KINETIC AND STRUCTURAL EVOLUTION OF FUNCTIONAL PEPTIDE ASSEMBLING NETWORKS

A Dissertation
Presented to
The Academic Faculty

by

Ming-Chien Hsieh

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemical & Biomolecular Engineering

Georgia Institute of Technology
August 2017

COPYRIGHT © 2017 BY MING-CHIEN HSIEH

KINETIC AND STRUCTURAL EVOLUTION OF FUNCTIONAL PEPTIDE ASSEMBLING NETWORKS

Approved by:

Dr. Martha A. Grover, Advisor
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. David G. Lynn, co-Advisor
Department of Chemistry
Emory University

Dr. Julie A. Champion
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. Peter J. Ludovice
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. Nicholas V. Hud
School of Chemistry & Biochemistry
Georgia Institute of Technology

Date of the defense: June 30, 2017

[To my family]

ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Grover and Dr. Lynn, from the bottom of my heart. Through these years, all the supports and teachings they provided made me become a scientist. Their insightful advises and constructive criticisms always brought me back to the right pathway whenever I felt that I cannot progress anymore.

I also want to take this opportunity to thank my thesis committees: Dr. Julie Champion, Dr. Peter Ludovice, and Dr. Nicholas Hud. Their valuable comments and informative feedbacks provided me new insight and drove me to have a better development of my work.

I am very grateful to my collaborators, Dr. Chenrui Chen, Dr. Tolulope Omosun, Dr. Li Zhang, and Dr. Junjun Tan. Their experimental analyses provided me enough information to build and construct the mathematic models. Through the discussion with them, I learn how to analyze multidiscipline projects and explain my ideas to people with different background.

I am very grateful to Dr. Anil Mehta for his friendship and priceless teaching about the knowledge in peptide self-assembly. He is always patient when answering my questions and always explains tough ideas in an easy and understandable way.

I would also like to thank my labmates in Georgia Tech: Dr. Andres Hernandez Moreno, Dr. Huayu Li, Dr. Daniel Griffin, Dr. Xun Tang, Dr. Christine He, Dr. Sheng-Sheng Yu, Dr. Michelle Kuykendal, Dr. Michael Casciato, Dr. Moran Pinter, Nils Persson, Tristan Kernick, Michael McBride, Matthew McDonald, Yi Sun, Chiamaka

Obianyor, and Youngjo Kim. It is a great pleasure to work with these talented people. I am also thankful to all the faculty members and all staffs at Georgia Tech.

I am also grateful to my colleges in Emory University: Dr. Jay Goodwin, Dr. Allisandra Mowles, Dr. Jillian Smith, Dr. Dibyendu Das, Dr. Savannah Johnson, Dr. Daniel Pierce, Dr. Chen Liang, Dr. Lisa Li, Noel Li, Fish (Ting) Pan, Rolando Rengifo, Michael Tseng and Yuka Watanabe. Special thanks to Hong Yi and Jeanette Taylor for helping with the microscopy measurement.

I also want to extend my gratitude to the members in CCE, especially Dr. Christine Conwell and Dr. Ramanarayanan Krishnamurthy. They had established a creative and helpful environment for me to start my research in Georgia Tech.

Other than the people I worked with, I also want to thank my classmates and friends for their company and advices: Guo Shiou Foo, Shilu Fu, Ho Yee Hui, Tang Zhenguan, Lu Xu, Luye He, Liren Xu, Boyi Fu, Chen Zhang, Xiaoling Zhan, Eric Parker, and Siyuan Zhang. I will treasure our friendship and wish them all the best in their bright future.

I also want to thank all my Taiwanese friends: Ping-Hsun Chu, Yu-Han Chu, Wei-Ming Yeh, Li-Chen Lee, Li-Wei Chou, Chi-Ta Lee and Ya-Dong Chiang. They generously provided me help and friendship when life became tough, especially when I was living for away from my home country.

