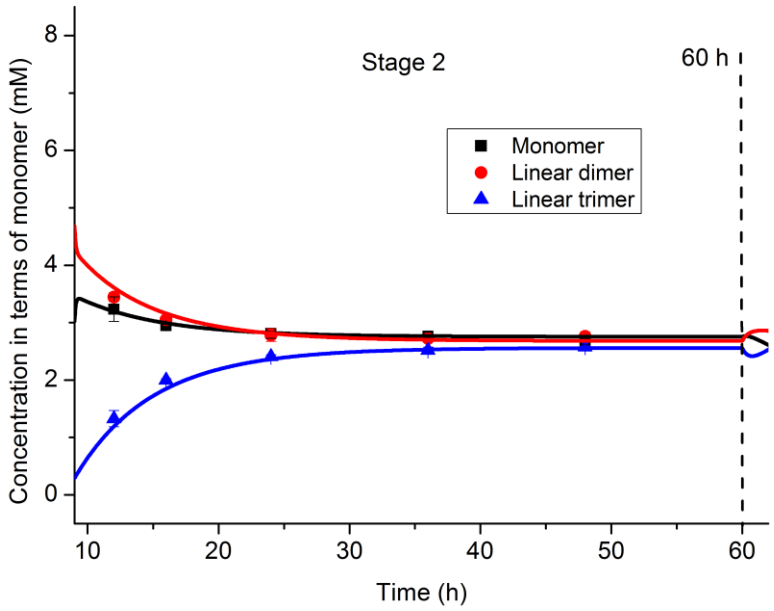


Figure 4-7 Kinetics of NFF network members in Stage 2 (from 9 hours to 60 hours) of network maturation (based on 3 sets of data).

At the end of the second stage, the emergence of paracrystalline structures (shown as white fiber-like structures) inside of the big particles at 48 hours observed by TEM (**Figure 4-8 B**) indicates that the association, possibly the hydrogen bonding of building blocks drive their packing geometry towards a more organized pattern from randomly dispersed fluid-like oligomers in particles (Cheon et al., 2007) as shown in the phase transition model below (**Figure 4-10**). In this

phase transition model, the red fibril represents the free peptide building blocks, which formed molten globules with fluid-like oligomers inside. The blue fibril represents the paracrystalline structures. Once the ordered paracrystalline structures reach the threshold to grow beyond the particles, another transition occurs from molten particle phase to paracrystalline phase in solution, where the ordered paracrystalline nuclei serves as templates for adsorption of soluble peptides at the template end (Stage 3). It can be seen clearly that appearance of abundant fibers around the aggregation of particles at 72 hours (**Figure 4-8 C**).

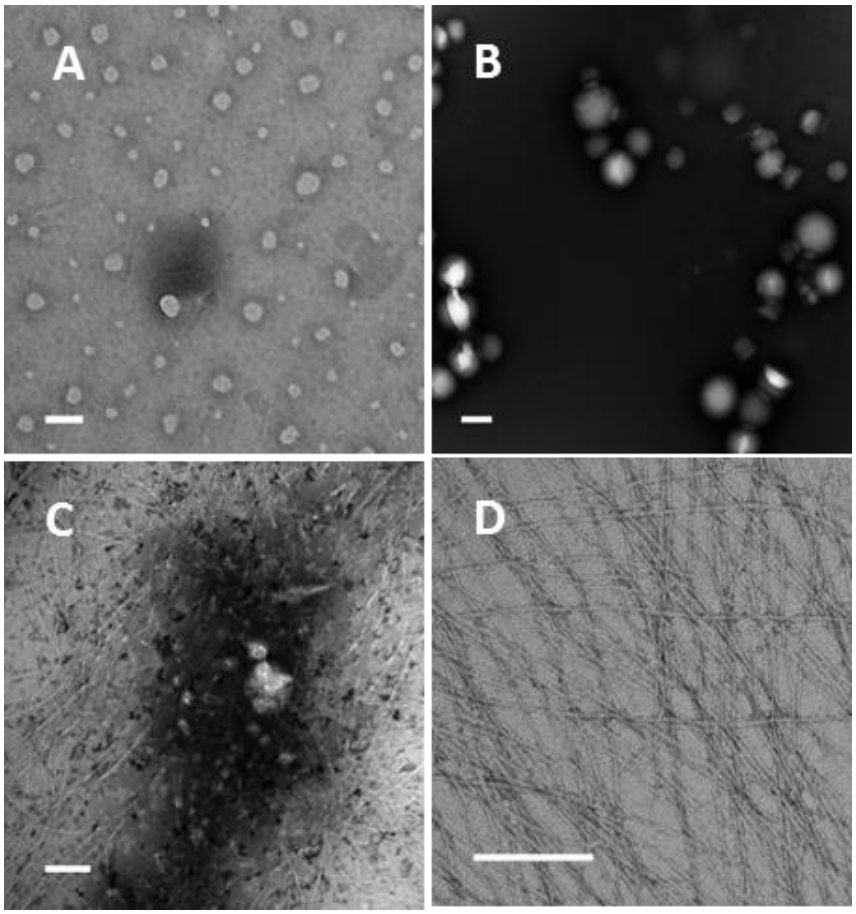


Figure 4-8 Transmission electron micrographs of samples taken directly from the NFF-DCN reaction network at (**A**) 24 hrs, particles with 44-120 nm diameter (**B**) 48 hrs, particles with 100-165 nm diameter (**C**) 72 hrs, and (**D**) 96 hrs fibers with 5-8 nm width. Scale bar: 200 nm.

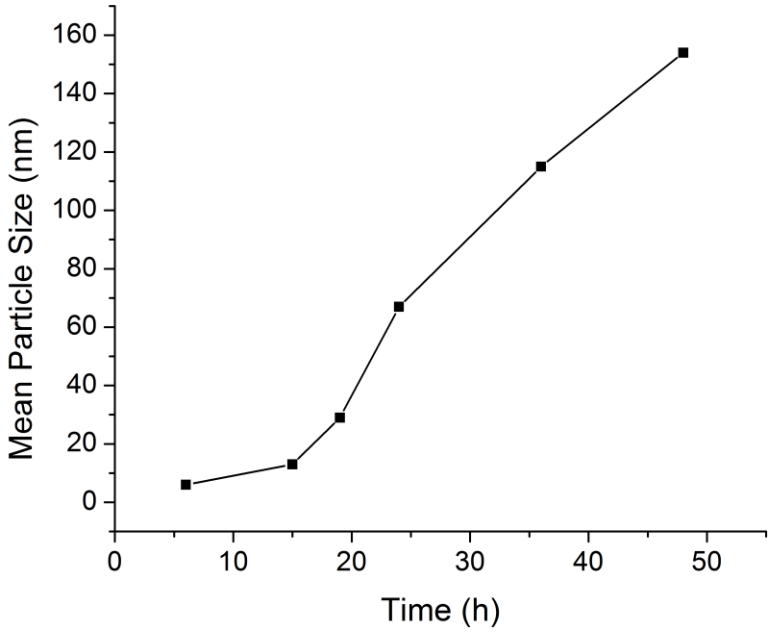


Figure 4-9 Mean particle size in NFF dynamic network vs. time

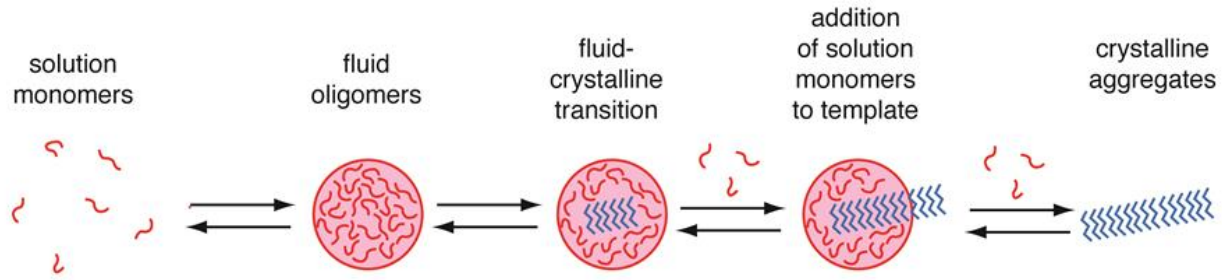


Figure 4-10 Model for the phase transition from monomeric state to disordered oligomer to an ordered crystalline state. Red denotes unstructured peptides and blue represents the transition to the more crystalline structures (Childers, 2010).

In the third stage, the linear trimers demonstrate another substantial growth shown after 60 hours with a considerable drop of monomer and linear dimers (**Figure 4-11**). After 1 week, the linear trimers are completely dominant in the networks at the cost of other species. Meanwhile, fibers gradually became the prevalent morphology observed on TEM grids after 4 days (**Figure 4-8 D**) while particles are gradually consumed by fibers growth and eventually disappeared. These evidences above suggest the network member for constituting the supramolecular structures might be the linear trimer.

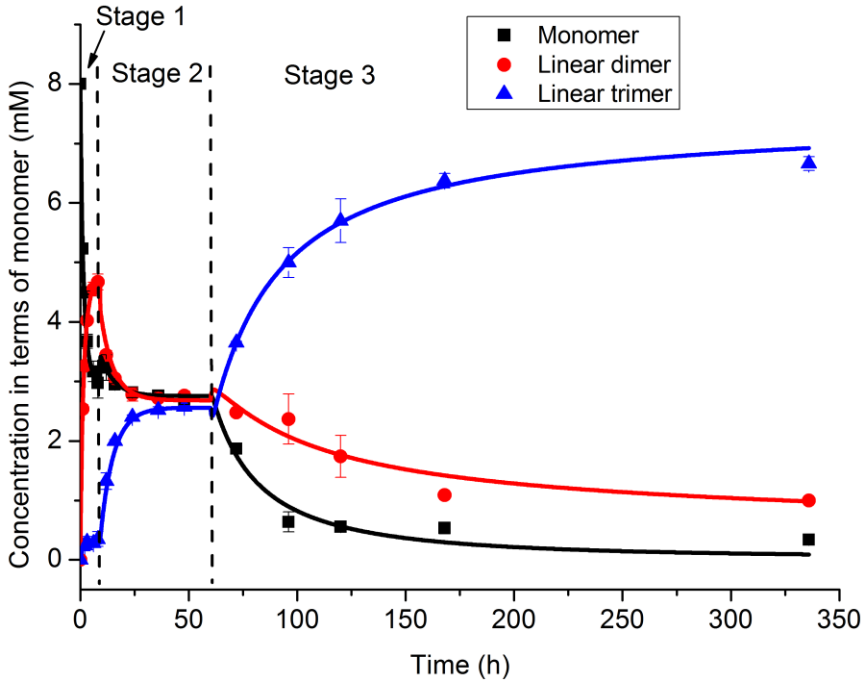


Figure 4-11 Kinetics of NFF network fitted into a model with 3 stages. Concentrations of each network member at different time points are calculated by peak integration on the HPLC chromatograms.

This kinetic model is sufficiently robust to fit change for each stage of kinetics in NFF-DCN and describe phase transitions leading to hierarchical molecular order construction in the network. Meanwhile, observation from direct TEM analysis confirms the prediction and helps explain the mechanism for network distribution change. The TEM data and HPLC kinetics corroborate each other well in this system.

4.3.5 Identification of Assemblies in the NF-DCN

The phase transition from particles to fibers is accompanied by the increase of linear trimers, suggesting linear trimers as the component of the fiber assemblies. The fibers in the NFF network at 96 hours are enriched by centrifugation at 16,000 × g for 30 minutes, re-suspended in water/acetonitrile (3/2, v/v), and analyzed by HPLC. Compared to the network before centrifugation (**Figure 4-12, top**), the chromatogram of the spun-down species is predominately linear trimers (**Figure 4-12, bottom**), even though at 96 hours ~ 1mM dimers remained in the network. As centrifugation (16,000 × g for 30 minutes) of network samples at early time points before the appearance of fibers did not enrich any network members, we reason that the linear dimers might exist in solution or particles which cannot be spun down at the above conditions and only the fibers are spun down.

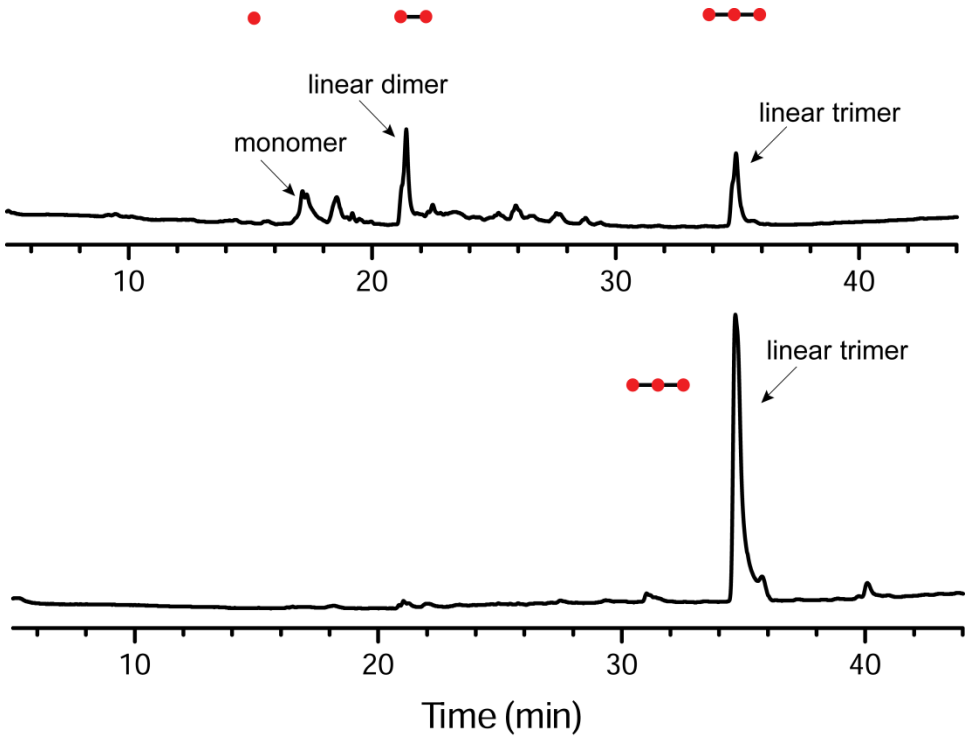


Figure 4-12 HPLC analyses of NFF-DCN (**top**) NFF-DCN prepared in 60% water/40% CH_3CN after 2 days at 25°C gradient elution from 10-90% MeCN with monomer, linear dimer, cyclic dimer, linear trimer, cyclic trimer assigned, and (**bottom**) following sedimentation at 16,000 × g, the pellet is re-suspended in 60% water/40% CH_3CN and analyzed under the same HPLC elution conditions. HPLC UV absorbance is monitored at 254nm; flow rate is 1.0 mL/min with gradient from 10% acetonitrile to 90% acetonitrile in 40 minutes.

4.3.6 Characterization of Fiber Assemblies in the NFF-DCN

Having verified the components of fibers in the NFF-DCN, next I investigated the nature and characteristics of these assemblies. The fibers themselves appear amyloid-like by TEM (**Figure 4-13, left**), with fiber width distribution of 6-8 nm. Since the building blocks of fibers are apparently altered from peptide amide backbone to reversible acetal linkages, the question arises

as to would these assemblies still exhibit typical β-sheet peptide signatures? Thioflavin T (ThT) binding assay is carried out to test the above question. ThT is a defining probe for β-sheet peptide assemblies as it displays increased fluorescence in quantum yield and a characteristic red shift of its emission spectrum upon binding to those assemblies (LeVine, 1999) (Wolfe, 2010). When the ThT dye solution is titrated into the 2.5mM NFF-DCN (with molar ratio 1:10), fluorescent fibers are observed (**Figure 4-13, right**) and an enhanced fluorescence and a slight redshift in the emission spectrum is detected (**Figure 4-14**), suggesting the possibility that the assemblies of linear trimers with acetal linkages still form β sheet like structures.

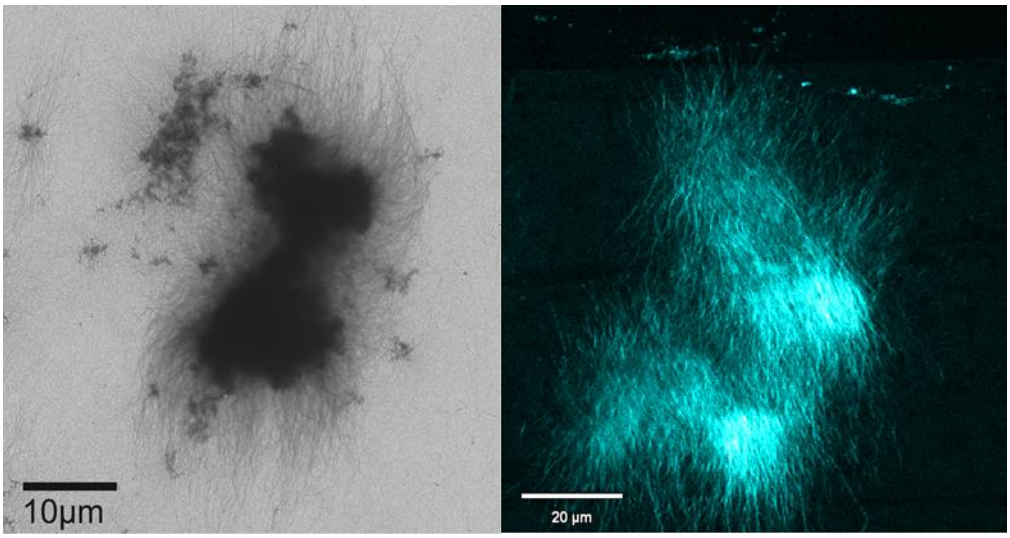


Figure 4-13 TEM image of fibers in NFF-DCN, scale bar: 10 μm (**left**). Fluorescence microscopy image of fibers in NFF-DCN upon binding with Thioflavin T (ThT), scale bar: 20 μm (**right**)

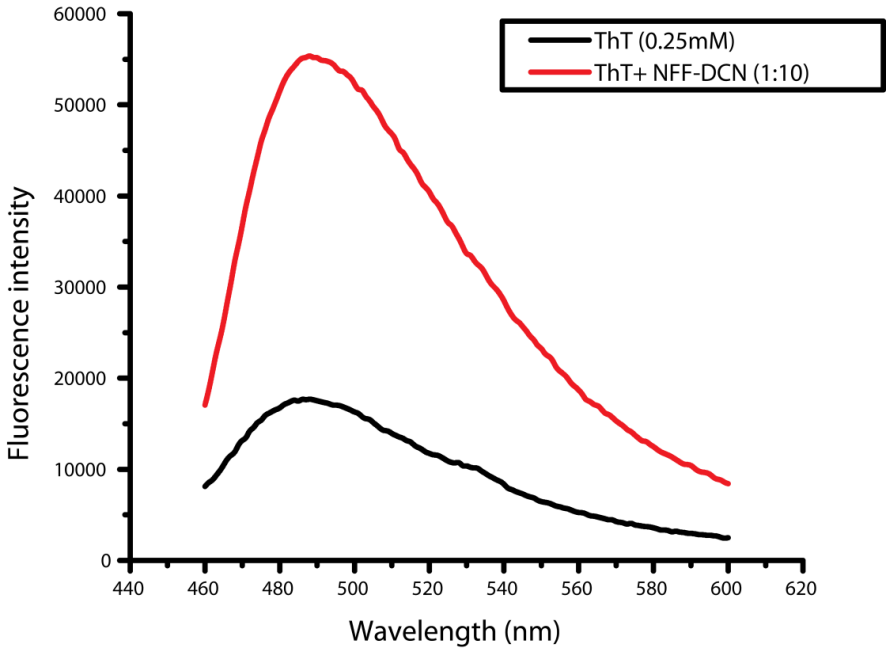


Figure 4-14 Fluorescence intensity of NFF-DCN. (**red**) *NFF*-DCN (2.5 mM) upon binding with ThT (0.25 mM); (**black**) ThT alone (0.25 mM).

For further investigation, the powder X-ray diffraction (XRD) analysis is used to further characterize the fiber assemblies. A *d*-spacing of 4.5± 0.5 Å is observed (**Figure 4-15**), which is diagnostic of the peptide inter-strand hydrogen bonded β-sheet repeat distance in amyloid assemblies (Mehta et al., 2008). A *d*-spacing of 19.9 ± 2.9 Å may correspond to the length of linear trimers.

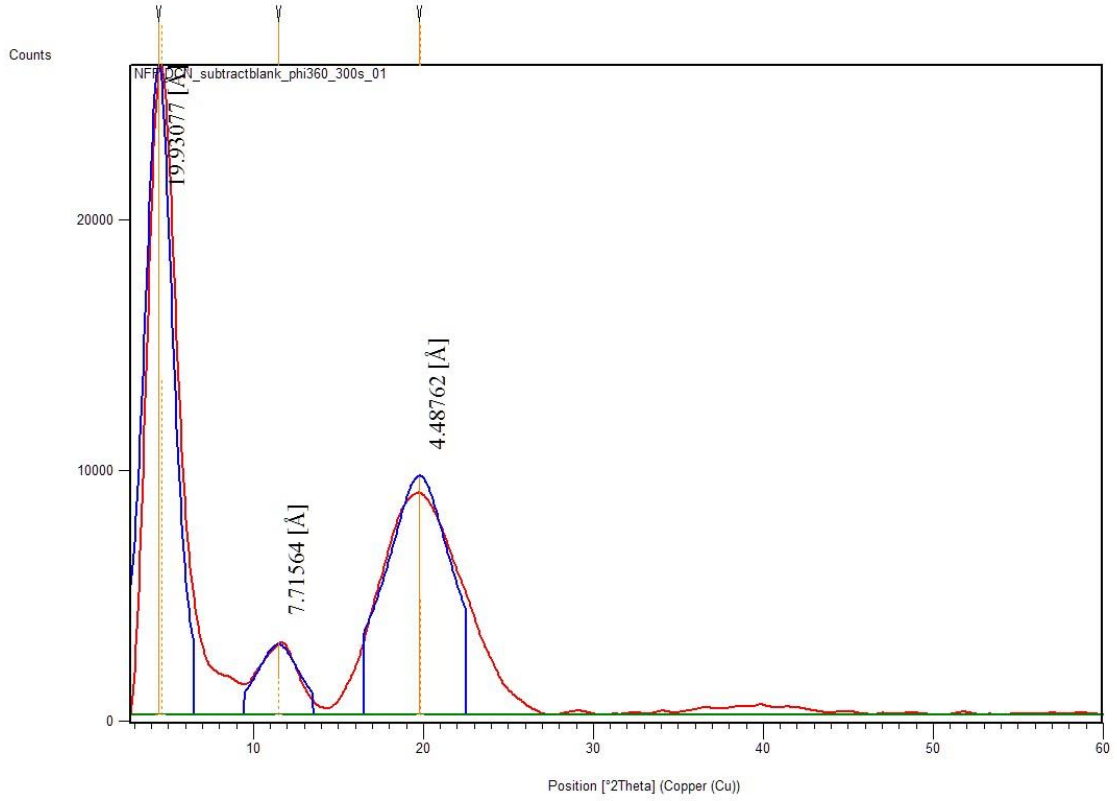


Figure 4-15 X-ray powder diffraction (XRD) patters of NFF-DCN fibers. The *d*-spacing of 4.5± 0.5 Å is diagnostic of the peptide inter-strand hydrogen bonded β-sheet repeat distance in amyloid like assemblies.