I also want to thank all people I met in Dharma Jewel Monastery, especially the shifus: Jiansou shifu, Jianyuan shifu, and Jiantong shifu. They helped me to develop a peaceful mind, and provided me great opportunities to volunteer in the monastery, which may help more people to find their peaceful mind back.

I am very appreciative that I have a supportive and caring family, my parents and my sister. I would not be able to stay in a foreign country for years without the unconditional love and endless support from them. Special thanks for my grandpa. He wanted to witness my graduation from the collage. Although he missed that moment, I now could make him proud in heaven with this thesis.

Finally, I want to sincerely thank all people that I met so far. You helped me to become who I am right now. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation

A	Alanine
ACN	Acetonitrile
Aβ	β–amyloid
AFM	Atomic force microscopy
AIC	Akaike information criterion
AICc	Corrected Akaike information criterion
Boc	*tert*-Butyloxycarbonyl group
CD	Circular dichroism
CR	Congo Red
DCN	Dynamic combinatorial networks
E	Glutamic acid
F, Phe	Phenylalanine
FTIR	Fourier transform infrared spectroscopy
HCl	Hydrogen chloride
HPLC	High-performance liquid chromatography
K	Lysine
MS	Mass spectrometry
N, Asn	Asparagine
Orn	Ornithine
PBS	Phosphate-buffered saline
Q	Glutamine

R	Arginine
SEM	Scanning electron microscopy
SSE	Sum of square error
TEM	Transmission electron microscopy
ter-L	*tert*-Leucine
TFA	Trifluoroacetic acid
ThT	Thioflavin T
V	Valine
1SN	One-step nucleation
2SN	Two-step nucleation

Nomenclature

C	free peptide concentration (mM)
C_{1*}	peptide solubility for particles (mM)
C_{1*}	peptide solubility for assemblies (mM)
j	Average oligomer length
p_1	Prefactor for particle nucleation rate (L^{-1} sec^{-1})
p_2	Constant in the exponential part of the particle nucleation rate
ρ_{mon}	Density of peptide (g cm^{-3})
χ	Flory-Huggins parameter
W_{mon}	Molecular weight of peptides (g $mole^{-1}$)
γ	Interfacial tension between peptide and solvent (g s^{-2})
k_b	Boltzmann constant (J K^{-1})
T	Temperature (K)
N_A	Avogadro’s number ($mole^{-1}$)
k_{g0}	Kinetic constant for particle growth ($mole^{-1}$ sec^{-1})
k_n	Assembly nucleation constant (sec^{-1})
k_{AG}	Assembly growth constant inside the particle (sec^{-1})
$\rho_{\pi\alpha\rho}$	Particle density (g cm^{-3})
k_g	Assembly growth constant in the solution phase ($mole^{-1}$ sec^{-1})
k_{bp}	Prefactor of Assembly dissolution (sec^{-1})
ΔG	Binding energy between peptides in the assembly (J)
C_0	Initial peptide concentration (mM)
N	Sample size
k_1	Oligomerization rate constant ($mole^{-1}$ sec^{-1})

K_1	Oligomerization equilibrium constant
k_{forS}	Binding constant for S-methodol ($s^{-1}\ \mu M^{-1}$)
k_{forR}	Binding constant for R-methodol ($s^{-1}\ \mu M^{-1}$)
K_{dS}	Dissociation constant for S-methodol (μM)
K_{dR}	Dissociation constant for R-methodol (μM)
K_d	Dissociation constant for the aldehyde product (μM)
k_a	Binding constant for the aldehyde product ($s^{-1}\ \mu M^{-1}$)
k_{catS}	catalytic rate constant for S-methodol (s^{-1})
k_{catR}	catalytic rate constant for R-methodol (s^{-1})

SUMMARY

The peptide assembly mechanism is important for development of both functional biomaterials and clinical therapies. Although the assembly structures and assembling pathways have been studied for decades, the mechanism remains to be clarified, as the assembly pathway is a complex multistep process with various assembly morphologies and structures.

The main goal of this thesis is to investigate the assembly mechanism of functional peptide assemblies. The crucial building blocks capable of assembly are synthesized and selected via the dynamic combinatorial networks (DCNs), which yield sequence-defined oligomer with high fidelity. Both experimental analyses and mathematic modeling are applied to confirm the interplay between the chemical distribution and emergent physical transitions in the networks; model discrimination shows that the chemical nature of the DCNs is affected and shifted at different physical stage.

Next, to simulate the two-step nucleation process observed in the DCNs, a peptide assembly model is developed with two-step nucleation. The model simulates the phase transitions between different physical phases, given combinations of the Flory-Huggins parameter, the particle growth constant, and the binding energy to assemblies. The final physical phase distribution is determined by the solubility of the particle and assembly phases, where the phase with the lower solubility accumulates. To explore the potential of this model, experimental results from the peptide, Ac-KLVFFAE-NH_2, is rationally explained by this model. The model is then extended for a polydisperse system, where the

peptides undergo oligomerization, to simulate the chemical and physical transitions in the DCNs. The polydisperse model shows that the physical and chemical distribution may reach the steady state independently.

The assembly pathway with two-step nucleation is experimentally investigated with the pH sensitive Aβ(16-22) peptide, which assembles into fiber at neutral pH but tubes at acidic pH. The difference between the nucleation and propagation environments for the two-step nucleation mechanism affects the assembly kinetics and morphological selection. Although Aβ(16-22) assembles into ribbons pH-independently, the ribbon intermediates undergo pH-dependent reaction pathway and transition into different morphologies under different pH conditions.

Finally, the catalytic peptide assemblies are analyzed. The methodol substrate is cleaved by Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 nanotubes enantioselectively. The Experimental results are analyzed with a modified Michaelis-Menten mechanism to resolve the nature of this catalytic system, including the number of peptides per binding site, the enantioselectivity, and the stability of the tubes. The simulations suggest similar binding pocket size for the both nanotubes, and the assemblies are stable throughout the entire reaction time. The enantioselectivity for the R- and S-methodol substrate may come from the chemical selectivity of the peptide nanotubes, but not from the difference of the binding affinity, based on the analysis.

The results from this thesis provide a comprehensive investigation of the construction of functional peptide assemblies. The building block synthesis and selection, the assembly structure and mechanism, and the functional properties of the peptide assemblies are reported and discussed.

CHAPTER 1. INTRODUCTION

1.1 Amyloid Peptide Assembly

Amyloid structures are highly ordered aggregates formed by assembled peptides or proteins, and they have triggered intense interest as they cause the neurodegenerative diseases such as Alzheimer's and Parkinson's diseases [1-5]. The design of new therapies against these diseases depends on the understanding of the structures and the aggregation mechanism of both the amyloid intermediates [3, 6-9] and final products [10, 11]. Other than clinical therapies, the amyloid assemblies have also been applied for the construction of functional nanomaterials, which may be used for drug delivery [12-15], nanofabrication [16, 17], and biocatalysts [18]. In the following sections, the amyloid-related diseases, the functional peptide-assemblies, and the techniques characterizing and quantifying the assemblies will be discussed.

1.1.1 Peptide Amyloid-related Diseases

The assembly and aggregation of peptides and proteins have caused many fatal diseases such as neurodegenerative diseases, type 2 diabetes, and prion diseases. For the neurodegenerative disease, such as Alzheimer's disease, the patients may undergo progressive dementia or decline in memory and cognitive functions. One of the peptides causing these diseases is the β-amyloid (Aβ) peptide, which is the cleavage product of the amyloid precursor protein. The accumulation of extracellular Aβ plaques is found in patient's brain, and the deposit of the plaques is necessary but not sufficient to cause the disease [19]. There are other amyloid assemblies of peptides/proteins neuropathologically related to diseases. The aggregation of α-synuclein [20] is related to Parkinson's disease, which cause movement disorder and gait disturbance. To efficiently develop the clinical therapies for these diseases, it is important to understand the assembly mechanism of the peptides and proteins.