Taken together, the presence of intermediate particle phases (Childers et al., 2012; Yan Liang et al., 2010), the observation of β-sheet *d*-spacing, ThT binding, and the preferential selection of the linear trimer from NFF-DCN is consistent with β-sheet structures for these assembled phases.

4.3.7 Supramoleclar Assemblies' Responsiveness to pH Change

The responsiveness of the network to pH changes is not only reflected in the oligomers at macromolecular level (**Figure 4-4**), but also demonstrate in the assemblies at the supramolecular level. At quasi neutral pH, e.g. pH 5.8, fibers dominate the morphology of assemblies in the NFF network (**Figure 4-16, left**). However at acidic conditions < pH 2.5 or basic conditions > pH 8, particles are the dominant morphology observed in the networks (**Figure 4-16, right**). This is likely because that at quasi neutral conditions, the N, N-acetal linkage is stable so linear trimers crossed the critical micelle concentration self-assemble to fibers. When pH falls out of the range 3-7, the N, N-acetal ring would open, which results in hydrolysis of oligomers to monomers. High concentration of monomers would aggregate to particles after three days.

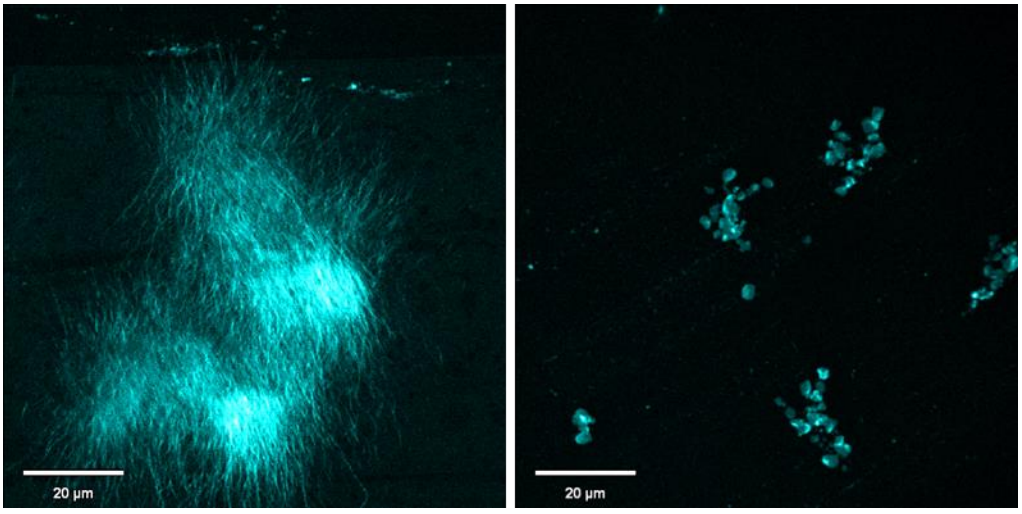


Figure 4-16 Fluorescence microscopy images of fibers at pH 5.8 , 6 days (**left**) and particles at pH 8.2, 6 days (**right**) in NFF-DCN.

4.4 Discussions

Similarly to NF network, the kinetics of NFF network aging is divided into three stages: solution stage (Stage 1), particles stage (Stage 2) and fiber stage (Stage 3). Transition from Stage 1 to Stage 2 occurred *ca*. 9 hours and the transition from Stage 2 to Stage 3 occurred *ca*. 60 hours. Although as early as 6 hours small particles are already observed (**Figure 4-6B**), those particles are probably aggregates of linear dimers instead of linear trimers as the concentration of linear trimers remain low and unchanged from 6 hours to 9 hours. Meanwhile, monomers are partitioning into the linear dimer particles. The ones incorporated into the particles condensed with linear dimers to linear trimers. The first exponential growth of linear trimers starting *ca*. 9 hours (**Figure 4-7**) suggests the emergence of particle phase impact the growth of oligomer chain length. As the equilibrium constant changes in the desolvated environment inside the particles, the condensation for linear trimers are more favored compared to the solution phase. In the first half of Stage 2, increases in the particle size from 10-15 nm to 20-45 nm (**Figure 4-6C, D**) together with the increase in linear trimer concentration, suggest that network members are kept being incorporated into particles and linear trimers are condensed within particles. In the second half of Stage 2, increases of particle size is observed without obvious change in network member distribution, which implies that the distribution of network members between liquid phase and particles phase reaches equilibrium so the size increase in particles is not driven by recruiting more network members from solution to particles as seen earlier but by gradually fusing of small particles into big particles, a thermodynamically-driven process to minimize the total surface area, resulting in shrinking and disappearance of small particles and growing of big particles. It is similar to molten globules seen in proteins (Muschol & Rosenberger, 1997) and in the pathway

for amyloid cross-β peptide paracrystalline assemblies (Childers et al., 2012; Dong et al., 2006; Yan Liang et al., 2010).

The first sigmoidal growth of linear trimers from 0 to 48 hour resulted from the emergence of the particle phase, which provided desolvated environment for condensation. The second sigmoidal growth of linear trimers from 50 to 336 hours associated with the emergence of fiber phase. In the third stage, the paracrystalline fibers released from the particles into solution, which provided templates for free linear trimers in solution to add at the ends of fibers, driving equilibrium towards linear trimer. Strand breaking of fibers due to collision, shearing force resulted in the exponential growth of linear trimer at the expense of other network members.

4.4.1 Comparison of the Kinetics of NFF-DCN to NF-DCN

Both the NF and the NFF dynamic network exhibit the same phase transitions from solution phase to particle/molten globule phase, and then to para-crystalline phase progressively, which dictate oligomerization kinetics and thermodynamic distributions of network members.

However, the main difference of NFF network from the NF network is that no obvious cyclic dimer is observed in NFF network. Is it because the linear dimer in NF network is easier to cyclize than the one in NFF network? Molecular Dynamic (MD) results of linear dimer (H-NFaNFaNF-CHO) in NF-DCN (**Figure 4-17 A**) and linear dimer (H-NFFaNFFaNFF-CHO) in NFF-DCN (**Figure 4-17 B**) indicate that both of them display "horseshoe" structures, with similar distance of two reactive ends (labeled as purple dashed lines), 8.238 Å, 7.800 Å for H-NFaNFaNF-CHO and H-NFFaNFFaNFF-CHO, respectively. Although the distances do not differ widely, the linear dimer in NF-DCN seems to have greater flexibility than the one in NFF-

DCN as the RMSD (Root-mean-square deviation) results suggest that H-NFFaNFFaNFF-CHO packs a little more tightly than H-NFaNFaNF-CHO.

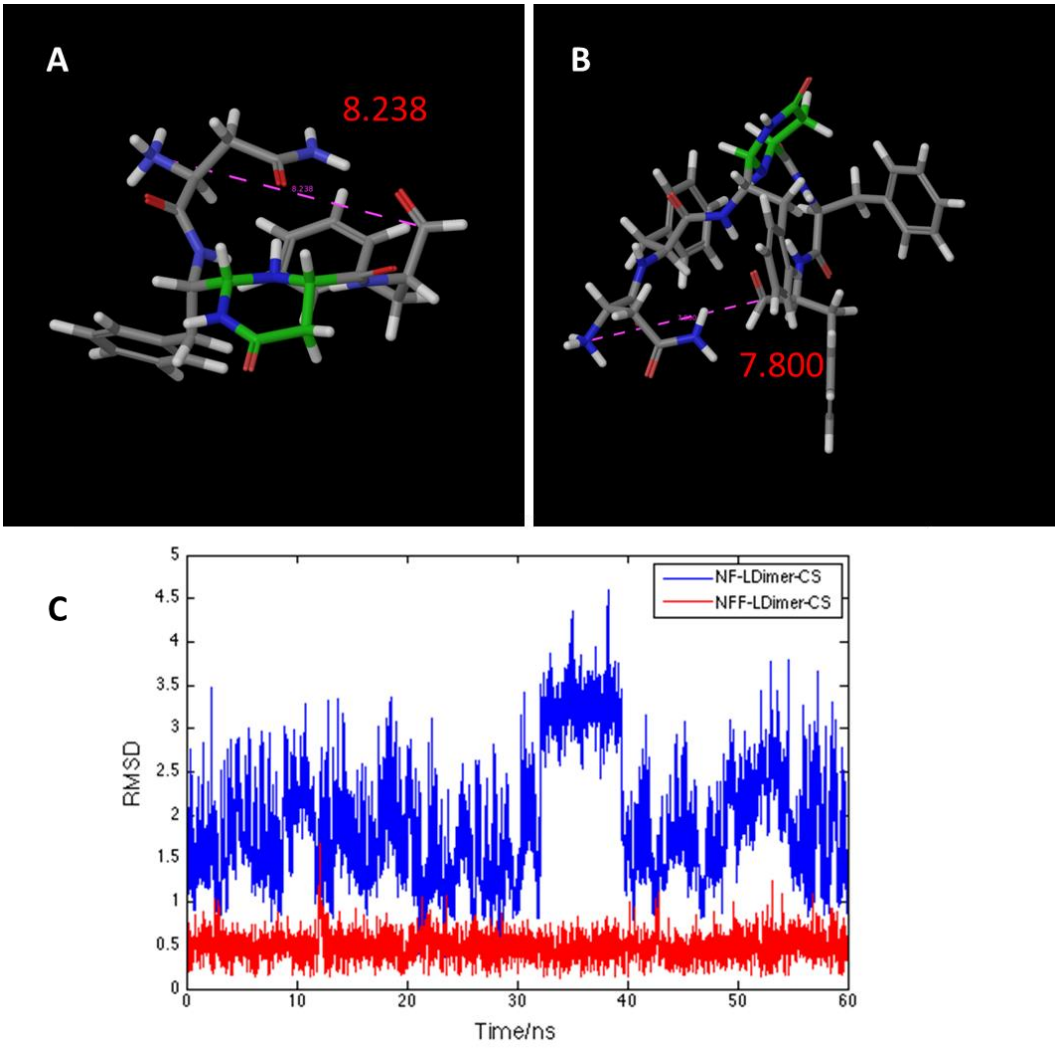


Figure 4-17 Molecular Dynamic results of linear dimer in NF-DCN (**A**) and NFF-DCN (**B**), respectively. RMSD results **(bottom)** of linear dimer in NF-DCN (blue) and NFF-DCN (red).

The simulations imply that both networks have the similar tendency to generate cyclic dimer and indeed cyclic dimers are observed in NFF-DCN within 1 hour by ESI mass spectrometry. But

why did cyclic dimers not observed at later time points? We reason that due to the additional phenylalanine residue in the building blocks of NFF-DCN, the more hydrophobic members of NFF network experienced more rapid hydrophobic collapse, which drove network members in NFF-DCN more quickly into small particles, converting linear dimers to linear trimers and the Phe-Phe stacking/ association stabilize the linear structure and prevent it from cyclization. Indeed, we observed faster transition from liquid phase to particle phase in NFF-DCN.

4.4.2 Similarity in Second Order Structures in NFF-DCN and NF-DCN

CD spectra of fibers from both networks (**Figure 4-18**) display a minimum band near 190 nm and two positive bands near 200 nm and 225 nm, most consistent with previously reported β-turn structures (Gao et al., 2002; Surewicz & Mantsch, 1988). The FTIR absorption from both networks (**Figure 4-19**) displays a maximum around 1670 cm^{-1} , characteristic of β-turn conformation (Adochitei & Drochioiu, 2001; Surewicz & Mantsch, 1988). Both IR and CD data of the network assemblies seem to be consistent with β-turn signatures within β-sheet arrays, suggesting a type I β-turn, consistent with asparagine stabilization (Wu, McElheny, Setnicka, Hilario, & Keiderling, 2012), but further structural characterization is required.

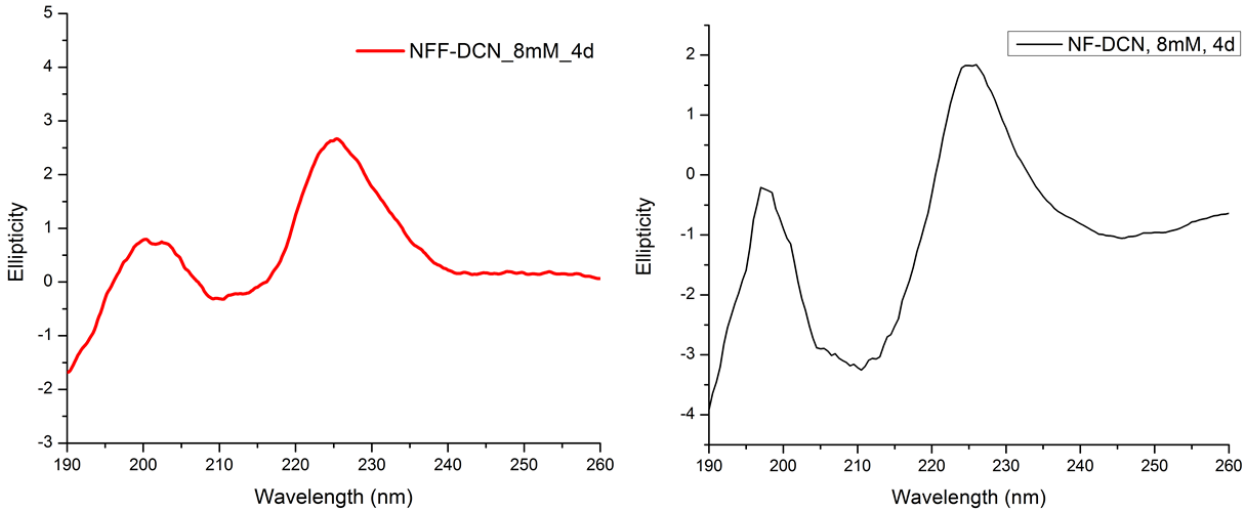


Figure 4-18 CD spectrum of fibers from NFF-DCN (left) and NF-DCN (right). Both networks are incubated at 40% acetonitrile/ 60% water, pH 4.

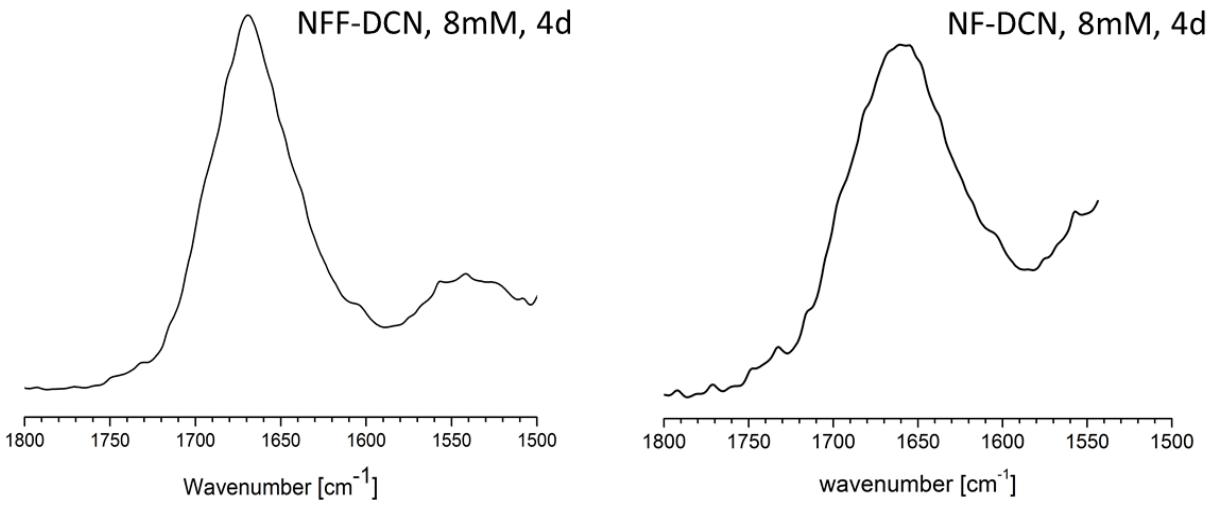


Figure 4-19 FT-IR spectra of fibers from NFF-DCN (left) and NF-DCN (right). Both networks are incubated at 40% acetonitrile/ 60% water, pH 4. The FTIR absorption from both networks displays a maximum around 1670 cm^{-1}, characteristic of β-turn conformation

4.4.3 Phase transitions lead to progressive growth in NFF-DCN and NF-DCN

These NF and NFF dynamic networks follow a phase transition-based assembly that appears correlated to the size of the initial particles. The observed dynamics are similar to those seen in emulsion polymerization, but these particle are composed of highly functionalized peptides and are more similar to intramolecular molten globule intermediates in protein folding and the intermolecular particles of amyloid assembly (Auer, Ricchiuto, & Kashchiev, 2012; Childers et al., 2012). The metastable nature of the particles (Anthony, Mehta, Lynn, & Berland, 2014; Childers et al., 2012; Yan Liang et al., 2010; Y. Liang, Pingali, et al., 2008) creates a tension between physical particle maturation and oligomer polymerization that allows for selection of a specific chain length. We propose that stage 1 and 3 occur largely in solution and that the cyclic intermediates that form in the NF-DCN do not partition into the particle phase but are templated as linear trimers once the fiber ends extend beyond the particles. Simulations show little thermodynamic differences in NF-CHO and NFF-CHO dimer cyclization propensity (**Figure 4-17**), consistent with cyclization of the more hydrophobic NFF-CHO being constrained by ordering in the particle phase.

It is the combination of chemical and physical dynamics that achieve a progressive growth of molecular order and allow for that autocatalytically amplification of chain-length specific peptide assembly.

Chapter 5: Understanding the pathway of assemblies in Chemical Networks and their responsiveness to templates

5.1 Introduction

As discussed in the previous chapters, the assemblies that emerged in the NF and NFF dynamic chemical networks with the reversible acetal linked oligomers demonstrate properties like β-sheet rich structures. The appearance of particles in both the NF and NFF networks preceding the emergence of mature fibers (**Figure 3-6**, **Figure 4-6**), resembles amyloid peptide assembly. Although the network's assembly growth coupled with acetal condensation reactions adds more complexity, it still demonstrates similar nucleation-dependent processes seen in the formation of amyloid fibers, which is commonly considered as a nucleation-dependent process (Childers et al., 2012; Harper & Lansbury, 1997; Jarrett & Lansbury, 1993; Yan Liang et al., 2010; Naiki & Gejyo, 1999). To better understand the assembling process in the dynamic networks, the formation of amyloid fibrils serves as a model system.

5.2 Two Steps of Amyloid Peptide Self-assembly

The amyloid peptide self-assembly process consists of two steps: (i) a lag phase (nucleation phase) where the peptide monomers forming the more ordered nuclei is the rate-limiting step, and (ii) a growth phase (elongation phase) which is exponential due to the addition of monomer at the continually breaking fiber ends (C. Liang et al., 2014). The kinetics is therefore sigmoidal

(**Figure 5-1, green curve**) with an initial lag phase. Addition of preformed seeds shortens the lag phase (**Figure 5-1, red curve**) by providing the initial templates.

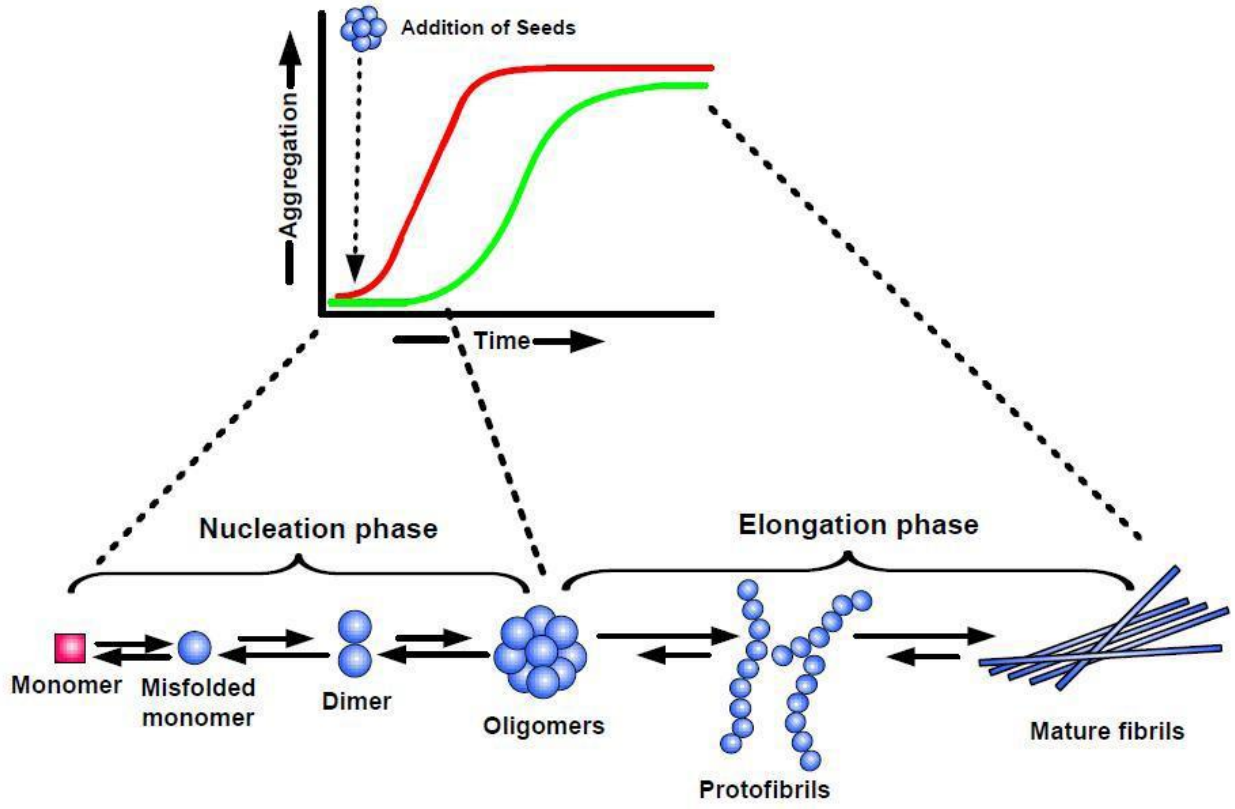


Figure 5-1 Nucleation-dependent process of amyloid peptide aggregation. The kinetics of amyloid aggregation consists of two phases: the lag phase (nucleation phase) and the growth phase, which can be represented by a sigmoidal curve with an initial lag phase (green curve). Adding preformed seeds will substantially reduce the lag time (red curve) (Kumar & Walter, 2011).

5.2.1 Are Oligomers on Pathway to Mature Fibers?

Even though the two steps in self-assembly have been observed in many cases (Auer et al., 2012; Kumar & Walter, 2011), the exact mechanisms of how amyloid peptides nucleate and grow remain elusive. One of the challenges for probing the pathway of assembly is to identify and characterize transient oligomeric intermediates en route to larger particle aggregates and fibrils (Carulla et al., 2009). In the lag phase (peptide nucleation phase), there has been extensive debate regarding whether oligomerization is the necessary pathway preceding the emergence of fibers. Small molecular inhibitors of Aβ (1-42) that block oligomerization are shown to have no effect on inhibiting fibrillization, suggesting oligomerization might not be an indispensable pathway to fibers (Necula et al., 2007).

However, multiple other studies (Blake & Serpell, 1996; Sikorski, Atkins, & Serpell, 2003; Sunde et al., 1997) have suggested that peptide monomers may form soluble oligomeric particles in the lag phase prior to self-assembling into fibers. In microscopic imaging, such as TEM and AFM, particle species are observed (Harper & Lansbury, 1997; Stine, Dahlgren, Krafft, & LaDu, 2003) and have been proposed to be peptide oligomers. A fluorescence microscopy study also reveals that KLVFFAE peptides indeed generated micron sized aggregates that are directly observed to serve as exclusive nucleation sites for amyloid nanotube assemblies shown in **Figure 5-2** (Yan Liang et al., 2010). However, these microscopic images cannot provide convincing evidence for on-pathway oligomerization of peptides and other analytical methods need to be utilized for the early stage characterization in self-assembly.

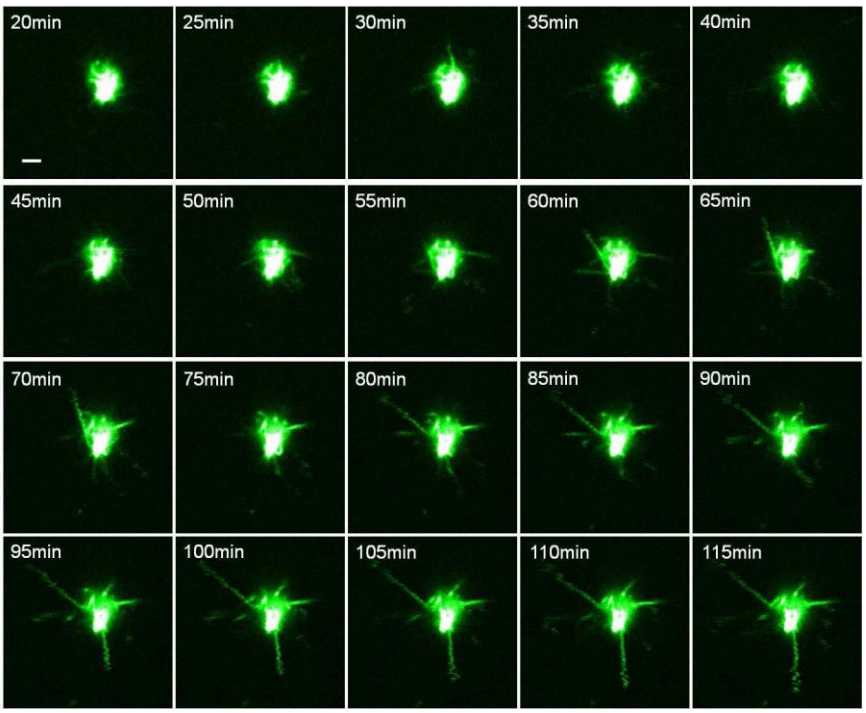


Figure 5-2 Fluorescence microscopy study of assembly of rhodamine-17-22/ KLVFFAE (1:150) (Yan Liang et al., 2010). Time lapse fluorescence images of aggregates with emerging tubes, Scale = 2μm.

5.2.2 IMS-MS Technique for Probing the Oligomeric Intermediates

Most biophysical techniques employed in the study of peptide assemblies, including CD, FTIR spectroscopy and fluorescent microscopy are limited to providing data of a global average of multiple, rapidly inter-converting species being co-populated in solution during the early stage. In identifying and characterizing transient oligomeric intermediates, mass spectrometry has been shown to be able to preserve non-covalently bound species (Loo, 1997; Smith & Light-Wahl, 1993). ESI-MS data collected for several protein systems are consistent with solution phase binding constants. The solvent evaporation in transferring the noncovalently bound complex

from solution to the gas phase is argued to be similar to the desolvated environments the oligomers experience in solution phase. Modern ion mobility spectrometry-mass spectrometry (IMS-MS) can resolve complex mixtures of species even at femtomolar concentrations, including transiently populated states (Bernstein et al., 2009; Vlad, Iurascu, Slamnoiu, Hengerer, & Przybylski, 2012; Woods et al., 2013). IMS-MS has been demonstrate for probing the oligomerization pathways of amyloid-β (Aβ) and α-synuclein peptide and oligomeric species up to 16-mer and 12-mer of Aβ40 and Aβ42 have been identified (Bernstein et al., 2009; Kloniecki et al., 2011). A recent study has utilized IMS-MS to monitor oligomer formation from both hIAPP and rIAPP (human amylin and rat amylin) (**Figure 5-3 a**). The IMS-MS driftscope plot shown in **Figure 5-3 b** suggests that oligomers up to hexamers are formed within 2 minutes of dilution of hIAPP into buffer (Young, Cao, Raleigh, Ashcroft, & Radford, 2014).

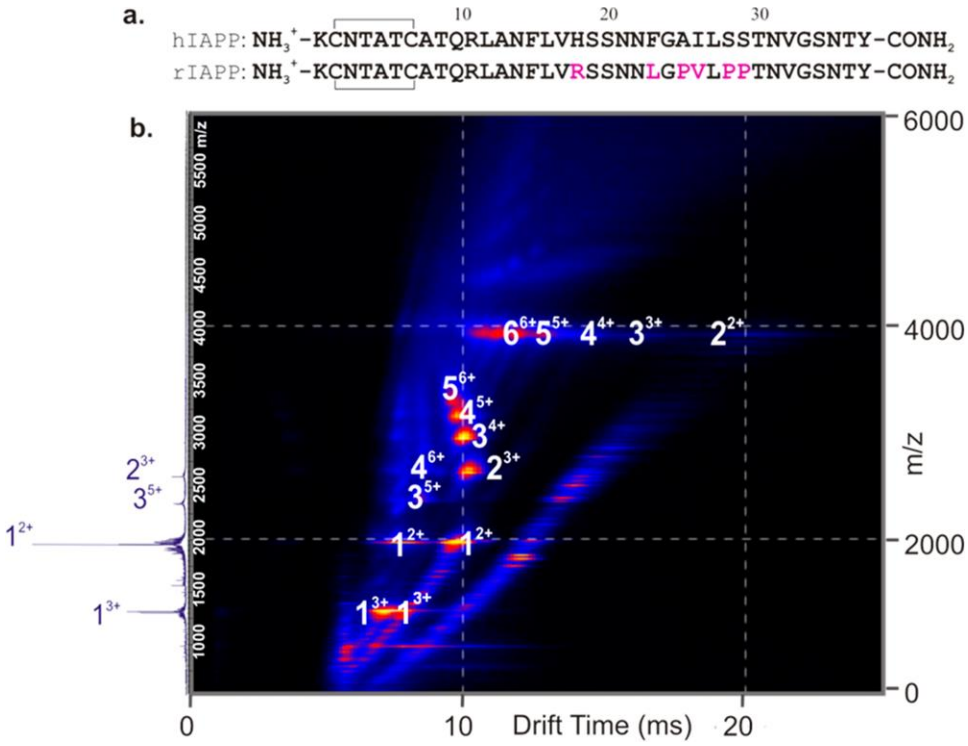


Figure 5-3 hIAPP forms an array of oligomeric species during fibril formation. (**a**) Comparison of hIAPP and rIAPP sequences. Difference in the rat sequence from those of the human peptide is labeled in pink. (**b**) ESI-IMS-MS driftscope plot of the hIAPP oligomers present 2 min after diluting the monomer to a final peptide concentration of 50 μM in 20 mM ammonium acetate. Driftscope plots show IMS drift time versus m/z versus intensity, and the corresponding mass spectrum is shown on the left-hand side. Numbers adjacent to peaks denote oligomer order, with the positive charge state of each oligomer ions in superscript. (Young et al., 2014).

When gaseous ions pass through travelling-wave ion mobility spectrometry (TWIMS) with a drift-tube filled with a buffer gas (e.g. helium), they are accelerated by an electric field and the mobility is governed by the mass to charge ratio (m/z), size and shape, which results in an ion's characteristic collision cross-section (CCS) (**Figure 5-4 A**). The peptide monomers' separation is

primarily dictated by the individual shapes. Compact conformations with smaller CCS have shorter drift times as they experience fewer collisions, while expanded conformations have longer drift times as they undergo more collisions with buffer gas in the drift tube (**Figure 5-4 B**). On the other hand, when separating non-covalent bound oligomeric species arising from the same peptide monomer, the effect of charge is more predominant. For example, when singly charged monomer, doubly charged dimer and triply charged trimer ions all share the same m/z, but the trimer tends to have a shorter drift time than the monomer and dimer, as multiple charges will propel them through the drift cell (**Figure 5-4 C**).

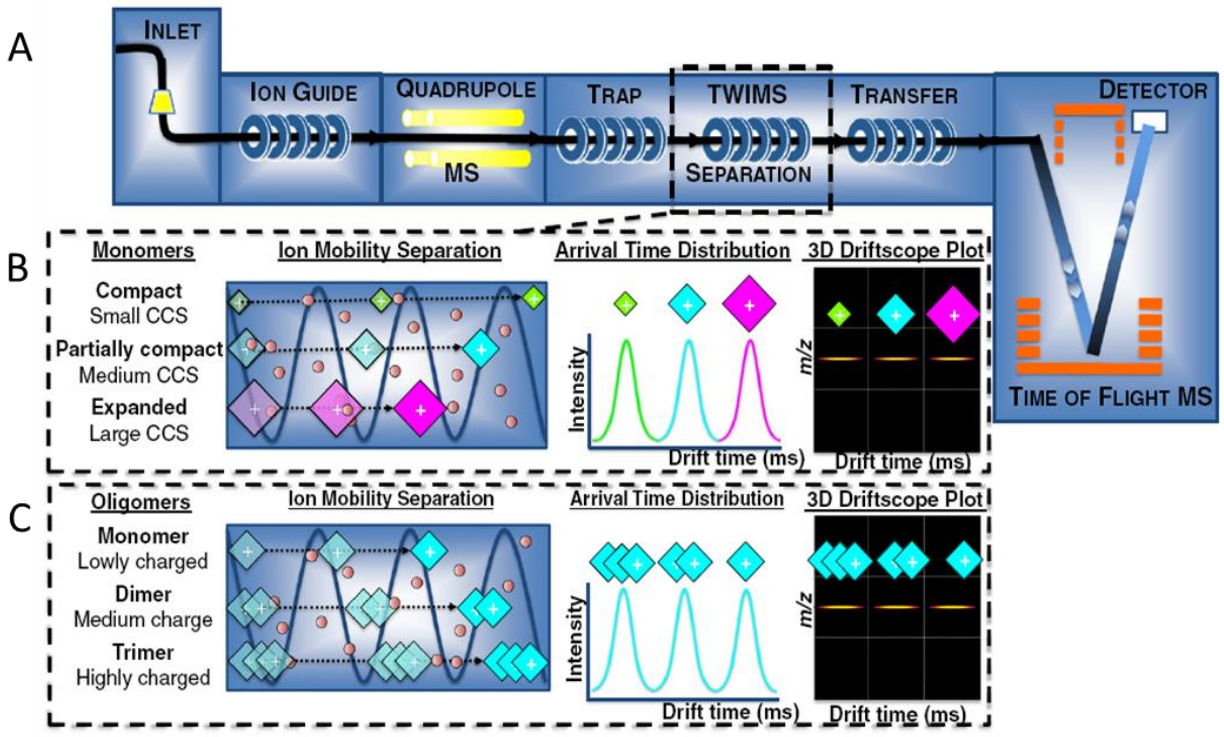


Figure 5-4 Illustration of IMS-MS. (**A**) Schematic showing travelling-wave ion mobility spectrometry (TWIMS) integrated with an orthogonal acceleration quadrupole-time-of-flight mass spectrometer. Ions are accelerated by an electric field and separated by their mobility through the drift cell. (**B**) Large ions with expanded characteristic collision cross-section (CCS)

have a longer drift time as they undergo more collisions with the buffer gas in the drift tube, whereas the smaller, more compact ions of the same molecular mass experience fewer collisions and traverse the drift tube in a shorter time. (**C**) Ions of monomer, non-covalent associated dimer, trimer and other oligomers with the same m/z but different mass and charges can be separated based on both their CCS and their number of charges. Typically, for a given m/z, the more highly charged ions are accelerated through the drift cell faster and have a shorter drift time. (Figure adapted from (Pringle et al., 2007; Woods et al., 2013))

Therefore, peptide species with different charges are typically separated into regions or bands shown in a two-dimensional drift time (t_D) to *m/z* plot, where multiply charged species have shorter drift time. The singly, doubly, and triply charged peptide ion families are clearly separated and marked with the white lines (**Figure 5-5**). Within these bands, drift times generally increase with mass. Based on previous theoretical and experimental studies, IMS-MS should detect the oligomeric state of network assemblies, and the data will be discussed in the results session.

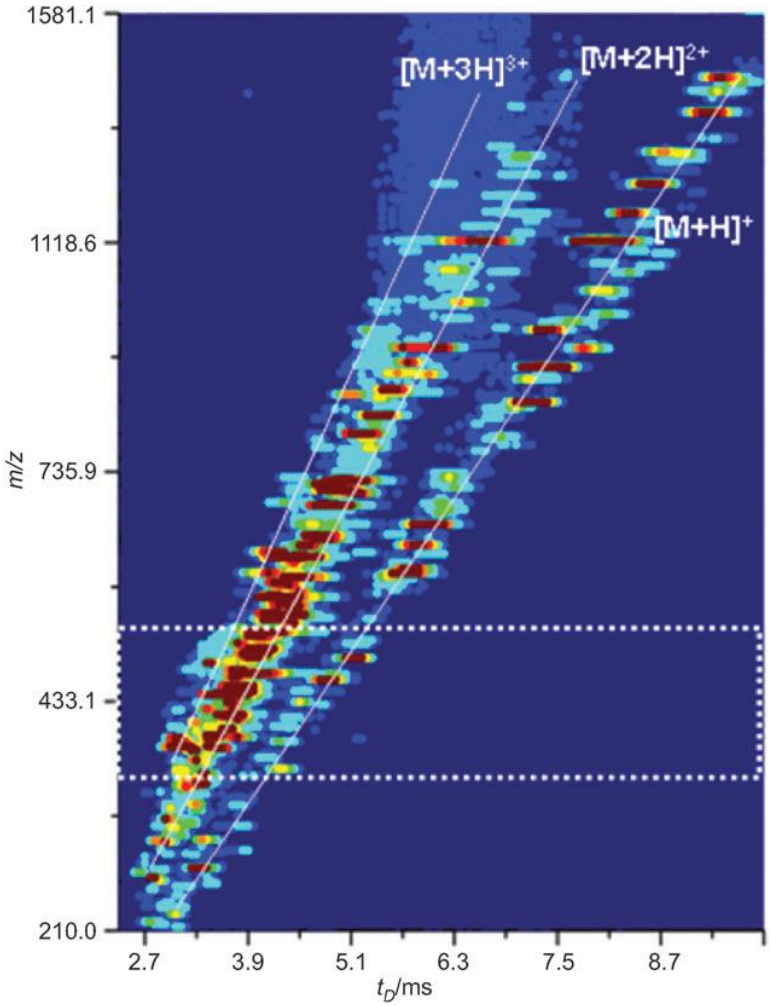


Figure 5-5 Plot of drift time (tD) to m/z by IMS-MS method, where the singly, doubly, and triply charged peptide ion families are clearly separated and marked with the white lines (Valentine et al., 2006).

5.2.3 Are Protofibrils on Pathway to Mature Fibers?

Besides the lag phase, the growth phase is also subject to debate as to whether protofibrils are on the pathway to mature fibrils. A study of huntingtin peptide (htt) fibrillization provided evidence from AFM and TEM, suggesting that these peptides formed globular assemblies first and

associated linearly to form single protofibrils (**Figure 5-6**) (Poirier et al., 2002). The transition from protofibrils to fiber can be blocked by Congo Red, leading to accumulation of protofibrils, which implied that protofibrils are on pathway to fibers. Another study (Habicht et al., 2007) also confirmed the appearance of Aβ(1-40) protofibrils prior to mature fibers by TEM (**Figure 5-7**) and show that using an antibody domain, termed B10, could prevent mature amyloid formation by stabilizing protofibrils. These studies led to a model that identifies protofibrils as intermediates to mature fibers, suggesting that protofibrils fuse to form mature fibrils, which may well co-exist with monomer addition to the fibril ends (**Figure 5-7**) (Habicht et al., 2007; Poirier et al., 2002). The fact that both Congo Red (CR) and B10 bind protofibrils and mature fibers implies both share a common surface structure.

However, preventing mature fiber formation through stabilizing protofibrils by CR or antibodies cannot rule out the possibility that protofibrils are not on the pathway, but rather a competing assembly.

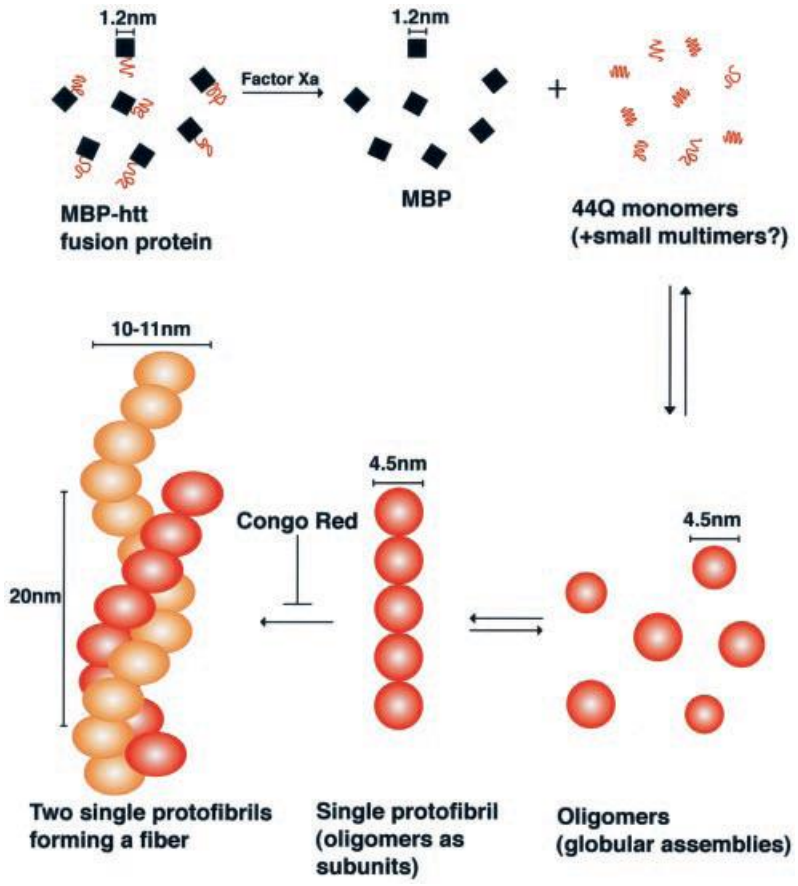


Figure 5-6 Hypothetical pathway for huntingtin peptide (htt) fibrilization . Cleaved from MBP fusion protein, htt 44Q monomers are released and undergo a conformational change. These htt monomers then associate to form globular assemblies with an average size of 4–5 nm and associate to form protofibrils with heights of 4–5 nm by AFM. The final fiber has a diameter of 10–11 nm and is suggested to consist of two protofibrils (Poirier et al., 2002).

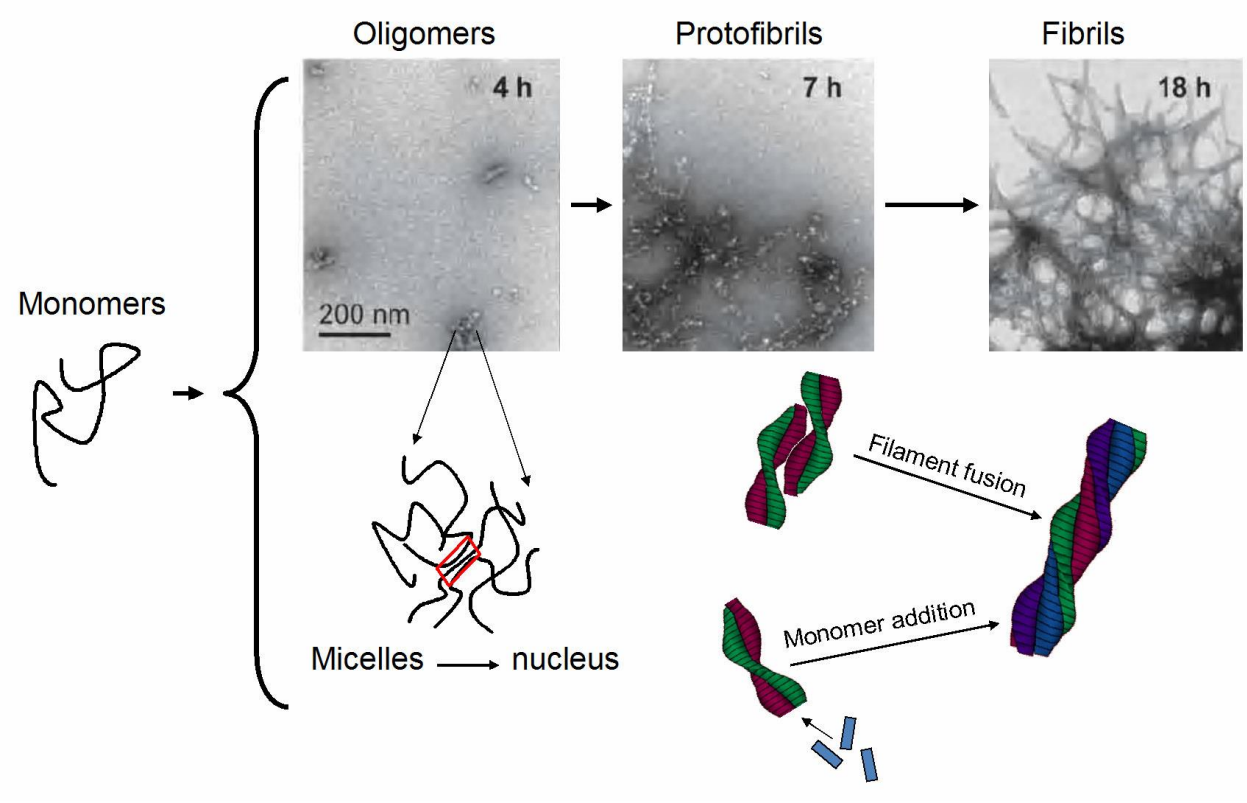


Figure 5-7 Proposed pathway of amyloid assembly. The observed particle-like species are proposed to be "micelles", which induce the nucleation. Protofibril fusion and monomer addition to the nuclei are proposed to be possible steps in the elongation phase. (The TEM images are adopted from (Habicht et al., 2007).)

Indeed, other studies suggest protofibrils are kinetically-trapped and semi-flexible species from oligomer fusion, which may be off pathway to mature fibers and do not contribute to the growth of straight long fibrils (Gosal et al., 2005; Kaylor et al., 2005; Modler, Gast, Lutsch, & Damaschun, 2003). For example, β_2-microglobulin is reveals to have two pathways with distinct kinetics by Thioflavin-T (ThT) binding (**Figure 5-8**) (Gosal et al., 2005). One pathway occurs through rapid linking of monomer/ oligomer particles to rod-like (RL) and worm-like (WL)

fibrils. By contrast, another pathway involves nucleation-dependent growth, which is characterized by lag-phase kinetics and results in long-straight (LS) fibers. However, the model for oligomer fusion to WL fibrils is not fully understood.