1.1.2 Functional Peptide Self-assemblies

Peptide assembly has become a powerful methodology to construct functional nanoscale materials. The morphology and function of the peptide assemblies may be manipulated at the molecular level by varying the sequence of the peptides. Through sequence modification, the structure-function relationships between different peptides may be studied systematically, which accelerates the development of functional peptide-based architectures. Also, some peptide assemblies are sensitive to external stimuli such as pH [21], temperature [21], and light [22], and these stimuli may be applied to diversify the structures and thus the functionalities of the peptide assemblies. In the following paragraphs, several functional peptide assembly-based materials are discussed.

Excellent catalytic activity [23] and selectivity [24] of protein enzymes have inspired scientists to synthesize catalysts with amino acids as the basic building blocks. The catalytic efficiency of protein enzymes relies on their structures, which are essential for substrate recognition, size of the binding pocket, and the catalytic pathway. However, the complex protein folding mechanism has become a problem for artificial enzyme development, as the enzymes need to fold properly for catalytic functions [25, 26]. On the other hand, without the needs to solve the complex protein folding mechanism, peptide assembly has shown promising potential as biocatalysts with a bottom-up fabrication mechanism. As one of the first examples for catalytic peptide assemblies, Guler and coworkers assembled peptide amphiphile nanostructures able to break ester bonds [27]. Peptides with histidine residues as the catalytically reactive group were designed, and the peptide reacted similar to a Michaelis-Menten enzyme after they assembled into nanotubes. Other peptide-based assemblies have also shown catalytic ability for aldol reactions [28], hydrolysis [29], and Diels-Alder reaction with metal ions as the cofactors [30].

Other than the chemical function, peptide assemblies have been used as templates for nanofabrication. The peptide assemblies with well-defined shapes and sizes show remarkable capability to synthesize inorganic functional nanomaterials [17]. Silver nanowires have been successfully fabricated with reduction of ionic silver inside the peptide nanotubes, which are degraded with enzymes [16]. The amine groups at the assembly surface may also serve as nucleation sites to fabricate metal nanoparticles [31, 32].

Compared to other organic materials, peptide assemblies are more suitable for drug delivery due to their biocompatibility, biodegradability, and multifunctionality [14, 15]. Both hydrophobic and hydrophilic drugs may be carried with peptide assemblies as the vesicles [33-35], and the adjustable assembly structures given the tailorable sequences also make peptide assemblies efficient carriers to different target sites [36-38]. Other advantages, including high efficiency for drug loading, low ratio for drug lose, and high stability [39], also makes the nanostructures of assembled peptides excellent drug carriers over the other candidates at the nanoscale.

1.1.3 Morphologies and Structures of Peptide Assemblies

The tailorable sequences of peptides make it possible to design and explore new structures and morphologies of the peptide assemblies. With different amino acid sequences, peptides assemble into diversified morphologies, such as tubes [40], sheets [41], vesicles [42], particles [21], and ribbons [43]. Common interactions driving peptide assembly are typically noncovalent, including hydrophobic interactions [44], electrostatic attraction [45], van der Waals interactions [46], π–π stacking [47], and hydrogen bonding [48]. The patterns of hydrogen bonding determine the secondary structure of the assembly; the assemblies possess α-helix structure with intramolecular hydrogen bonds or β-sheet structure with intermolecular hydrogen bonds. All of these morphologies and structures may be probed with appropriate instruments.