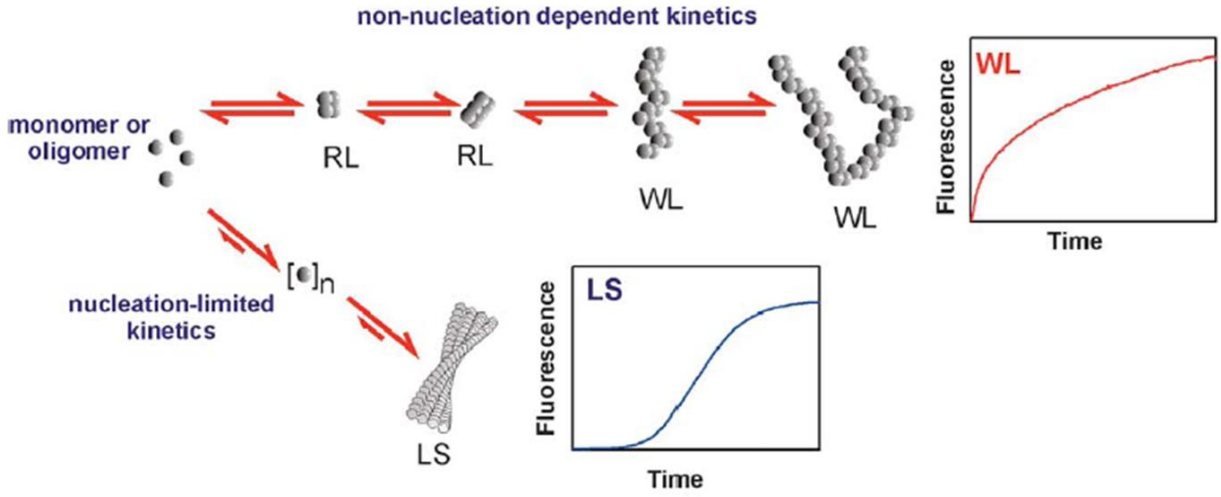


Figure 5-8 Proposed model for β2-microglobulin self-assembly, including competing pathways that lead to the formation of WL or LS fibrils. The two pathways are bifurcated and compete for assembly-competent molecules. The two pathways also display different kinetic mechanisms as indicated by Thioflavin-T (ThT) fluorescence (Gosal et al., 2005).

In summary, there is not yet a commonly accepted mechanism / pathway for amyloid assembly. Maybe peptides with different sequences would access different pathways, depending on their structures. Then questions arise as to what pathway the assemblies from the dynamic networks follow in self-assembling. In previous chapters, centrifugation data of network assemblies confirm that linear trimers, H-NFaNFaNF-CHO and H-NFFaNFFaNFF-CHO, are the building blocks for assemblies in NF-DCN and NFF-DCN, respectively. However, it is not clear that 1) if the linear trimers directly generate those particles through hydrophobic collapse or experience

some extent of non-covalent oligomerization preceding the emergence of particles. (note: the term 'non-covalent oligomerization' here refers to non-covalent associations among building block linear trimers, which is to differentiate 'oligomerization' through covalent condensation reactions referred in DCNs); 2) if the protofibrils are intermediates en route to mature fibers and 3) if the network assemblies can be affected by exogenous seeds and have cross-β compatibility. In this chapter, I will employ IMS-MS and a series of seeding experiments to address these questions.

5.3 Materials and Methods

5.3.1 Materials

All commercially available chemicals are purchased from Sigma-Aldrich, AnaSpec and Nova Biochem. Anhydrous solvents are either dried over molecular sieve (4 Å) that had been pre-treated overnight at 300 ºC or purchased from EMD or Acros organics. HPLC grade acetonitrile and water are obtained from Sigma-Aldrich and/or Fisher Scientific. TLC plates are purchased from EMD (silica gel 60 F_{254}). Fmoc-amino acids, resins and solid phase peptide synthesizer reagents are purchased from AnaSpec. Distilled deionized water for sample preparation is obtained from EMD chemicals Inc.

5.3.2 Dynamic Peptide Network Preparation

After final deprotection of the precursor of the network, Boc-NF-CHO by trifluoroacetic acid, resulting network building blocks of NH_2-NF-CHO are dissolved to a concentration of 8mM in water/acetonitrile (3/2, v/v) at ambient temperature under N_2 protection. All the solvents are

flushed with N_2 gas for at least 15 min before usage. Dissolution is assisted by ~2 minutes of continuous vortexing, followed by ~15 minutes of bath sonication until solution became clear. The pH of the solution is adjusted by titrating aliquots of 10mM NaOH and the final pH value is determined by pH meter (Fisher Scientific Accumet Basic AB15 pH meter).

5.3.3 HPLC and LC-MS Analyses

HPLC analyses are performed on Waters Delta 600 equipped with a photodiode array UV/Vis detector at room temperature using a reversed-phase HPLC column (Kromasil 100-5C18, 4.6 × 250mm). Solvent A: water. Solvent B: acetonitrile. UV absorbance is monitored at 258nm (for Phenyl ring side chain absorption) and 222nm (for amide bond absorption). Flow rate is 1.0 mL/min. Gradient is from 10% acetonitrile to 90% acetonitrile, 2% acetonitrile/ min. LC-MS analyses are performed on Waters Synapt G2 MS/Acquity UPLC system. Positive-ion mass spectra are obtained using electrospray ionization.

Time (mins)	Solvent A	Solvent B	Note
0	90%	10%	Gradient starts
40	10%	90%	Gradient ends
41	0%	100%	Cleaning starts
48	0%	100%	Cleaning ends
49	10%	90%	Re-equilibrium starts
56	10%	90%	Re-equilibrium ends

5.3.4 IMS-MS

Direct total ion current analysis is performed on Synapt G2 High Definition Mass Spectrometry system (Waters Corporation, Manchester, UK), which is a hybrid quadrupole-ion mobility-orthogonal acceleration time-of-flight instrument, with typical resolving power of 20,000 m/Δm (FWHM) and mass accuracy of 9 ppm at m/z 556.2771. The instrument is operated in positive ion mode with a probe capillary voltage of 3 kV, and a sampling cone voltage of 45 V. The source and desolvation temperatures are set to 120 °C and 250 °C, respectively; and the nitrogen desolvation flow rate is set to 650 L h^{-1}. The mass spectrometer is calibrated across the 50-1200 m/z range using a 0.5 mM sodium formate solution prepared in 90:10 2-propanol:water v/v. Data are mass corrected during acquisition using a leucine encephalin (m/z 556.2728) reference spray (LockSpray) infused at 2 µL min^{-1}. The scan time is set to 1 s. Data acquisition and processing is carried out using MassLynx v4.1 and Drift Scope v2.1 (Waters Corp.) Column: Waters Acquity UPLC BEH C18 column 1.7 µm, 2.1 x 50 mm)

5.3.5 Transmission Electron Microscopy and Electron Diffraction

Aliquots (10 µl) of sample solutions are dropped on TEM grids (200 mesh copper grid covered with a thin carbon film, purchased from Electron Microscopy Sciences) for 3 minutes before excessive solution is blotted with filter paper. Uranyl acetate (10 µl of 5% solution) is added for 3 minutes for negative staining. Extra fluid is blotted with filter paper. The grids are analyzed on a Hitachi H-7500 transmission electron microscope with a LaB6 emission filament at an accelerating voltage of 75 kV.

5.3.6 Microwave Assisted Solid-Phase Peptide Synthesis

Peptides H-NFNFNF-NH_2 and Ac-KLVFFAL-NH_2 (E22L) are synthesized on a Liberty CEM Microwave Automated Peptide Synthesizer utilizing FMOC-Rink Amide MBHA Resins purchased from AnaSpec. All FMOC protected amino acids are from Anaspec, and other chemicals from Sigma-Aldrich or Fisher Scientific. Each peptide synthesis is performed at 0.1 mmol using a 45 mL reaction vessel at a scale of 0.1mmol. FMOC-Rink Amide MBHA Resin is initially swollen using ~7 ml dimethylformamide (DMF) for 15 minutes. FMOC deprotection is achieved by addition of 20% piperdine 0.1M N-Hydroxybenzotriazole (HOBt) in DMF with microwave power set to maintain temperature between 45-55°C for 180 sec, followed by 3X flushing with DMF. Each coupling step is performed using 0.1M FMOC protected amino acid, and activated with 0.1 M 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate (HBTU), and 0.2M N,N- Diisopropylethylamine (DIEA) in DMF. Coupling temperatures are maintained between 75-82°C 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°C) diethyl ether. Precipitated product is centrifuged at 3000 rpm for 15 min, and the pellet is subjected to 3 additional rounds of ishing 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.

Purified peptides are dissolved in water/acetonitrile (3/2, v/v) with 0.1% TFA for nanofiber assemblies. Dissolution is assisted by ~2 minutes of continuous vortexing, followed by ~15 minutes of sonication until solution became clear.

5.3.7 Alexa Binding of Seeded Assemblies in NFF-DCN

5 mg of Alexa Fluor® 633 (excitation maxima: 632nm; emission maxima: 647 nm) is dissolved in acetonitrile/ water (3/2, v/v) as the stock solution stored at -20°C. An aliquot of 20 µl solution is taken out and diluted with 980 µl of acetonitrile/ water (3/2, v/v) to 0.1g/ml (8.33×10^{-2} mM) and titrated to the NFF-DCN seeded with E22L. The molar ratio of NFF-DCN : E22L: Alexa= 240: 30: 1, and the concentrations of NFF-DCN, E22L and Alexa are 3.33×10^{-1} mM, 4.20×10^{-2} mM, 1.39×10^{-3} mM, respectively.

5.3.8 Fluorescent Peptide Seeds Preparation

Fluorescence peptides (Rh17-22) are synthesized by replacing the lysine residue of Aβ(16-22) with Rhodamine 110 (Rh110) through standard FMOC peptide synthesis protocol described above. Fluorescent nanotubes are prepared by mixing Rh17-22 and E22L with a molar ratio of 1: 250 (8 µM/ 2 mM) in water/acetonitrile (3/2, v/v) with 0.1% TFA.

5.3.9 Dual Color Fluorescence Imaging of Alexa 633 and Rhodamine Labeled Peptide

The rhodamine fluorescent peptides are added into the seeded NFF-DCN which is pre-stained with Alexa 633. The molar ratio of rhodamine fluorescent peptides, NFF-DCN (seeded with E22L tubes) is 1:8 and concentrations are 46 µM, 0.37 mM, respectively. After mixing for one day, the mixture is diluted by 20 folds for preparing samples for fluorescence imaging. The two-photon excitation experimental is under 780nm excitation. By applying synchronized two channel imaging with 570nm dichroic, and filters 530/50 for rhodamine labeled peptide and

645/75 for Alex633, Alexa 633 allows for visualizing the assemblies in the seeded NFF network, and rhodamine labeled seeds for detecting the growth in the seeding event.

5.4 Results

5.4.1 Using IMS-MS to Probe the Oligomeric Intermediates on Pathway of Network Assemblies

The network stock solution prepared (8mM, pH4) in water/acetonitrile (3/2, v/v) at ambient temperature within an hour is diluted to 20 μM before injection into the spectrometer. IMS-MS driftscope plot shows drift time versus m/z versus intensity. The singly and doubly charged network species ion families are clearly separated and marked as $[M+H]^+$, $[M+2H]^{2+}$, respectively (**Figure 5-9**). Within these bands, drift time generally increase with mass. To note here, as the X-axis is not long enough to show the entire bands of singly and doubly charged species, the two bands with larger m/z values are cut off and shown at the upper- left corner in the **Figure 5-9**.

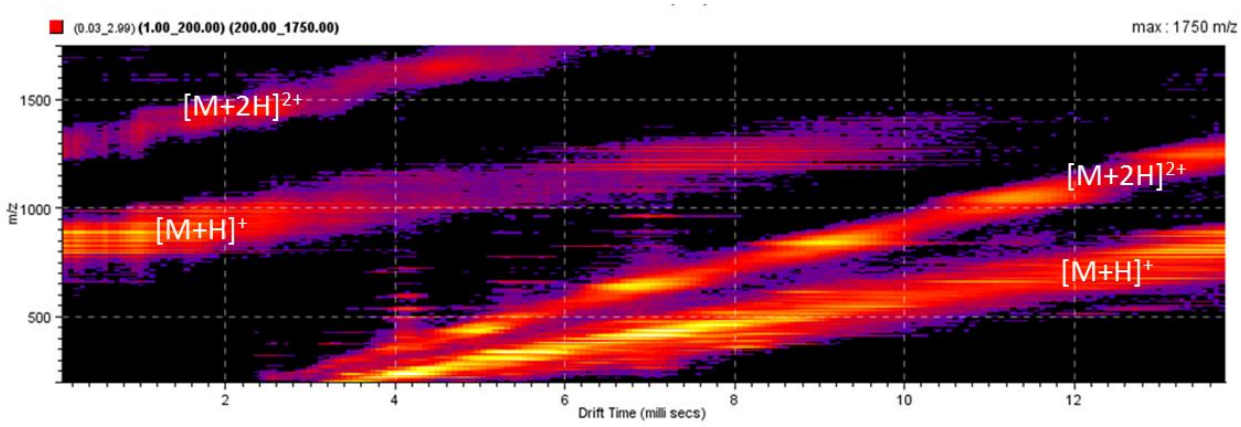


Figure 5-9 IMS-MS driftscope plot of covalent and non-covalent oligomers presented in NFF network (8mM stock solution diluted to 20 μM) shown as drift time (milliseconds, ms) to *m/z*, where the singly, doubly charged network species ion families are clearly separated. The signal amplitude is color-coded, increasing from purple (low intensity) to bright yellow (high intensity).

To get more insight into the non-covalent association of network members, each individual network species' m/z value is selected for further analysis. Shown in **Figure 5-10**, singly charged monomer (NFF-CHO) (MON^{+}), doubly charged noncovalent dimer of monomers ($2MON^{2+}$), and triply charged noncovalent trimer of monomers ($3MON^{3+}$) are observed. These ions all share the same *m/z* value, but have different drift times, ~ 5 ms (milliseconds), 7 ms and 8 ms for triply, doubly and singly charged species, respectively. For network linear dimer (NFFaNFF-CHO), singly charged linear dimer (DIM^{+}), doubly charged noncovalent dimer of linear dimers ($2DIM^{2+}$) are detected with the same *m/z* value but different drift time, ~ 13 ms and 9 ms, respectively (**Figure 5-11**). No oligomeric state of cyclic dimers is observed. For linear trimer (NFFaNFFaNFF-CHO), singly charged linear trimer (TRI^{+}) and doubly charged ($2TRI^{2+}$) ions with the same *m/z* value but different drift times are detected. The doubly charged ion has a

shorter drift time of ~13 ms compared to singly charged linear trimer ~ 21 ms (14 ms + 7 ms) (**Figure 5-12**). Because the charged ion bands are cut off as shown in **Figure 5-9**, the actual drift time for singly charged linear trimer is 14 ms + 7 ms. The IMS-MS analysis of oligomeric state of the network species in the dynamic network at the early stage (within an hour) suggests that the noncovalent oligomerization indeed occurred fast and is on pathway preceding the emergence of particles and fibers.

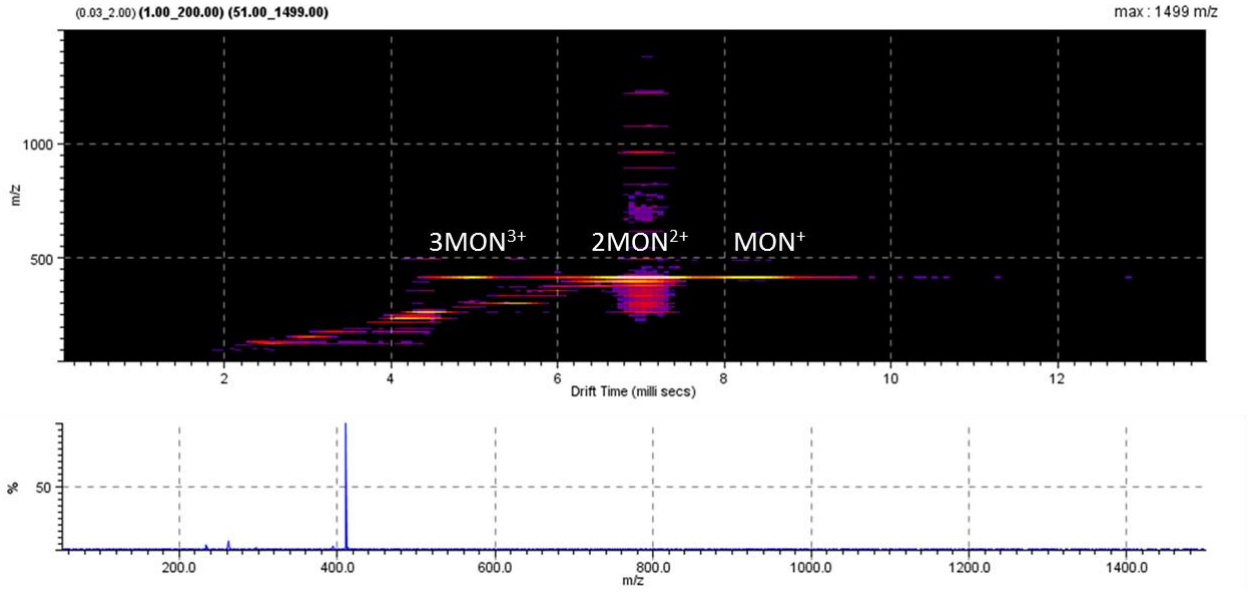


Figure 5-10 IMS-MS analysis of oligomeric state of monomers (MON) in NFF-DCN. (**Top**) Ion mobility drift time (ms) are shown on the horizontal axis and *m/z* values on the vertical axis. Singly charged monomer (NFF-CHO), doubly charged noncovalent dimer of monomers and triply charged noncovalent trimer of monomers with the same *m/z* value but different drift time are shown. Numbers in front MON denote oligomer order, with the positive charge state of each oligomer ions in superscript. (**Bottom**) *m/z* value of 411.21 for the species mentioned above.

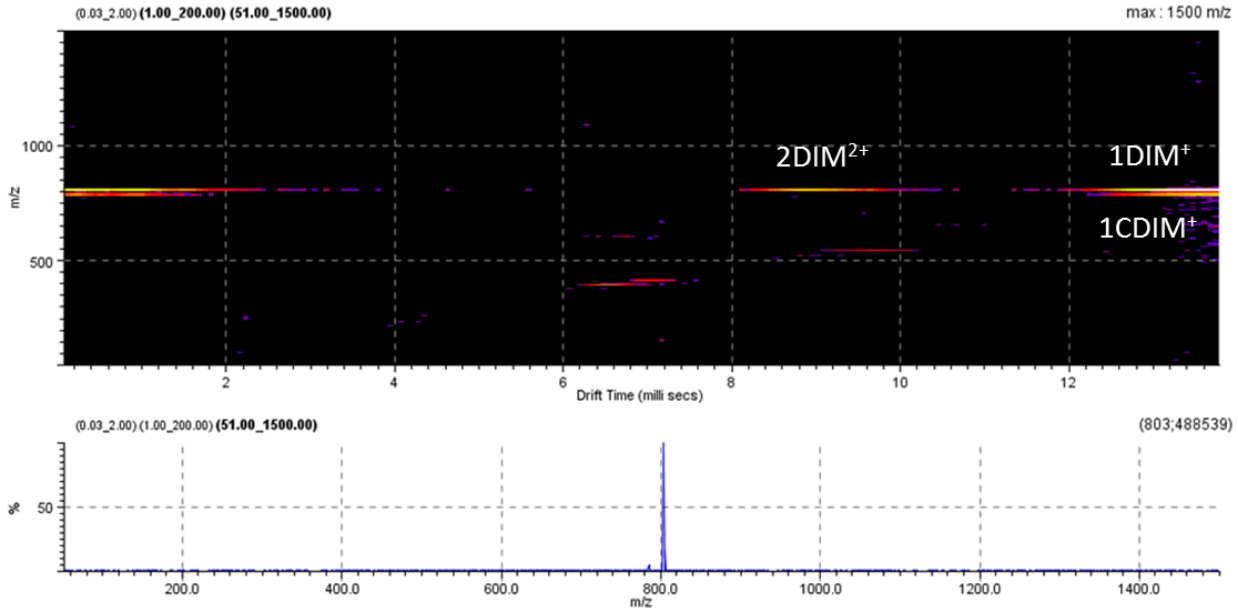


Figure 5-11 IMS-MS analysis of oligomeric state of linear dimers (DIM) and cyclic dimers (CDIM) in NFF-DCN. (**Top**) Ion mobility drift time (ms) are shown on the horizontal axis and *m/z* values on the vertical axis. Singly charged linear dimer (NFFaNFF-CHO), doubly charged noncovalent dimer of linear dimers with the same *m/z* value but different drift time are shown. No oligomeric state of cyclic dimers is observed. Numbers in front denote oligomer order, with the positive charge state of each oligomer ions in superscript. (**Bottom**) *m/z* value of 802.38 for linear dimers and *m/z* value of 784.37 for cyclic dimers.

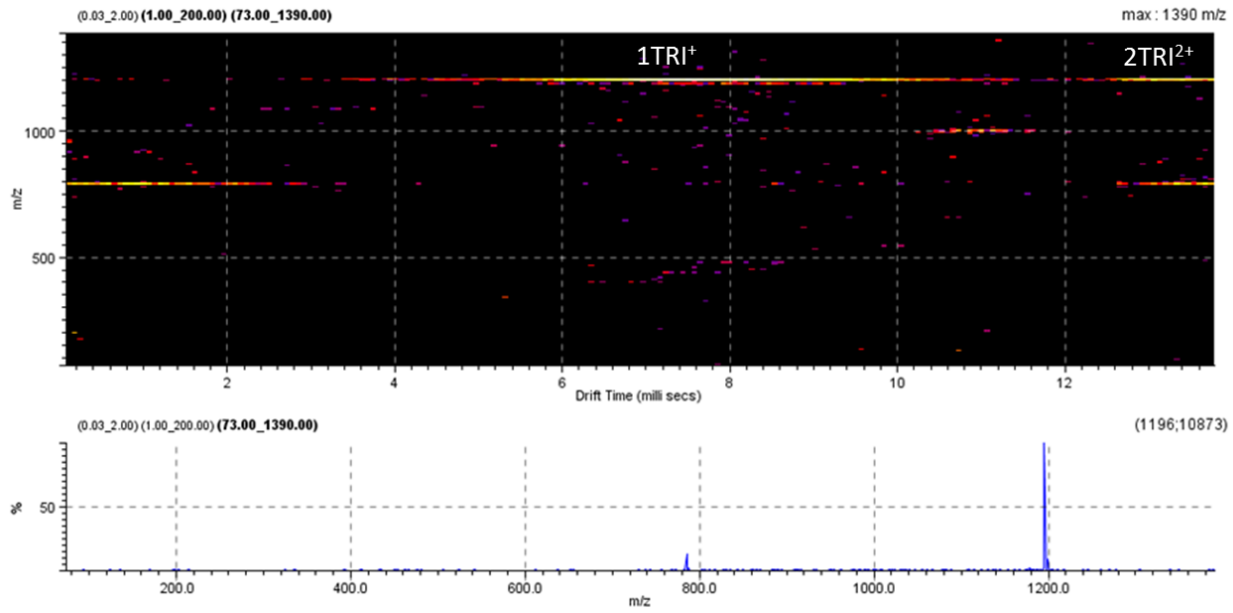


Figure 5-12 IMS-MS analysis of oligomeric state of linear trimers (TRI) in NFF-DCN. (**Top**) Ion mobility drift time (ms) are shown on the horizontal axis and *m/z* values on the vertical axis. Singly charged linear trimer (NFFaNFFaNFF-CHO), doubly charged noncovalent dimer of linear trimer with the same *m/tz* value but different drift time are shown. Numbers in front denote oligomer order, with the positive charge state of each oligomer ions in superscript. (**Bottom**) *m/z* value of 1194.57 for linear dimers.

5.4.2 Are Particles and Molten Globule on Pathway for Network Assemblies to Mature Fibers?

It has been shown in previous chapters that both NF-DCN and NFF-DCN demonstrate phase transitions marked by emergence of particles, molten globules and fibers, consistent with a similar metastable phase seen with amyloid peptide assemblies (**Figure 5-13**) (Auer et al., 2012; Childers et al., 2012). To better understand the transitions, time-dependent changes of particle populations are assessed quantitatively. Before 15 hours, the particles size distribution appears

homogenous by TEM and can be fit into single Gaussian population with a width of 13.9 ± 0.1 nm (fwhm = 4.3 ± 0.2 nm) (**Figure 5-14**). However, starting from 19 hours, two populations of particle sizes are observed (**Figure 5-15**) and the ratio of small particles to larger ones decreased over time, consistent with the growth being Ostwald ripening (Baldan, 2002). As the larger particles have less surface tension, there is a net diffusion of molecules on the surface of small particles to the larger particles. The number of smaller particles continues to shrink, while larger particles continue to grow. By 36 and 48 hours, almost all of small particles with width from 15 to 40 nm have disappeared. When larger particles (molten globules) transition to paracrystalline phase, the growth of fibers is accompanied by the disappearance of particles.

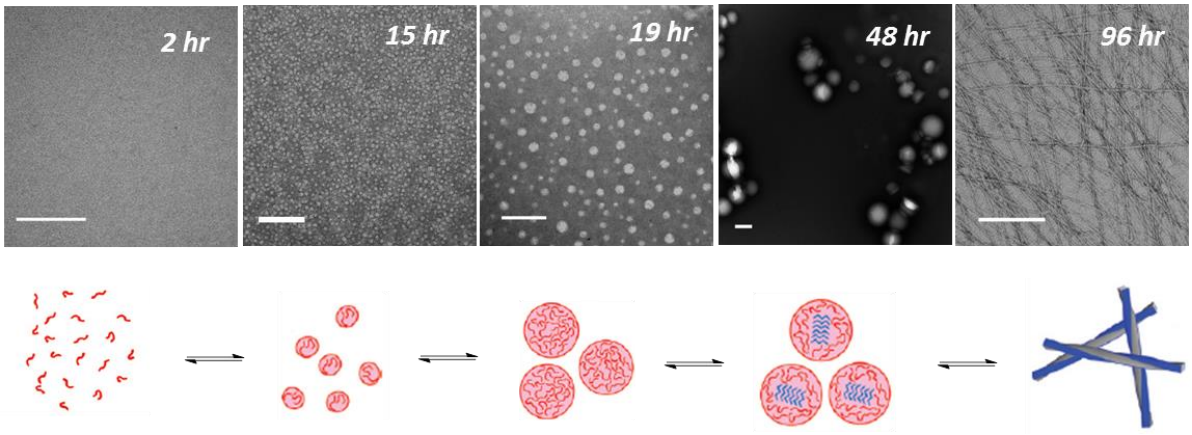


Figure 5-13 Transmission electron micrographs of aliquots taken directly from the NFF-DCN at 2 hrs, 15 hrs, 19 hrs, 48 hrs and 96 hrs with each scale bar at 200 nm (**top panel**). Models of phases based on previous peptide analyses (Childers et al., 2012) (**bottom panel**).

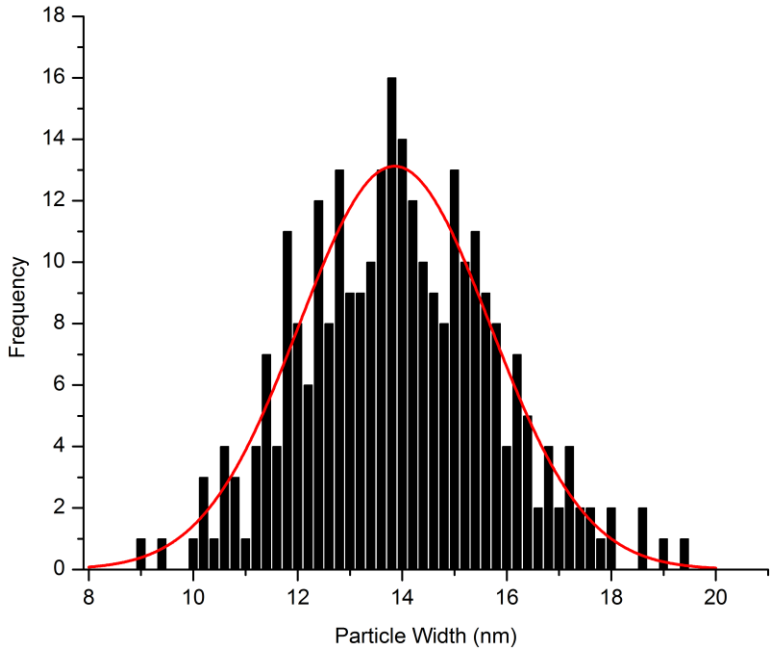


Figure 5-14 Gaussian distribution of measured particle widths fit at 15 h with center at 13.9 ± 0.1 nm (fwhm = 4.3 ± 0.2 nm)

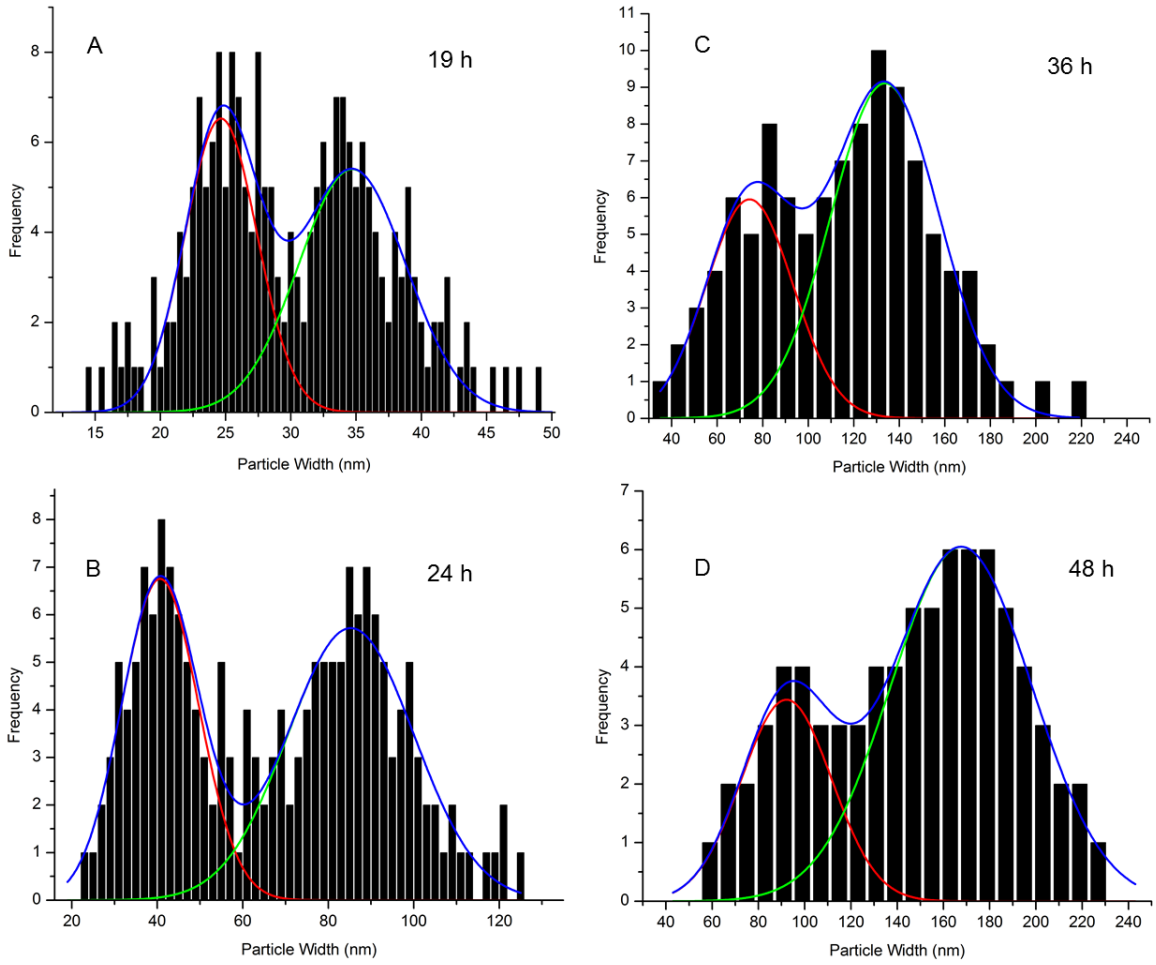


Figure 5-15 Gaussian distribution of measured particle widths fit to two individual populations (red/ green line) and the sum (blue line) at (**A**) 19 h, centers at 24.6 ± 0.3 nm (fwhm = 6.5 ± 0.7 nm) and 34.7 ± 0.5 nm (fwhm = 9.7 ± 1.2 nm), (**B**) 24 h, centers at 40.7 ± 0.7 nm (fwhm = 21.3 ± 1.6 nm) and 85.3 ± 1.0 nm (fwhm = 34.8 ± 2.6 nm), (**C**) 36 h, centers at 72.4 ± 4.0 nm (fwhm = 45.2 ± 8.2 nm) and 133.8 ± 3.1 nm (fwhm = 56.1 ± 6.6 nm), (**D**) 48 h, centers at 92.1 ± 1.8 nm (fwhm = 46.0 ± 4.3 nm) and 168.0 ± 1.4 nm (fwhm = 72.4 ± 3.6 nm)

5.4.3 Are Protofibrils (Twisted Fibers) on Pathway for Network Assemblies to Mature Fibers?

Besides particles, protofibrils are also observed in both of NF-CHO and NFF-CHO dynamic networks. In NF networks, twisted fiber structure became apparent by TEM analysis at ~72 hours (**Figure 5-16**), co-existing with some long straight fibers. However, as the network mature past 4 days, the long straight fibers dominate and twisted fibers disappear, suggesting the latter one as an intermediate. But it cannot rule out the possibility that the twisted fibers and straight fibers adopted two different pathways and straight fibers have lower free energy which causes twisted fibers to disassemble and re-form the long straight ones. This evidence implies that twisted fibers are kinetic intermediates, whereas straight fibers are thermodynamic products.

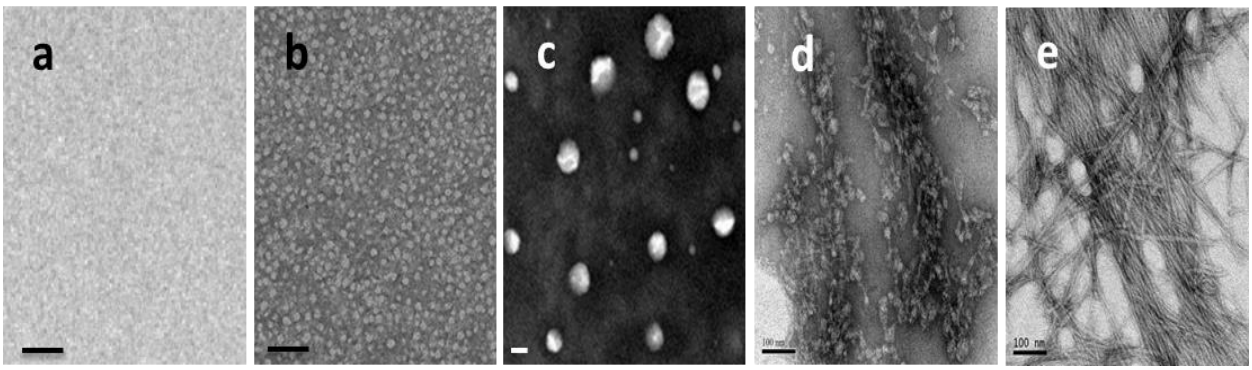


Figure 5-16 Transmission electron micrographs of aliquots taken directly from the NF-DCN in water/acetonitrile (3/2, v/v), pH 4 at (**a**) 4 hrs, (**b**) 17 hrs , (**c**) 48 hrs, (**d**) 72 hrs and (**e**) 96 hrs with each scale bar at 100 nm.

In the NFF network, the twisted fibers are the only species observed after particles transitioned to the paracrystalline phase at pH 5.1 (**Figure 5-17 B**) and did not change to straight fibers even after 15 days. However, at pH 4.6 and pH 7.6, only straight fibers form (**Figure 5-17 A, C**). The reason for this pH dependence may be obtained from **Figure 4-4**, where the population of linear trimers is highest near pH 5 and kinetics of trimer growth is faster than those at other pH values (data not shown). The hypothesis is that at pH 5.1, linear trimers accumulate to achieve a critical concentration for assembling faster and twisted fibers may result from a non-nucleation dependent pathway, where small particles fuse together in a coalescence process (Kumar & Walter, 2011; Poirier et al., 2002) (Gosal et al., 2005); or more likely it underwent some extent of nucleation in particles, but due to fast hydrophobic collapse, the packing registry of the peptide-like strands is less rigid than the straight fibers, providing a the kinetic trap. On the other hand, the formation of straight fibers appears to require a relatively long nucleation time, where small particles gradually form big molten globules and paracrystalline nuclei slowly accumulate (Cheon et al., 2007), packing in a way that is most thermodynamically stable. The emergence of paracrystalline nuclei only in molten globules suggests that larger population is required for nucleation.

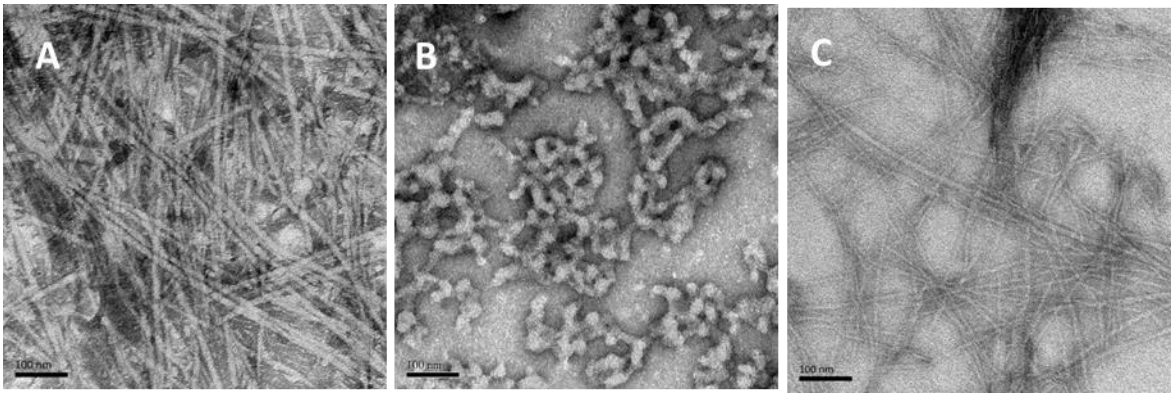


Figure 5-17 Transmission electron micrographs of NFF-DCN in water/acetonitrile (3/2, v/v) incubated after 5 days at (**A**) pH 4.6, (**B**) pH 5.1,(**C**) pH 7.6 with each scale bar at 100 nm.

5.4.4 Using Seeding experiment to Probe Protofibrils on Pathway of Network Assemblies

To test the above hypothesis, preformed straight fibers seeds (**Figure 5-18 A**) are added to seed the NFF network at pH 5.1 within 2 hours of construction of the network. After 2 days, homogenous straight fibers are observed (**Figure 5-18 B**), where only twisted fibers formed in the absence of seeds (**Figure 5-18 C**). This result supports the hypothesis that the twisted fibers are kinetically trapped species, straight fiber seeds help overcome the energy barrier and provide templates for continued growth.

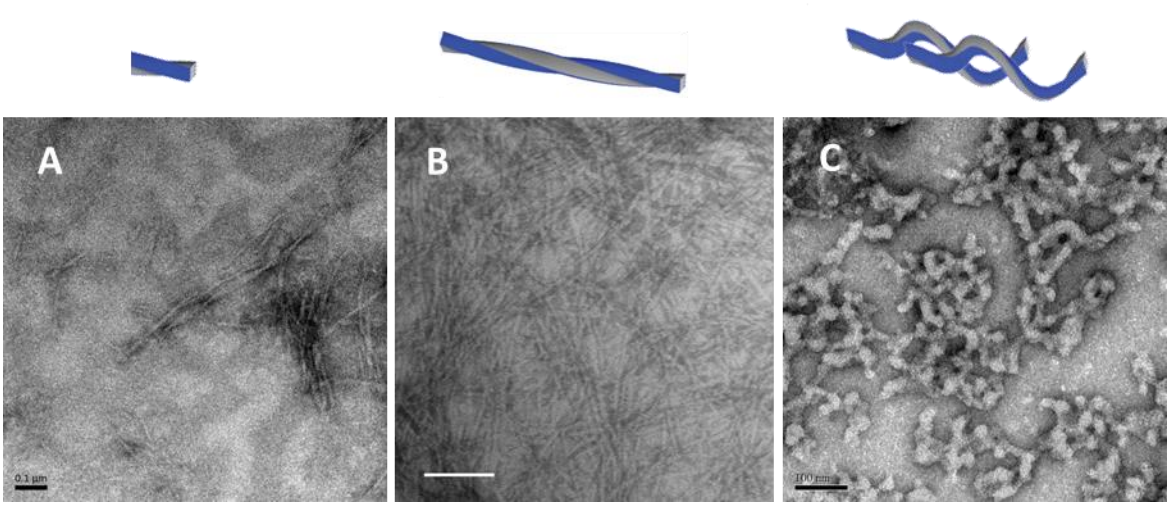


Figure 5-18 Seeding experiment of straight fibers formed at pH 4.6 (**A**) to seed NFF-DCN at pH 5.1 within 2 hours of construction of the network at molar ratio of 1:8. Homogenous straight fibers are observed at this pH at 2 days (**B**), in contrast to NFF-DCN at pH 5.1 in the absence of straight fiber seeds. Scale bar of (**A**) and (**C**): 100 nm; (**B**): 200 nm

Twisted fibers are consumed by straight fibers in NF-DCN (**Figure 5-16**) and twisted fiber assembly can be bypassed by addition of straight fiber seeds (**Figure 5-17**). Can straight fibers assembly be passed by twisted fiber seeds? When sonicated twisted fiber seeds are added into NFF-DCN at pH 4.6 within 2 hours of construction of the network at molar ratio of 1:8, homogenous twisted fibers form within 2 days, while only straight fibers form in the absence of seeds. Therefore, the twisted fiber seeds bypass the nucleation phase required for more rigid straight fibers to form and result in the kinetic trap of the system. The twisted fibers remained stable after 15 days and are not converted to straight ones.

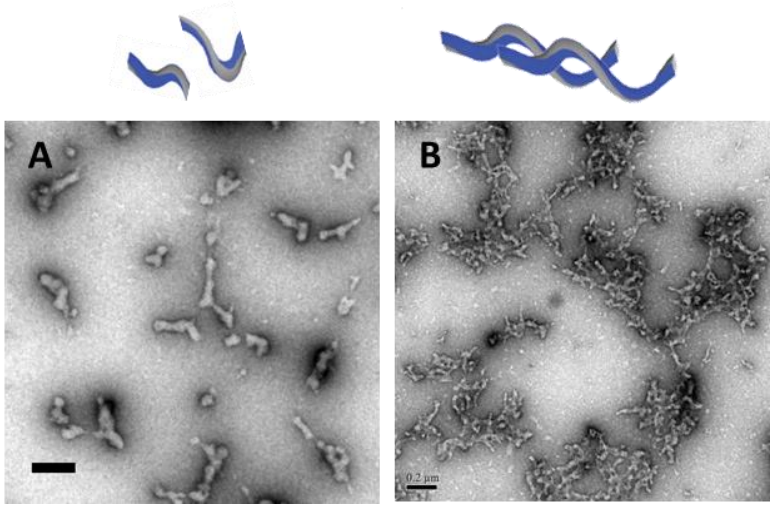


Figure 5-19 Preformed twisted fibers seeds (**A**) are added into NFF-DCN at pH 4.6 within 2 hours of construction of the network at molar ratio of 1:8, resulting in homogenous twisted fibers within 2 days (**B**). Scale bar: 200 nm

5.4.5 Exogenous Seeding of NFF-DCN by H-NFNFNF-NH_2

As shown above, assemblies in these dynamic networks with acetal linked oligomers can adopt the template morphology of endogenous seeds (with same chemical structures) much like amyloid and prion peptides. Next, it is more intriguing to see if these dynamic networks would be susceptible to infectious exogenous assemblies, like a biological network being infected by a virus. The all-amide-containing peptide H-NFNFNF-NH_2, incubated under identical solvent conditions as dynamic chemical networks in water/acetonitrile (3/2, v/v), assembles as straight fibers with 7.9 ± 3.1 nm widths (**Figure 5-20**). Since previous data has reveals that growth of fiber/nanotube assemblies occur though peptide monomers adding at the ends of templates (Childers et al., 2012; Yan Liang et al., 2010), to make more effective seeds with increased fiber

ends, the straight fibers of H-NFNFNF-NH_2 are also exposed to mild bath sonication for 30 minutes to generate multiple short fiber seeds (**Figure 5-21 A**).

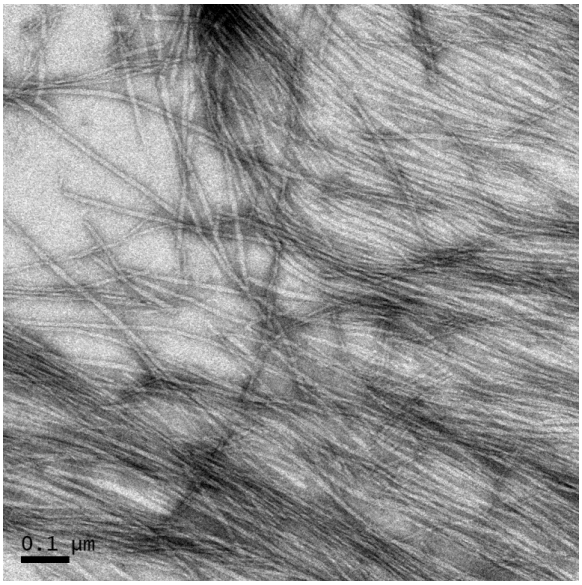


Figure 5-20 Transmission electron micrographs (TEM) of H-NFNFNF-NH2 peptides

at 2.2mM, which self-assemble as straight fibers with several micron length and 7.9 ± 3.1 nm width.