1.1.4 Measuring the Kinetics and Structural Change of Peptide Assemblies

Characterization and quantification of the peptide assemblies can be done with spectroscopic methods investigating the chemical, physical, and electrical properties of the assemblies. For example, peptide assemblies interact with clockwise- and counterclockwise-spinning light differently due to the chirality of the assemblies, which is affected by the secondary structure of the assemblies. The chirality generates a non-zero circular dichroism (CD) signal [49], and thus CD spectroscopy provides reliable information to determine of the secondary structures of peptide assemblies [50] and to quantify the assembled peptides [51].

To probe the complex pathways and registries of the peptide assemblies with sufficient time resolution, FTIR has become a useful technique to address the peptide assembly kinetics [45, 52]. FTIR is sensitive to the secondary structure of amyloid aggregation. The amide I bond from the antiparallel β-sheet structure shows one band around 1625 cm^{-1} and another around 1690 cm^{-1} on the IR spectrum. For parallel β-sheet structure, the amide I bond shows only one band around 1635 cm^{-1} [53-57]. More detailed structural information can be obtained with isotope-edited peptides, which provide structural information at the amino acid level [52, 58, 59]. The assemblies of ^{13}C-enriched peptides show IR spectra with split amide I bands; the ^{13}C band is red shifted to smaller wavenumbers and is separated from the ^{12}C band as a result. The separated bands allow specific determination of the coupling registries and the structural transitions of the assemblies [45, 60]. Also, the small mass requirement and rapid data acquisition make FTIR useful for measuring the kinetics of amyloid formation and maturation [52, 61].

The fluorescence dye is another tool related to peptide assembly investigation. Fluorescence dyes are used to quantify the concentration of assembled peptides or proteins with β-sheet amyloid structures. Congo Red (CR) shows a distinct apple-green birefringence upon binding to amyloid structure and thus serves as an indicator of amyloid assemblies. Thioflavin T (ThT) is another quantitative probe for peptide assemblies with cross β-sheet structure, which fluoresces on binding. The high sensitivity, simplicity, and fast measurement make ThT an universal diagnostic for peptide assemblies [62].

Finally, microscopy has been used to observe the morphology of the assemblies. Multiple imaging microscopies have been used to visualize the amyloid nanostructures, such as transmission electron microscopy (TEM), atomic force microscopy (AFM), scanning electron microscopy (SEM), and fluorescence microscopy [63], which provide the morphological information with high resolution.

1.1.5 Modeling the Mechanism and Kinetics of Peptide Self-assembly

The range of peptide and protein sequences involved in amyloid assembly has complicated defining the assembly pathways and modeling, particularly the critical nucleation step [64]. One-step nucleation (1SN) [65-68] is the most direct model for assembly nucleation, which assumes that the free peptide monomers nucleate into ordered structures in solution. However, strong desolvation energy barriers are known to prevent assemblies from nucleating directly in solution [69, 70]. Instead, according to Ostwald's rule of stages [71], the peptides may form a metastable phase prior to nucleation of fibers [72]. This two-step nucleation (2SN) has been observed widely not only in peptide amyloid assemblies [73-77], but also in the crystallization of proteins [70, 78-83], colloids [84-86], minerals [87, 88] and polymers [89]. The morphologies of the metastable phases range from molecular clusters to macroscopic dense solution phases [90], which may either a) be incorporated as the assembly phases develop [21, 69, 79], b) coexist with the assembly phases [21, 78, 91], or c) dominate if assembly nucleation is slow or thermodynamically disfavored [11, 21, 72, 92].