The NFF network is incubated in water/acetonitrile (3/2, v/v) at pH 4 for 36 hours, where the assemblies are particles (chapter 4), and then the seeds of H-NFNFNF-NH_2 are added to the NFF-CHO network at 1:8 molar ratio (0.25mM: 2.0mM). Concentration of seeds is below the critical concentration of itself. After mixing, TEM of aliquots taken from the mixture show a combination of particles and fibers at 2 hour (**Figure 5-21 B**). The particles observed are probably from NFF-CHO network and fibers started elongating. At 2 days, most of the particles are gone and a dense array of long tangled fibril with 8.1 ± 3.2 nm widths dominate the

population in the seeded network (**Figure 5-21 C**). The observed fast transition is consistent with efficient template seeding of the network by this peptide. The seeded fibers show dimensions similar to the seeding fiber H-NFNFNF-NH_2, but changed morphology from straight fibers to tangled fibers, implying the fundamental structural change of the packing registry of seeded fibers.

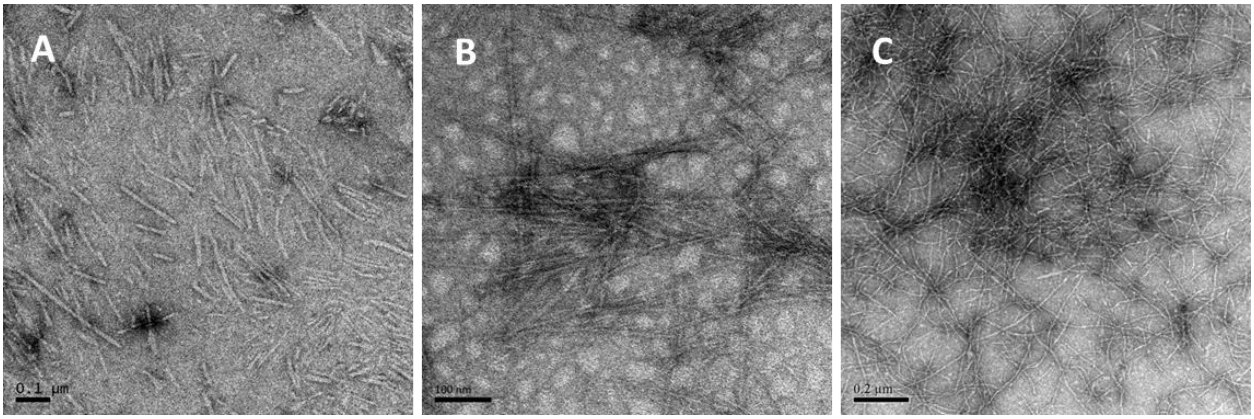


Figure 5-21 Templating of NFF-DCN networks with seeds of H-NFNFNF-NH2 peptide assemblies . All samples are prepared in water/acetonitrile (3/2, v/v). TEM micrographs of (**A**) Templating H-NFNFNF-NH_2 fiber seeds are made through mild bath sonication (30 min) to break the long fibers and increase the number of fiber ends. (**B**) *NFF*-DCN seeded with sonicated H-NFNFNF-NH_2 fibers (final concentration is 2.0mM, 0.25mM, respectively) at 2 hour. (**D**) *NFF*-DCN seeded with sonicated NFNFNF-NH_2 fibers at 2 days. Scale bar: (**A**), (**B**):100nm; (**C**):200nm

Further investigation of the seeded fiber by Circular dichroism (CD) is shown in **Figure 5-22**. Before seeding experiment, both CD spectra of the assemblies in NFF-DCN and H-NFNFNF-NH_2 display a minimum band near 190 nm and two positive bands near 200 nm and 224 nm, 197

nm and 220 nm, respectively, most consistent with previously reported β-turn structures (Gao et al., 2002; Surewicz & Mantsch, 1988). In the CD spectrum of seeded NFF network, two positive band blue shifted and became more similar to seed peptide, suggesting the template effect.

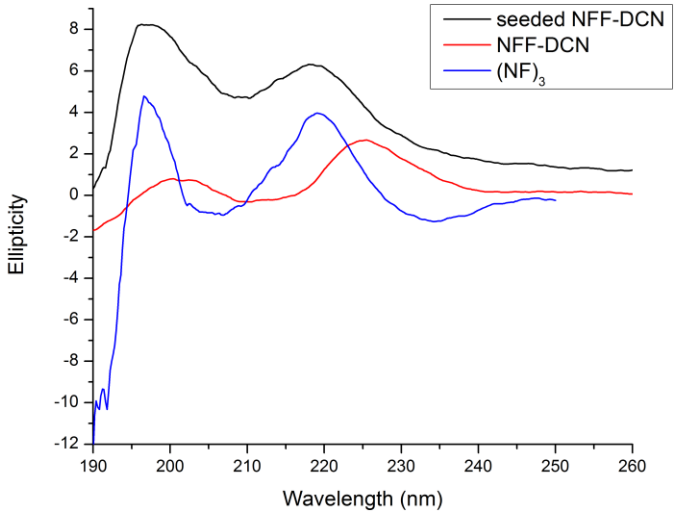


Figure 5-22 Circular dichroism (CD) spectra of NFF-DCN, 8mM, 4 days (red), H-NFNFNF-NH_2 fiber seeds at 20 days, 2mM (blue) and seeded NFF-DCN at 2 days, concentrations for seeds and network are 0.25mM, 2.0mM, respectively (black).

5.4.6 Exogenous Seeding of NFF-DCN by Ac-KLVFFAL-NH_2 Peptide Nanotubes

The seeding experiments shown above either utilize endogenous fiber seeds from similar dynamic networks or exogenous H-NFNFNF-NH_2 fiber seeds with similar secondary structures as assemblies in the network. It is of great interest to test whether these acetal linkages are able to access more of the secondary structures of amide peptides. A more complex template with cross-β secondary structure and nanotube morphology, Ac-KLVFFAL-NH_2 (E22L) self-assembles in water/acetonitrile (3/2, v/v) and in a wide range of pHs as hollow nanotubes with diameter range of 38.0 ± 5.0 nm. These nanotubes are characterized by the parallel white lines of

the tube walls that exclude uranyl acetate stain (Mehta et al., 2008) (**Figure 5-23 A**). Its cross-β structure is ideal for testing if the assemblies in dynamic networks are prone to be infected by prion-like species. To increase the seeding efficiency, again E22L peptide nanotubes are exposed to mild bath sonication for 30 minutes to generate multiple short nanotubes seeds (**Figure 5-23 B**).

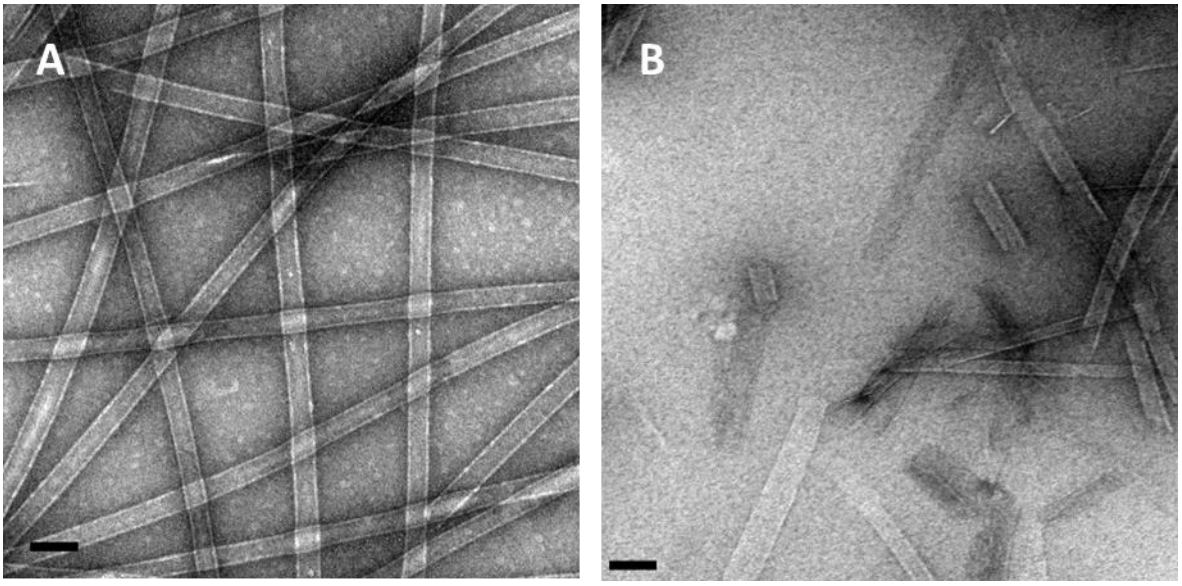


Figure 5-23 TEMs of Ac-KLVFFAL-NH2 peptide nanotubes at 2.5 mM (**A**) and short Ac-KLVFFAL-NH_2 peptide nanotube seeds generated by mild bath sonication for 30 min (**B**). Scale bar: 100nm

The NFF network is incubated in water/acetonitrile (3/2, v/v) at pH 4 for 36 hours, where the assemblies are particles (chapter 4), and then the seeds of E22L are added to the NFF-CHO network at a 1:8 molar ratio with concentration of 0.25mM and 2.0mM, respectively. The seeded network at 20 hours shows affinity of network particles attracted to the inside of the hollow nanotubes (shown as white dots inside the nanotubes in **Figure 5-24 A**), though there are many

small particles outside the nanotubes. These particles gradually transition to fibers both inside and outside the nanotubes shown by TEM at day 3 with a concomitant decrease in the number of small particles (**Figure 5-24 B**). At day 5, some fibril-like structures grow at the ends of tubes and sheet structures form with the disappearance of almost all the big particles (**Figure 5-24 C, D**). Newly generated tubes are observed at 7 days (**Figure 5-24 E**) and become more homogenous after 2 weeks with tube diameter size of 29.0 ± 4.6 nm, 9 nm smaller than the E22L nanotube seeds (**Figure 5-24 F**).

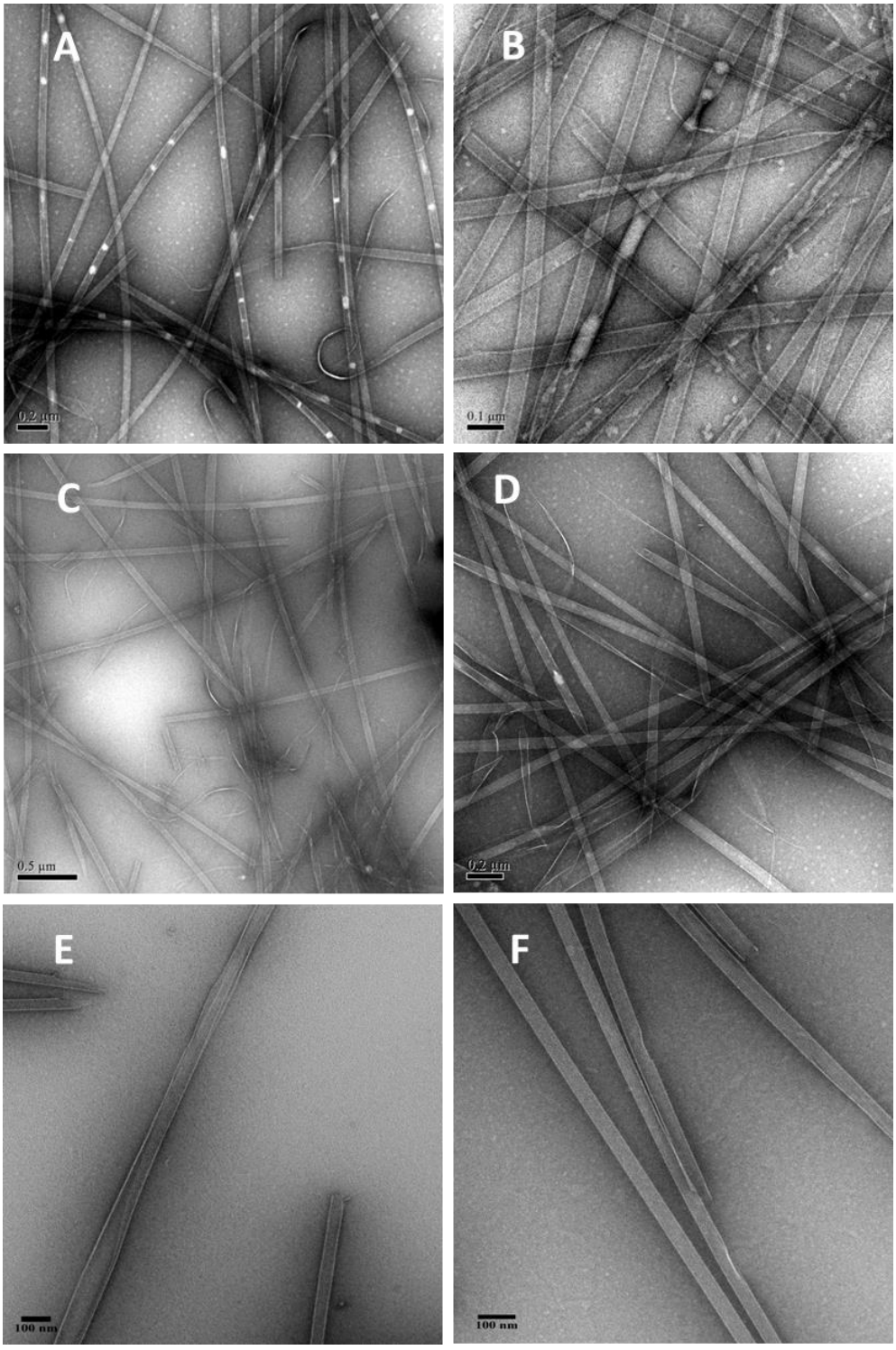


Figure 5-24 Templating of NFF-DCN networks with seeds of Ac-KLVFFAL-NH2 (E22L) peptide assemblies. Molar ratio of seeds to NFF-DCN is 1:8 with concentration of 0.25 mM and

2.0 mM, respectively. All samples are prepared in water/acetonitrile (3/2, v/v). TEM micrographs of (**A**) NFF-DCN seeded with sonicated E22L nanotubes at 20 hours (scale bar 200 nm). (**B**) NFF-DCN seeded with E22L at 3 days (scale bar 100 nm). (**C**) NFF-DCN seeded with E22L at 5 days (scale bar 500 nm). (**D**) NFF-DCN seeded with E22L at 5 days (scale bar 100 nm). (**E**) NFF-DCN seeded with E22L at 7 days (scale bar 100 nm). (**F**) NFF-DCN seeded with E22L at 16 days (scale bar 100 nm).

While this pathway is complex, the original fiber morphology of network assemblies is cleanly converted by the cross-β E22L nanotubes. Further investigation of the seeded nanotubes by CD is shown in **Figure 5-25**. The secondary structure of newly generated nanotubes is converted from β-turn to cross-β structure, adopting the seeds' packing registry. To get more insight of what network members being selected to form the new nanotubes, the seeded assemblies are spun down (30 min at 16,000 × g) and analyzed by HPLC, which show the nanotubes to contain a stoichiometry of trimers, dimers, and the infecting peptide at a ratio of 7: 0.6: 1, respectively, suggesting that the 7-residue seeds template associates most of 9-residue trimers and some 6-residue dimers, which probably accounts for the size difference seen in resulted nanotubes from the seeds. However, more characterization will be required for detailed structure assignments.

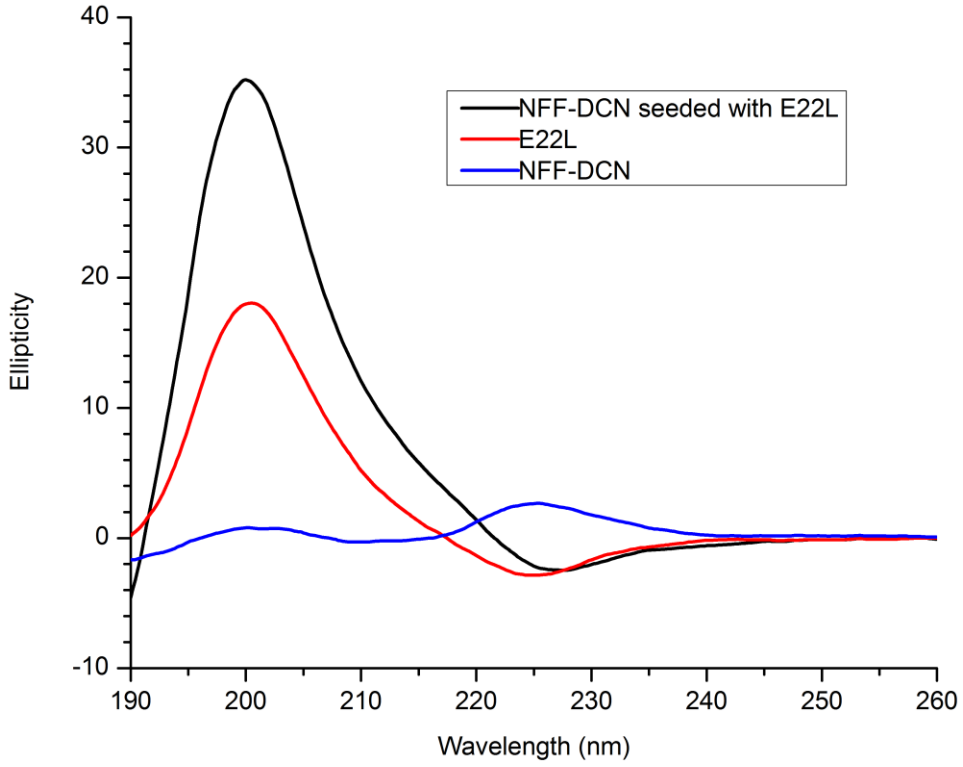


Figure 5-25 Circular dichroism (CD) spectra of NFF-DCN (blue), E22L nanotube seeds (red) and seeded NFF-DCN (black).

5.4.7 Alexa 633 Binds Co-assembly of NFF-DCN with E22L

The TEMs shown above follows the process of E22L cross-seeding the network assemblies and suggests the templating effect of E22L in directing the morphology change. The questions then arise to as: How similar is the surface of newly generated nanotubes compared to the E22L tubes seeds? Are the network particles incorporated in the new nanotube structures?

Alexa 633 has been demonstrate for detecting Aβ (16-22) nanotube assemblies under fluorescence microscopy (Yan Liang et al., 2010) as it has high affinity to the nanotubes probably due to the charge complementarity between the negatively charged fluorophore and

positively charged surface on Aβ (16-22) nanotubes (Mehta et al., 2008) and the hydrophobic interaction between them. The assemblies generated from mixing NFF-DCN with E22L demonstrate similar nanotube morphology and should also be positively charged, which might bind the dye as well. To test this, Alexa 633 is added into the seeded NFF network. The Alexa dye to peptide molar ratio is 1: 270 with concentrations of 1.39×10^{-3} mM and 3.75×10^{-1} mM, respectively. The fluorescent microscopy image of newly formed nanotubes with Alexa dye clearly demonstrates its binding on the tube surface (**Figure 5-26**), implying that the network assemblies have similar charge attraction and hydrophobic interaction as the E22L tubes. However, the bundling of nanotubes suggests that the newly formed tubes may have denser positive charges resulting from the positively charged acetals from network members.

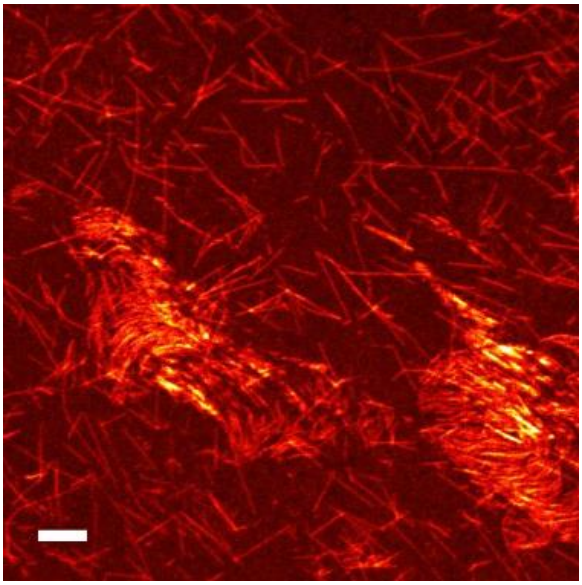


Figure 5-26 Fluorescent microscopy image of Alexa Fluor 633 binding on the co-assembly of NFF-DCN with E22L (molar ratio of peptide to Alexa 633 = 270:1, the concentrations of NFF-DCN, E22L and Alexa are 3.33×10^{-1} mM, 4.20×10^{-2} mM, 1.39×10^{-3} mM, respectively). Scale bar: 5 μm

5.4.8 Dual Color Experiment for Visualizing Growth of New Assemblies

To gain more insight into how two species of assemblies interact and how co-assemblies grow, a dual color experiment utilizing two chromophores is designed to visualize the dynamics of the growth of new assemblies. Since Alexa 633 and Rhodamine 110 (Rh110) have minimal spectral overlap (**Figure 5-27**), Alexa 633 is used to bind the assemblies in the seeded NFF network, and Rh110 is covalently attached to the seeds for detecting the growth in the seeding event.

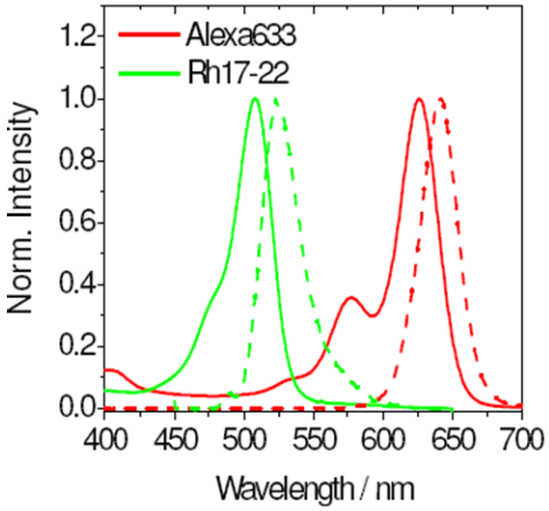


Figure 5-27 Absorbance and emission spectra of Alexa 633 and Rh110 show minimal spectral overlap of the fluorescence emission (Yan Liang et al., 2010).

The Rh110 labeled peptide seeds are prepared by replacing the lysine residue of Aβ(16-22) (KLVFFAE) peptide with Rh110 via standard solid phase synthesis developed previously (Y. Liang, Guo, et al., 2008), which gives rise to a fluorescent peptide Rh17-22. It self-assembles to nanofibers in acetonitrile/ water (2/3, v/v) with 0.1% TFA (Y. Liang, 2009). Further, mixing

Rh17-22 with E22L nanotubes at molar ratio of 1: 250 convert those nanofibers to homogeneously fluorescent nanotubes (**Figure 5-28**), which are indistinguishable from those obtained with pure E22L by TEM, indicating that templating effect of E22L plays a major role in determining the final morphology of the mixed assemblies. These homogeneously fluorescent nanotubes allows for direct observation of assemblies under fluorescent microscopy.

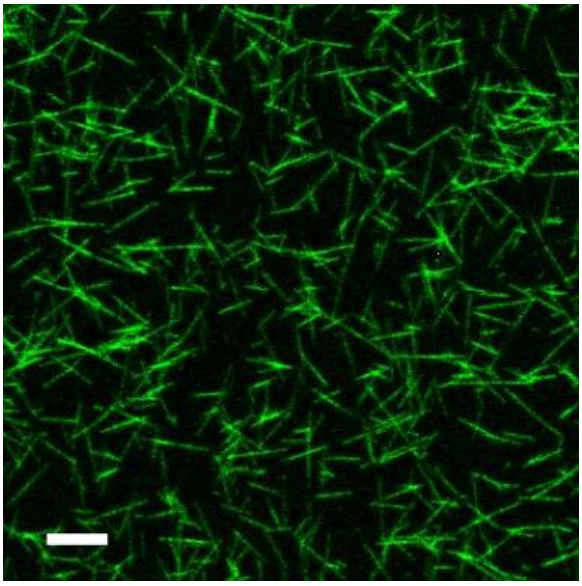


Figure 5-28 Fluorescent nanotubes resulted from mixing Rh17-22 with E22L nanotubes at molar ratio of 1: 250. The concentration of Aβ(16-22) is 2mM and the concentration of Rh17-22 is 8μM. Scale bar: 10 μm

To make the fluorescent nanotubes as more effective seeds, they are exposed to bath sonication for 30 minutes to generate multiple short nanotubes with reactive ends and then diluted by 20 folds for preparing samples for fluorescence imaging (**Figure 5-29)**.

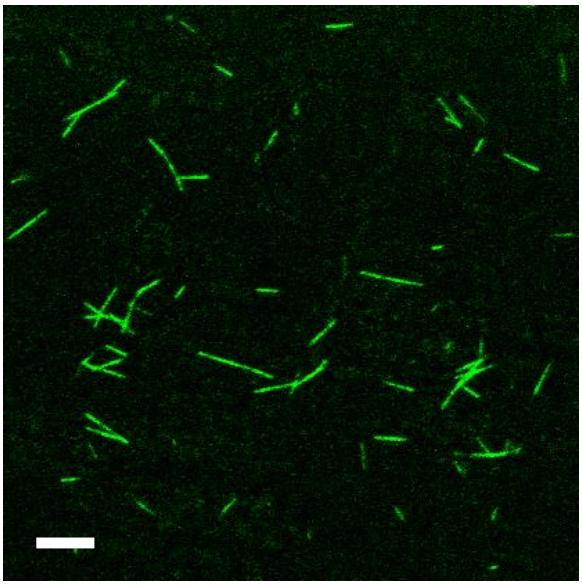


Figure 5-29 Fluorescent microscopy image of Rhodamine labeled peptide seeds prepared by bath sonication. Scale bar: 1 μm

The rhodamine fluorescent nanotube seeds are then added in NFF network with co-assembly pre-stained by Alexa dye (shown in **Figure 5-26**). Upon adding seeds to NFF-DCN at molar ratio of 1:8, the mixture is examined within an hour using synchronized two-channel imaging at 530 nm and 645 nm at their emission wavelength. In contrast to Alexa 633, rhodamine labeled peptide does not stain pre-existing assemblies in NFF-DCN but can be a probe for where the new assemblies grow. As shown in **Figure 5-30**, the red colored peptides are the pre-stained assemblies in the NFF-DCN, while the green colored peptides represent the seeds. It appears that all the red pre-stained assemblies grow at the places where the green seeds are. No randomly distributed rhodamine labeled peptide is observed. In the zoomed-in **Figure 5-31**, it can be seen more clearly that the green and the red assemblies associate tightly with each other. The network assemblies (red) grow at the end of Rhodamine labeled peptide (green), indicating the templating

effect exerted by the seeds. The yellow color coded peptides probably indicate the region where the dynamic exchange of building blocks in the assemblies occurs.

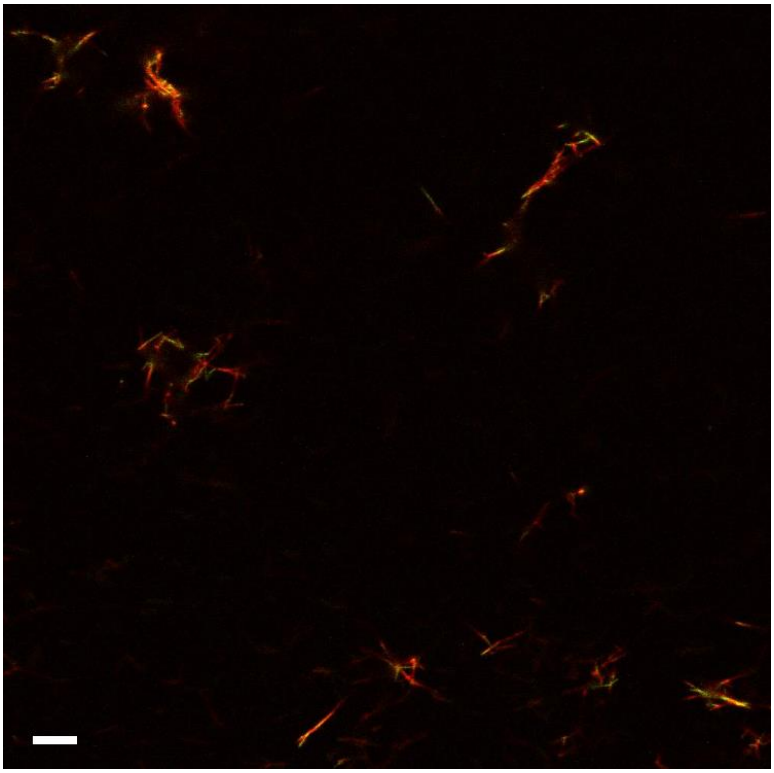


Figure 5-30 Fluorescence microscopy image of mixture of pre-stained assemblies in NFF-DCN and Rhodamine labeled peptide seeds (molar ratio 8:1; concentrations are 0.37 mM, 46µM, respectively). Scale bar: 10 µm

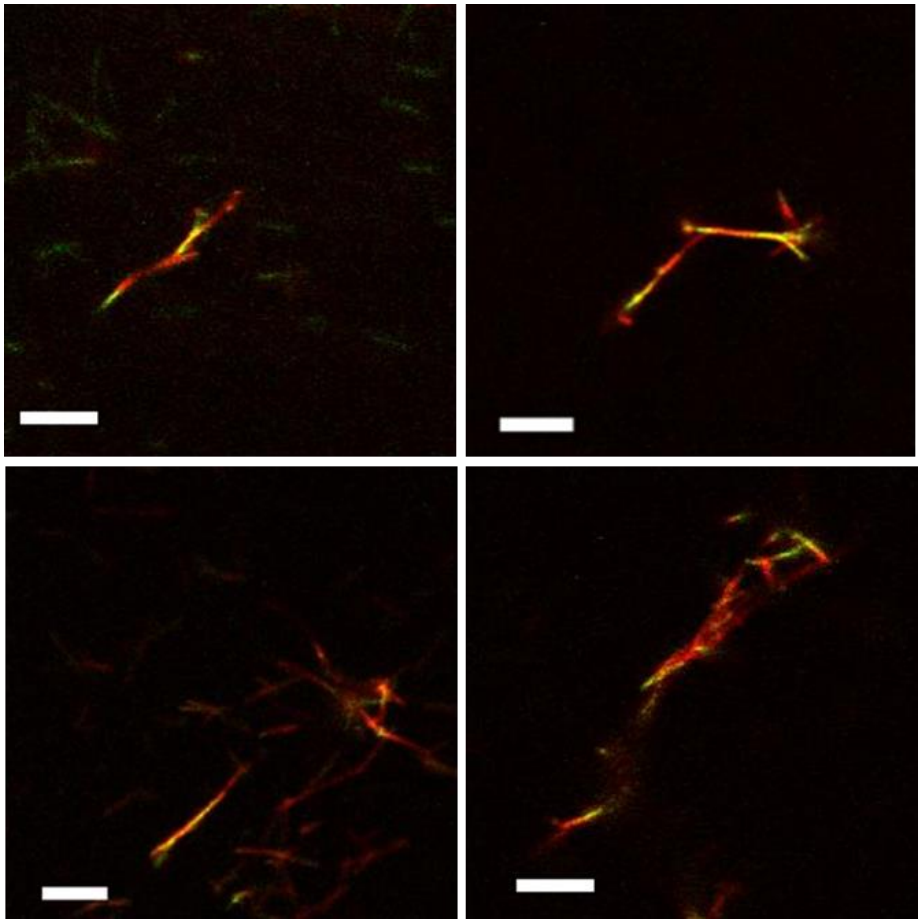


Figure 5-31 Fluorescence microscopy image of mixture of pre-stained assemblies in NFF-DCN and Rhodamine labeled peptide seeds (molar ratio 8:1; concentrations are 0.37 mM, 46µM, respectively). Scale bar: 5 µm

5.5 Discussions

The IMS-MS analysis of dynamic networks within an hour provides evidence for fast noncovalent oligomerization of building blocks (linear trimers) on the pathway preceding the emergence of particles and fibers. Interestingly, no oligomeric state of cyclic dimer is observed, which may be explained as the cyclic dimer is not a typical amphiphilic molecule and the two

positively charged N,N-acetal rings might repulse cyclic dimers to associate as oligomers, whereas amphiphilicity of the monomer, linear dimer and linear trimer makes them susceptible to the noncovalent oligomerization and particle formation. The tendency of cyclic dimers remain in the solution instead of partitioning into early-stage particles implicates the cyclic dimer as a transient species and consumed by linear trimer templates. It also helps to explain the phenomenon described in Chapter 3, where no perceptible change of cyclic species in NF-DCN during the transition from solution phase to particle phase.

The emergence of particles and molten globules in the dynamic networks preceding the fibers and their disappearance due to growth of fibers supports the model that the particles and molten globules are en route to mature assemblies.

The twisted fiber formation due to fast linear trimer condensation and accumulation in NFF-DCN at pH 5.1 suggests that twisted fibers as kinetic product formed at faster rate, whereas, NFF-DCN at pH 4.6 and 7.6 with slower rate for condensation generated straight fibers as thermodynamic product. This is also supported by the observation in NF-DCN at pH 4, where the twisted fibers are consumed by straight fibers when the latter one continued to grow. These data lead to a model proposed in **Figure 5-32**, which implies the two kinds of fibers adopt two different pathways that result from different nuclei. The nuclei resulting in straight fibers need longer time to allow molecular recognition via hydrogen bonding through backbone/ side chains. However, the fast formed nuclei for twisted fibers resulted in less rigid structures. The cross-seeding experiments of using either twisted or straight fibers can induce the network to adopt the template's morphology, suggesting their pathways can be altered by addition of seeds.

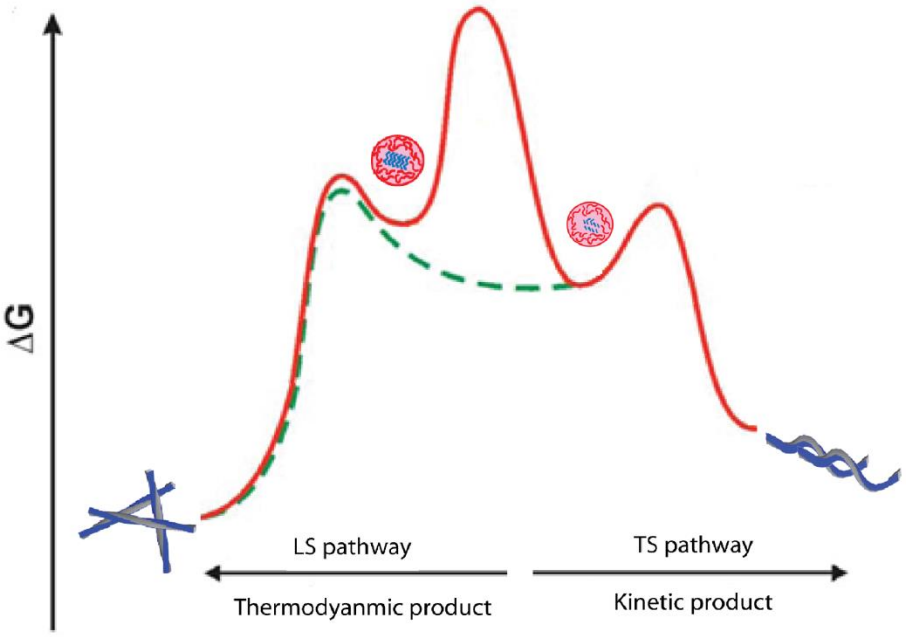


Figure 5-32 Proposed free energy landscape for twisted fibers and straight fibers.

More experiments are required to fully elucidate the assembly pathway. For example, incubating the NFF-DCN (pH 4.6) sample seeded by twisted fibers for longer time to check if twisted fiber will be converted to straight fibers; or using twisted fibers to seed the NFF network after 2 days' incubation with paracrystalline structures formed in molten globules to test whether twisted fiber win or lose the competition. Maybe the seeds will be captured by newly formed paracrystalline structures of straight fibers. Moreover, in the transition of twisted fibers to straight fibers, would twisted fibers disassemble to particles and re-form the long straight ones or go through different mechanism? More detailed knowledge of the kinetics and structural relationship among different species in the networks will help fully understand the pathways and map out the free energy difference of two species under different pH, solvent condition, temperature and concentrations.

Even though the dynamic network's species have altered acetal linkages, they adapt to the templates and show remarkable plasticity to cross-β amyloid infecting the dynamic network, which implies that the N, N-acetal can be positioned well within the H-bonding distance in peptide chain associations. Based on this evidence, we propose a structural model shown in **Figure 5-33** of linear trimer β-sheet assemblies in NFF-DCN. Initial structural models using these structural constraints are consistent with the 4-pyrimidinone system being incorporated within the β-sheet scaffold, but these modules do not position the heterocycle optimally for H-bonding between adjacent peptide stands.

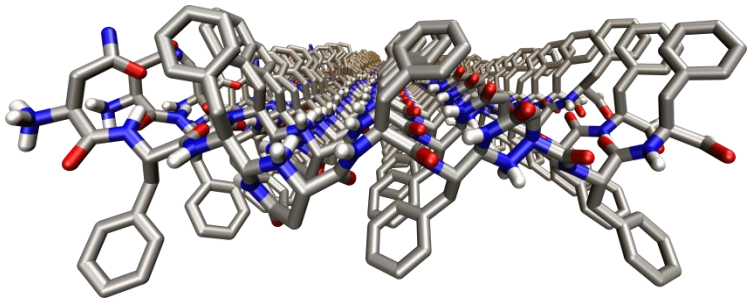


Figure 5-33 Proposed structural model of one sheet of a β-sheet fiber of oligomers in NFF-DCN containing two acetal linkages.

The fact that these networks can be affected by amyloid peptide poses more interesting questions in understanding of amyloid and prion diseases' conformational information transfer. Whether these acetal linkages are able to access more of the secondary structures of amide peptides and create surfaces with catalytic competencies sufficient for the progressive growth for functional selection remains to be determined, but the acetal has proven to be remarkably robust, extending to alternative infectious peptides and possibly the use of multiple monomers to provide access to more interacting networks and complex systems. The simplicity of these peptide-like networks

suggests that other chemical scaffolds, operating either orthogonally to existing biology (Davies, 2011) or coupled to biological functions like the prions, can emerge far more broadly (Rubinov et al., 2012) (Carnall et al., 2010). As the ribosome serves the Darwinian threshold for cellular diversification (C. R. Woese, 2002), these dynamic peptide assemblies provide a chemical threshold for biopolymer emergence through metastable phases.

Chapter 6: Conclusion and Perspectives

Charles Darwin, in his musings on emergence from a warm pond, suggested that life and its origins in biological evolution could be an inherent property of matter. Life is fundamentally a complex dynamic system, fashioned by amino acids, nucleic acids, lipids and other small molecules, which is molecular foundation for biological evolution. Two strategies (genetic and epigenetic information replication) for molecular evolution are manifested in the eukaryotic cell (shown in Chapter 1), but both require the dynamic chemical networks (DCNs) of the cell to produce the biopolymers necessary for evolution. Therefore, designing and construction of self-replicating DCNs is very important in understanding the biopolymer emergence and evolution.

In Chapter 2, I systematically studied the three acetal condensations (N,O-; N,N- and N,S-acetals) in a series of solvents and at different temperatures. The structural investigation of the acetals in this chapter extends our understating of reversible linkages and set foundations for utilizing these linkages for constructing novel dynamic chemical networks in aqueous media. Due to the great diastereoselectivity of N,N-acetal in acetonitrile, this linkage is utilized to build NF-CHO and NFF-CHO networks.

The construction of NF-CHO and NFF-CHO dynamic networks are described in Chapter 3 and 4. These networks cross kinetic barrier in the absence of complex catalysts and allow peptide oligomerization in aqueous solutions and they are environmentally responsive from which specific chain-lengths can be kinetically selected through self-replication. The resulting assemblies are capable of seeding both chain-length and conformational information (Chapter 5), establishing a minimal system for the chemical evolution of biopolymers.

Both viruses and prions are parasitic genotypic systems that require host machinery for informational polymer construction. These networks are capable of autocatalytically producing its own informational polymers. Even though built on a single biopolymer scaffold, the network displays a tension between thermodynamic and kinetic growth dynamics that reflects the forms of information content of biological information transfer in a simpler molecular form.

These peptide networks combine chemical and physical dynamics to achieve a progressive growth of molecular order. A solution phase dynamic chemical network sets up a member distribution sufficient for a physical phase change. The resulting particle phase is metastable based on the thermodynamic frustration created by strongly self-complementary amphiphilic monomers, much as in protein folding transitions. These particles set up transitions to paracrystalline phases that autocatalytically amplify chain-length specific peptide assembly. It offers the possibility of defining a new 'genotypic code', one with the potential of being broadly applicable to the creation of a general materials genome (J.T. Goodwin et al., 2014), just like the metastable phases of prion propagation that create the Darwinian gain-of-function selection in disease (Serio et al., 2000).

With this demonstrated chain length specificity, mixed monomer populations and alternative ligation strategies could move to sequence-specific genotypic selection. Moreover, the surfaces of the paracrystalline phase can further extend the functional capability of the assemblies to include a subsequent gain-of-function relevant to productive feedback in chemical evolution (Jay T. Goodwin et al., 2012).

Reference

Adler-Abramovich, L., Reches, M., Sedman, V. L., Allen, S., Tendler, S. J. B., & Gazit, E. (2006). Thermal and Chemical Stability of Diphenylalanine Peptide Nanotubes: Implications for Nanotechnological Applications. *Langmuir, 22*, 1313-1320.

Adochitei, A., & Drochioiu, G. (2001). Rapid Characterization of Peptide Secondary Structure by FT-IR Spectroscopy *Revue Roumaine de Chimie*, *56*(8), 783-791.

Anthony, N. R., Mehta, A. K., Lynn, D. G., & Berland, K. M. (2014). Mapping amyloid-b(16-22) nucleation pathways using fluorescence lifetime imaging microscopy. *Soft Matter, 10*(23), 4162-4172. doi: 10.1039/c4sm00361f

Auer, S., Ricchiuto, P., & Kashchiev, D. (2012). Two-step nucleation of amyloid fibrils: omnipresent or not? *J Mol Biol, 422*(5), 723-730. doi: 10.1016/j.jmb.2012.06.022

Azema, L., Bathany, K., & Rayner, B. (2010). 2′-O-Appended Polyamines that Increase Triple-Helix-Forming Oligonucleotide Affinity are Selected by Dynamic Combinatorial Chemistry. *ChemBioChem, 11*, 2513-2516.

Bader, R., Bamford, R., Zurdo, J., Luisi, B. F., & Dobson, C. M. (2006). Probing the mechanism of amyloidogenesis through a tandem repeat of the pi3-sh3 domain suggests a generic model for protein aggregation and fibril formation. *J Mol Biol, 356*, 189–208.

Baldan, A. (2002). Progress in Ostwald ripening theories and their applications to nickel-base superalloys. *Journal of Materials Science, 37*, 2171-2202.

Bartel, D. P., & Szostak, J. W. (1993). Isolation of new ribozymes from a large pool of random sequences. *Science, 261*, 1411-1418.

Bernstein, S. L., Dupuis, N. F., Lazo, N. D., Wyttenbach, T., Condron, M. M., Bitan, G., . . . Bowers, M. T. (2009). Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat Chem, 1*(4), 326-331. doi: 10.1038/nchem.247

Besenius, P., Cormack, P. A. G., Ludlow, R. F., Otto, S., & Sherrington, D. C. (2008). Polymer-supported cationic templates for molecular recognition of anionic hosts in water. *Chemical Communications*, 2809-2811.

Blake, C., & Serpell, L. (1996). Synchrotron X-ray Studies Suggest that the Core of the Transthyretin Amyloid Fibril is a Continuous b-sheet helix. *Structure, 4*, 989-998.

Bru, M., Alfonso, I., Bolte, M., Burguete, M. I., & Luis, S. V. (2011). Structurally disfavoured pseudopeptidic macrocycles through anion templation. *Chemical Communications, 47*, 283-285.

Bru, M., Alfonso, I., Burguete, M. I., & Luis, S. V. (2006). Anion-Templated Syntheses of Pseudopeptidic Macrocycles. *Angew. Chem. Int. Ed, 45*, 6155-6159.

Carnall, J. M., Waudby, C. A., Belenguer, A. M., Stuart, M. C., Peyralans, J. J., & Otto, S. (2010). Mechanosensitive self-replication driven by self-organization. *Science, 327*(5972), 1502-1506. doi: 10.1126/science.1182767

Carulla, N., Zhou, M., Arimon, M., Gairi, M., Giralt, E., Robinson, C. V., & Dobson, C. M. (2009). Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc Natl Acad Sci U S A, 106*(19), 7828-7833. doi: 10.1073/pnas.0812227106

Cavalier-Smith, T. (2006). Cell evolution and Earth history: stasis and revolution. *Philosophical Transactions of the Royal Society of London, Series B, 361*, 969–1006.

Cheon, M., Chang, I., Mohanty, S., Luheshi, L. M., Dobson, C. M., Vendruscolo, M., & Favrin, G. (2007). Structural reorganisation and potential toxicity of oligomeric species formed during the assembly of amyloid fibrils. . *PLoS Comput. Biol., 3*, 1727-1738.

Cheon, M., Favrin, G., Chang, I., Dobson, C. M., & Vendruscolo, M. (2008). Calculation of the free energy barriers in the oligomerisation of Aβ peptide fragments. *Front. Biosci., 13*, 5614−5622.

Childers, W. S. (2010). Dissertation.

Childers, W. S., Anthony, N. R., Mehta, A. K., Berland, K. M., & Lynn, D. G. (2012). Phase Networks of Cross-β Peptide Assemblies. *Langmuir, 28*(15), 6386-6395. doi: 10.1021/la300143j

Corbett, P. T., Leclaire, J., Vial, L., West, K. R., Wietor, J.-L., Sanders, J. K. M., & Otto, S. (2006). Dynamic Combinatorial Chemistry. *Chemical Reviews, 106*, 3652-3711.

Cougnon, F. B. L., & Sanders, J. K. M. (2011). Evolution of Dynamic Combinatorial Chemistry. *Acc. Chem. Res., 45*(12), 2211-2221.

Creaser, C., & Thomas, P. (2004). Ion mobility spectrometry: a review. Part 1. Structural analysis by mobility measurement. *The Analyst, 129*(984-994).

Custelcean, R. (2012). Dynamic chemistry of anion recognition. *Top Curr Chem., 322*, 193-216.

Davies, P. C. (2011). Searching for a shadow biosphere on Earth as a test of the 'cosmic imperative'. *Philos Trans A Math Phys Eng Sci, 369*(1936), 624-632. doi: 10.1098/rsta.2010.0235

de Bruin, B., Hauwert, P., & Reek, J. N. H. (2006). Dynamic Combinatorial Chemistry: The Unexpected Choice of Receptors by Guest Molecules. *Angew. Chem., Int. Ed., 45*, 2660–2663.

Demetriades, M., Leung, I. K. H., Chowdhury, R., Chan, M. C., McDonough, M. A., Yeoh, K. K., . . . Schofield, C. J. (2012). Dynamic Combinatorial Chemistry Employing Boronic Acids/Boronate Esters Leads to Potent Oxygenase Inhibitors. *Angewandte Chemie International Edition, 51*, 6672-6675.

Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., & Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. *Trends in Biochemical Sciences, 25*(7), 331–339.

Dobson, C. M. (1994). Solid evidence for molten globules. *Curr. Biol., 4*(7), 636-640.

Dong, J., Lu, K., Lakdawala, A., Mehta, A. K., & Lynn, D. G. (2006). Controlling amyloid growth in multiple dimensions. *Amyloid, 13*, 206.

Eichinger, B. E. (2000). Cyclization in reversible and irreversible step-growth polymerizations. *Computational and Theoretical Polymer Science, 10*, 83-88.

Eliel, E. L., & Wilen, S. H. (1994). *Stereochemistry of Organic Compounds*. New York: Wiley.

Fernandez, A., Kardos, J., & Goto, Y. (2003). Protein Folding: Could Hydrophobic Collapse be Coupled with Hydrogen-bond Formation? *FEBS Letters, 536*, 187-192.

Fernandez, A., Sosnick, T. R., & Colubri, A. (2002). Dynamics of Hydrogen Bond Desolvation in Protein Folding. *Journal of Molecular Biology, 321*, 659–675.

Fife, T. H., & Hutchins, J. E. C. (1980). General-Acid-Catalyzed Ring Opening of Oxazolidines. Hydrolysis of 2-[4-(Dimethylamino)styryl]-N-phenyl-1,3-oxazolidine. *J. Org. Chem, 45*, 2099-2104.

Frederix, P. W., Ulijn, R. V., Hunt, N. T., & Tuttle, T. (2011). Virtual Screening for Dipeptide Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *J Phys Chem Lett, 2*(19), 2380-2384. doi: 10.1021/jz2010573

Gao, F., Wang, Y., Qiu, Y., Li, Y., Sha, Y., Lai, L., & Wu, H. (2002). b-turn formation by a six-residue linear peptide in solution. *J. Peptide Res., 60*, 75-80.

Giuseppone, N., & Lehn, J. M. (2006). Protonic and temperature modulation of constituent expression by component selection in a dynamic combinatorial library of imines. *Chemistry, 12*(6), 1715-1722. doi: 10.1002/chem.200501038

Godoy-Alcántar, C., Yatsimirsky, A. K., & Lehn, J. M. (2005). Structure-stability correlations for imine formation in aqueous solution. *Journal of Physical Organic Chemistry, 18*(10), 979-985. doi: 10.1002/poc.941

Goodwin, J. T., & Lynn, D. G. (1992). Template-Directed Synthesis: Use of a Reversible Reaction. *J.AM.CHEM.SOC, 114*(23), 9197-9198.

Goodwin, J. T., Mehta, A. K., & Lynn, D. G. (2012). Digital and Analog Chemical Evolution. *Acc. Chem. Res., 45*(12), 2189-2199. doi: 10.1021/ar300214w

Goodwin, J. T., Walker, S. I., Amin, S., Armbrust, G., Burrows, C. J., & Lynn, D. G. (2014). Alternative Chemistries of Life: Empirical Approaches. *A report from a National Meeting on Alternative Chemistries of Life: NASA & NSF*.

Gosal, W. S., Morten, I. J., Hewitt, E. W., Smith, D. A., Thomson, N. H., & Radford, S. E. (2005). Competing Pathways Determine Fibril Morphology in the Selfassembly of b2-Microglobulin into Amyloid. *Journal of Molecular Biology, 351*, 850-864.

Habicht, G., Haupt, C., Friedrich, R. P., Hortschansky, P., Sachse, C., Meinhardt, J., . . . Fandrich, A. M. (2007). Directed Selection of a Conformational Antibody Domain that Prevents Mature Amyloid Fibril Formation by Stabilizing Ab Protofibrils. *Proc. Natl. Acad. Sci. USA., 104*, 19232-19237.

Harish, A., & Caetano-Anolles, G. (2012). Ribosomal history reveals origins of modern protein synthesis. *PLoS ONE, 7*(3), e32776.

Harper, J. D., & Lansbury, P. T., Jr. (1997). Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annual Review of Biochemistry, 66*, 385-407.

Hoffmann, R. W. (1989). Allylic 1,3-Strain as a Controlling Factor in Stereoselective Transformations. *Chemical Reviews, 89*, 1841-1860.

Huc, I., & Lehn, J. M. (1997). Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly. *Proc. Natl. Acad. Sci. U.S.A., 94*(6), 2106-2110.

Jarrett, J. T., & Lansbury, P. T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell, 73*, 1055-1058.

Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E., & Bartel, D. P. (2001). RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension. *Science, 292*, 1319-1325.

Joyce, G. F. (2004). Directed evolution of nucleic acid enzymes. *Annual Review of Biochemistry, 73*, 791-836.

Kanu, A. B., Dwivedi, P., Tam, M., Matz, L., & Hill, H. H. (2008). Ion mobility-mass spectrometry. *Journal of Mass Spectrometry, 43*(1), 1-22.

Kayed, R., Head, E., Sarsoza, F., Necula, M., Margol, L., Wu, J., . . . Glabe, C. G. (2007). Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibers and fibrillar oligomers that is distinct from prefibrillar oligomers. *Mol. Neurodegeneration, 2*, 1-11.

Kaylor, J., Bodner, N., Edridge, S., Yamin, G., Hong, D. P., & Fink, A. L. (2005). Characterization of Oligomeric Intermediates in a-Synuclein Fibrillation: FRET Studies of Y125W/Y133F/Y136F a-Synuclein. *Journal of Molecular Biology, 353*, 357-372.

Kim, W., & Hecht, M. H. (2006). Generic Hydrophobic Residues Are Sufficient to Promote Aggregation of the Alzheimer's Ab42 Peptide. *Proc. Natl. Acad. Sci. U.S.A., 103*, 15824–15829.

Kloniecki, M., Jablonowska, A., Poznanski, J., Langridge, J., Hughes, C., Campuzano, I., . . . Dadlez, M. (2011). Ion mobility separation coupled with MS detects two structural states of Alzheimer's disease Abeta1-40 peptide oligomers. *J Mol Biol, 407*(1), 110-124. doi: 10.1016/j.jmb.2011.01.012

Kumar, S., & Walter, J. (2011). Phosphorylation of amyloid beta (Aβ) peptides – A trigger for formation of toxic aggregates in Alzheimer's disease. *3*(803-812).

Lehn, J. M. (2007). From supramolecular chemistry towards constitutional dynamic chemistry and adaptive chemistry. *Chemical Society Reviews, 36*(2), 151-160. doi: 10.1039/b616752g

LeVine, H. (1999). *Methods in Enzymology, 309*, 274.

Li, J. (2014). *Self-Assembly in Complex Chemical Systems.* (Ph. D.), University of Groningen, The Netherlands.

Li, J., Browning, S., Mahal, S. P., Oelschlegel, A. M., & Weissmann, C. (2010). Darwinian evolution of prions in cell culture. *Science, 327*(5967), 869-872. doi: 10.1126/science.1183218

Li, J., Nowak, P., & Otto, S. (2013). Dynamic combinatorial libraries: from exploring molecular recognition to systems chemistry. *Journal of the American Chemical Society, 135*(25), 9222-9239. doi: 10.1021/ja402586c

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

Li, X., Zhan, Z.-Y. J., Knipe, R., & Lynn, D. G. (2002). DNA-Catalyzed Polymerization. *J.AM.CHEM.SOC, 124*(5), 746-747.

Liang, C., Ni, R., Smith, J. E., Childers, W. S., Mehta, A. K., & Lynn, D. G. (2014). Kinetic intermediates in amyloid assembly. *Journal of the American Chemical Society, 136*(43), 15146-15149. doi: 10.1021/ja508621b

Liang, Y. (2009). *Energetic Contribution of Amyloid Self-assembly.* (Ph. D.), Emory University.

Liang, Y., Guo, P., Pingali, S. V., Pabit, S., Thiyagarajan, P., Berland, K. M., & Lynn, D. G. (2008). Light harvesting antenna on an amyloid scaffold. *Chem. Commun., 48*, 6522-6524. doi: 10.1039/b814262a

Liang, Y., Lynn, D. G., & Berland, K. M. (2010). Direct Observation of Nucleation and Growth in Amyloid Self-Assembly. *J. AM. CHEM. SOC., 132*(18), 6306-6308.

Liang, Y., Pingali, S. V., Jogalekar, A. S., Snyder, J. P., Thiyagarajan, P., & Lynn, D. G. (2008). Cross-Strand Pairing and Amyloid Assembly. *Biochemistry, 47*, 10018-10026.

Lifshitz, I. M., & Slyozov, V. V. (1961). *J. Phys. Chem. Solids, 19*, 35−50.

Loo, J. A. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews, 16*, 1–23.

Lu, K., Jacob, J., Thiyagarajan, P., Conticello, V. P., & Lynn, D. G. (2003). Exploiting Amyloid Fibril Lamination for Nanotube Self-Assembly. *J. AM. CHEM. SOC., 125*(6391-6393).

Malakoutikhah, M., Peyralans, J. J., Colomb-Delsuc, M., Fanlo-Virgos, H., Stuart, M. C., & Otto, S. (2013). Uncovering the selection criteria for the emergence of multi-building-block replicators from dynamic combinatorial libraries. *Journal of the American Chemical Society, 135*(49), 18406-18417. doi: 10.1021/ja4067805

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A COMPARISON OF THREE METHODS FOR SELECTING VALUES OF INPUT VARIABLES IN THE ANALYSIS OF OUTPUT FROM A COMPUTER CODE. *Technometrics, 21*(2), 239-245. doi: 10.2307/1268522

McNaughton, B. R., Gareiss, P. C., & Miller, B. L. (2007). Identification of a Selective Small-Molecule Ligand for HIV-1 Frameshift-Inducing Stem-Loop RNA from an 11,325 Member Resin Bound Dynamic Combinatorial Library. *129*, 11306-11307.

Mehta, A. K., Lu, K., Childers, W. S., Liang, Y., Dublin, S. N., Dong, J., . . . Lynn, D. G. (2008). Facial Symmetry in Protein Self-Assembly. *J. AM. CHEM. SOC., 130*, 9829-9835.

Meijer, J. T., Roeters, M., Viola, V., Lowik, D. W. P. M., Vriend, G., & Hest, J. C. M. (2007). Stabilization of Peptide Fibrils by Hydrophobic Interaction. *Langmuir, 23*, 2058-2063.

Metz, T. O., Page, J. S., Baker, E. S., Tang, K., Ding, J., Shen, Y., & Smith, R. D. (2008). High-resolution separations and improved ion production and transmission in metabolomics. *Trends in Analytical Chemistry, 27*(3), 205. doi: 10.1016/j.trac.2007.11.003

Miller, S. M., & Rawlings, J. B. (1994). Model Identification and Control Strategies for Batch Cooling Crystallizers. *AIChE Journal, 40*(8), 1312-1327.

Mills, D. R., Peterson, R. L., & Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences, USA, 58*, 217-224.

Modler, A. J., Gast, K., Lutsch, G., & Damaschun, G. (2003). Assembly of Amyloid Protofibrils via Critical Oligomers—A Novel Pathway of Amyloid Formation. *Journal of Molecular Biology, 325*, 135-148.

Muschol, M., & Rosenberger, F. (1997). Liquid−liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/ crystallization. *J. Chem. Phys., 107*, 1953-1962.

Naiki, H., & Gejyo, F. (1999). Kinetic analysis of amyloid fibril formation. *Methods in Enzymology, 309*, 305-318.

Necula, M., Kayed, R., Milton, S., & Glabe, C. G. (2007). Small molecule inhibitors of aggregation indicate that amyloid beta oligomerization and fibrillization pathways are independent and distinct. *Journal of Biological Chemistry, 282*, 10311–10324.

Nguyen, P. H., Li, M. S., Stock, G., Straub, J. E., & Thirumalai, D. (2007). Monomer adds to preformed structured oligomers of Abeta-peptides by a two-stage dock-lock mechanism. *Proc. Natl. Acad. Sci. USA., 104*, 111–116.

Ostwald, W. (1901). *Phys. Chem, 385*.

Otto, S., Furlan, R. L., & Sanders, J. K. (2002). Selection and amplification of hosts from dynamic combinatorial libraries of macrocyclic disulfides. *Science, 297*(5581), 590-593. doi: 10.1126/science.1072361

Pace, C. N., Shirley, B. A., Mcnutt, M., & Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *FASEB Journal, 10*(1), 75–83.

Pellarin, R., & Caflisch, A. (2006). Interpreting the aggregation kinetics of amyloid peptides. *Journal of Molecular Biology, 360*, 882–892.

Petty, S. A., & Decatur, S. M. (2005). Experimental evidence for the reorganization of beta-strands within aggregates of the Ab(16–22) peptide. *J. Am. Chem. Soc, 127*, 13488–13489.

Plakoutsi, G., Bemporad, F., Calamai, M., Taddei, N., & Dobson, C. M. (2005). Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. *Journal of Molecular Biology, 351*, 910–922.

Poirier, M. A., Li, H., Macosko, J., Cai, S., Amzel, M., & Ross, C. A. (2002). Huntingtin Spheroids and Protofibrils as Precursors in Polyglutamine Fibrilization. *Journal of Biological Chemistry, 277*, 41032–41037.

Press, W. H. T., S. A.; Vetterling, W. T.; Flannery, B. P. (1988). Numerical recipes in C: the art of scientific computing. *Cambridge University Press*, 698-699.

Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinos, K., . . . Scrivens, J. H. (2007). An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom, 261*, 1-12.

Reches, M., & Gazit, E. (2006). Designed aromatic homo-dipeptides: formation of ordered nanostructures and potential nanotechnological applications. *Phys. Biol., 3*, S10-S19.

Rose, G. D., Fleming, P. J., Banavar, J. R., & Maritan, A. (2006). A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. U.S.A., 103*(45), 16623–16633.

Rubinov, B., Wagner, N., Matmor, M., Regev, O., Ashkenasy, N., & Ashkenasy, G. (2012). Transient fibril structures facilitating nonenzymatic self-replication. *ACS Nano, 6*(9), 7893-7901. doi: 10.1021/nn302223v

Scanlon, S., & Aggeli, A. (2008). Self-assembling peptide nanotubes. *Nano Today, 3*(3-4), 22-30. doi: 10.1016/s1748-0132(08)70041-0

Schmeing, T. M., oorhees, R. M., Kelley, A. C., Gao, Y. G., Murphy, F. V., Weir, J. R., & Ramakrishnan, V. (2009). The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science, 326*, 688-694.

Schopf, J. W. (2006). Fossil evidence of Archaean life. *Philosophical Transactions of the Royal Society of London, Series B, 361*, 869–885.

Scott, D. E., Dawes, G. J., Ando, M., Abell, C., & Ciulli, A. (2009). A fragment-based approach to probing adenosine recognition sites by using dynamic combinatorial chemistry. *ChemBioChem, 10*, 2772-2779.

Seebach, D., Aebi, J. D., Gander-Coquoz, M., & Naef, R. (1987). Stereoselektive Alkylierung an C(α) von Serin, Glycerinsäure, Threonin und Weinsäure über heterocyclische Enolate mit exocyclischer Doppelbindung. *Helvetica Chimica Acta, 70*(4), 1194-1216.

Seebach, D., Lamatsch, B., Amstutz, R., Beck, A. K., Dobler, M., Martin Egli,, Robert Fitzi, M. G., Bernard0 Herrad6n3), Pirmin C. Hidber, John J. Irwin, Rita Locher,, Miguel Maestro, T. M., Antonio Mourilo, Elmar Pfammatter, Dietmar A. Plather,, & Christof Schickli, W. B. S., Paul Seiler, and Gerhard Stncky. (1992). Structure and reactivity of five- and six-ring N,N-, N,O-, and O,O-acetals: A lesson in allylic 1,3-strain. *Helvetica Chimica Acta, 75*.

Seebach, D., Sommerfeld, T. L., Jiang, Q., & Venanzi, L. M. (1994). Preparation of Oxazolidine-Containing Peptides: Unusual effects in Rh(III)-catalyzed acetalizations of aldehydes with urethane-protected serine and threonine esters and with dipeptides containing serine or threonine residues at the N-terminus. *Helvetica Chimica Acta, 77*(5), 1313-1330.

Seebach, D., Zimmermann, J., Gysel, U., Ziegler, R., & Ha, T. (1988). Totally Stereoselective Additions to 2,6-Disubstituted 1,3-Dioxin-4-ones (Chiral Acetoacetic Acid Derivatives). Synthetic and Mechanistic Aspects of Remote Stereoselectivity. *J. AM. CHEM. SOC., 110*, 4763-4772.

Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., & Moslehi, J. J. (2000). Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science, 289*, 1317–1321.

Shi, B. L., Stevenson, R., Campopiano, D. J., & Greaney, M. F. (2006). *J. AM. CHEM. SOC., 128*, 8459−8467.

Sikorski, P., Atkins, E. D. T., & Serpell, L. C. (2003). Structure and Texture of Fibrous Crystals Formed by Alzheimer's Ab(11–25) Peptide Fragment. *Structure, 11*, 915-926.

Smith, R. D., & Light-Wahl, K. J. (1993). The observation of non-covalent interactions in solution by electrospray ionization mass spectrometry: promise, pitfalls and prognosis. *Biol. Mass Spectrom., 22*, 493–501.

Stine, W. B., Dahlgren, K. N., Krafft, G. A., & LaDu, M. J. (2003). In vitro Characterization of Conditions for Amyloid-b Peptide Oligomerization and Fibrillogenesis. *Journal of Biological Chemistry, 278*, 11612–11622.

Stoddart, J. F. (1971). *Stereochemistry of Carbohydrates*. New York: Wiley-Interscience.

Sulloway, F. J. (1982). Darwin and his finches: The evolution of a legend. *J History Biol, 15*, 1-53.

Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., & Blake, C. C. F. (1997). Common Core Structure of Amyloid Fibrils by Synchrotron X-ray Diffraction. *Journal of Molecular Biology, 273*, 729-739.

Surewicz, W. K., & Mantsch, H. H. (1988). New insight into protein secondary structure from resolution-enhanced infrared spectra. *Biochimica et Biophysica Acta, 952*, 115-130.

Szilagyi, L., & Gyorgydeak, Z. (1979). Comments on the putative stereoselectivity in cysteine-aldehyde reactions. Selective C(2) inversion and C(4) epimerization in thiazolidine-4-carboxylic acids. *J. AM. CHEM. SOC., 101*(2), 427-432.

Theobald, D. L. (2010). A formal test of the theory of universal common ancestry. *Nature, 465*, 219-223.

Thiel, J., Yang, D., Rosnes, M. H., Liu, X., Yvon, C., Kelly, S. E., . . . Cronin, L. (2011). Observing the Hierarchical Self-Assembly and Architectural Bistability of Hybrid Molecular Metal Oxides Using Ion-Mobility Mass Spectrometry. *Angewandte Chemie International Edition, 50*, 8871-8875. doi: 10.1002/anie.201102340

Valade, A., Urban, D., & Beau, J.-M. (2006). Target-Assisted Selection of Galactosyltransferase Binders from Dynamic Combinatorial Libraries. An Unexpected Solution with Restricted Amounts of the Enzyme. *ChemBioChem, 7*, 1023-1027.

Valentine, S. J., Plasencia, M. D., Liu, X., Krishnan, M., Naylor, S., Udseth, H. R., . . . Clemmer, D. E. (2006). *J. Proteome Res., 5*, 2977–2984.

Valters, R. E., & Flitsch, W. (1985). Ring-Chain Tautomerism.

Vlad, C., Iurascu, M. I., Slamnoiu, S., Hengerer, B., & Przybylski, M. (2012). Characterization of oligomerization-aggregation products of neurodegenerative target proteins by ion mobility mass spectrometry. *Methods Mol. Biol., 896*, 399−412.

Voytek, S. B., & Joyce, G. F. (2009). Niche partitioning in the coevolution of 2 distinct RNA enzymes. *Proceedings of the National Academy of Sciences, USA, 106*, 7780-7785.

Wagner, C. Z. (1961). *Elektrochem, 65*, 581−591.

Wang, F., Richards, V. N., Shields, S. P., & Buhro, W. E. (2014). Kinetics and Mechanisms of Aggregative Nanocrystal Growth. *Chemistry of Materials, 26*(1), 5-21. doi: 10.1021/cm402139r

Wessjohann, L. A., Rivera, D. G., & Leo´n, F. (2007). Freezing Imine Exchange in Dynamic Combinatorial Libraries with Ugi Reactions: Versatile Access to Templated Macrocycles. *Organic Letters, 9*(23), 4733-4736.

Wickner, R. B., Edskes, H. K., Bateman, D. A., Kelly, A. C., Gorkovskiy, A., Dayani, Y., & Zhou, A. (2014). Amyloid diseases of yeast: prions are proteins acting as genes. *Essays Biochem, 56*, 193-205. doi: 10.1042/bse0560193

Woese, C. R. (2002). On the evolution of cells. *Proc. Natl. Acad. Sci. U.S.A., 99*, 8742-8747.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences, USA, 87*, 4576-4579.

Wolfe, L. S. (2010). Protein-induced photophysical changes to the amyloid indicator dye thioflavin T. *Proceedings of the National Academy of Sciences, USA, 107*, 16863.

Woods, L. A., Radford, S. E., & Ashcroft, A. E. (2013). Advances in ion mobility spectrometry-mass spectrometry reveal key insights into amyloid assembly. *Biochim Biophys Acta, 1834*(6), 1257-1268. doi: 10.1016/j.bbapap.2012.10.002

Wright, M. C., & Joyce, G. F. (1997). Continuous in vitro evolution of catalytic function. *Science, 276*, 614-617.

Wu, L., McElheny, D., Setnicka, V., Hilario, J., & Keiderling, T. A. (2012). Role of different beta-turns in beta-hairpin conformation and stability studied by optical spectroscopy. *Proteins, 80*(1), 44-60. doi: 10.1002/prot.23140

Wuitschik, G., Carreira, E. M., Wagner, B., Fischer, H., Parrilla, I., Schuler, F., . . . Muller, K. (2010). Oxetanes in drug discovery: structural and synthetic insights. *Journal of Medicinal Chemistry, 53*(8), 3227-3246. doi: 10.1021/jm9018788

Young, L. M., Cao, P., Raleigh, D. P., Ashcroft, A. E., & Radford, S. E. (2014). Ion mobility spectrometry-mass spectrometry defines the oligomeric intermediates in amylin amyloid formation and the mode of action of inhibitors. *Journal of the American Chemical Society, 136*(2), 660-670. doi: 10.1021/ja406831n

Zameo, S., Vauzeilles, B., & Beau, J.-M. (2006). Direct Composition Analysis of a Dynamic Library of Imines in an Aqueous Medium. *European Journal of Organic Chemistry*, 5441–5444. doi: 10.1002/ejoc.200600859

Zhan, Z.-Y. J., & Lynn, D. G. (1997). Chemical Amplification through Template-Directed Synthesis. *J.AM.CHEM.SOC, 119*, 12420-12421.

Zhan, Z. J., & Lynn, D. G. (1997). Chemical Amplification through Template-Directed Synthesis. *J. AM. CHEM. SOC., 119*, 12420-12421.