Both 1SN and 2SN have been experimentally observed [75] and mathematically analyzed for peptide assembly. Several models describe 1SN nucleation occurring in solution directly [65-68, 93, 94], while 2SN requires an intermediate metastable phase to initiate the assembly phase. Pan and coworkers have simulated two-step nucleation of proteins with a phenomenological kinetic model as a function of temperature [95]. The reversible formation of the metastable phase and irreversible nucleation of the assembly phase are assumed to be Arrhenius-type. They conclude that the rate-determining step of the assembly phase is assembly nucleation, which is controlled by both viscosity and monomer concentration in the metastable phase. Since the amount of the metastable phase significantly impacts the maturation of the assembly phase [96], Kashchiev and coworkers simulate the assembly nucleation rate as a function of the individual metastable particle size [97]. Later Auer *et al.* compare the difference between 1SN and 2SN [72] using classical nucleation theory to simulate the formation of the metastable phases. However, all these efforts focus solely on assembly nucleation rate without including assembly maturation or monomer depletion. To explain protein crystallization data, Sauter and coworkers in 2015 proposed a 2SN kinetic model [90] that included the formation of a metastable phase, monomer depletion, and transition to the assembly phase. However, the model used the total peptide concentration to describe the metastable phase because of its poorly defined morphology. For peptide assembly, the peptide particle is a common metastable phase, and the individual size of the metastable particles or droplets are critically tied to nucleation rate with assemblies; the assemblies inside a particular particle do not extend into the solution phase until they become longer than that particular particle size.

After the assembly nucleated and is exposed into the solution, they propagate through consuming the other free peptides. In some models the assembly growth is assumed to be irreversible due to the stable assembly structures. Finke and Watzky proposed a pseudoelementary model [66] to describe the kinetics of peptide assembly. With a first order reaction for assembly nucleation and a second order reaction for assembly autocatalytic growth, the model is able to describe the aggregation of several proteins and peptides, including α-synuclein, amyloid β peptide, and polyglutamine. However, when the assembly dissolution is significant and thus the assembly growth is reversible, the peptides may dissociate from the end of the assemblies [68]. In addition, when mechanical forces are applied to the peptide solution by stirring or shacking, the assemblies will break in the middle and thus the number of active ends increase [67, 98], which accelerates the autocatalytic assembly kinetics [65, 68].

1.2 Dynamic Combinatorial Networks

In addition to assembling the peptides through non-covalent bonding, researchers have applied covalently interacting molecules to build a system with emergent functionality. In dynamic combinatorial networks (DCNs), the chemical complexity of the networks is achieved with reversible covalent linkages between different building blocks, and the composition of the networks is thermodynamically determined [99, 100]. Species, including the assemblies, in these complex networks are sensitive to the surroundings; they respond to external stimuli, and thus may be manipulated for different functions. This multicomponent assembly network may also shed a light on the origin of life, due to its environmental sensitivity and accumulating complexity [98].

1.2.1 Reversible Linkages in Dynamic Combinatorial Networks

The properties of the DCN may be determined with different reversible linkages and the building blocks. The dynamic combinatorial chemistry is different from the traditional combinatorial chemistry because of its reversible linkages [101], through which the distribution between different network members can change. Many reversible linkages have been used to construct dynamic combinatorial networks, including the imine linkages [102] and the acetal linkages [103, 104].

The first imine-based dynamic combinatorial network was constructed to polymerize DNA [102], where the trimers reversibly make hexamers. Later the imine linkages have been applied to mediate the networks of aldehydes, amines, and macrocyclic molecules [105-108]. However, due to its high vulnerability to hydrolysis, these imine-based species are difficult to analyze with HPLC in aqueous solution, which is a common tool to separate and quantify the species in the DCNs. One solution for this problem is to incubate the network in water-free conditions [109].

For reversible acetal linkages, acid-catalyzed cyclic products may be produced with under thermodynamic control. The condensation of glycerol with isobutyraldehyde results in both five-membered ring and six-membered ring compounds (Figure 1.1) [110]. In this network, the five-membered ring product is kinetically favored, while the six-membered ring product is more thermodynamically stable. The reversible nature of the acetal linkages gradually shifts the five-membered rings into the six-membered rings, and the later predominate in the end. Acetal linkages have also been applied to construct DCNs of diaminobutanediols [103, 111], benzene hexathiol with p-tolualdehyde [112], and pyrazolotriazone with aldehydes [113].

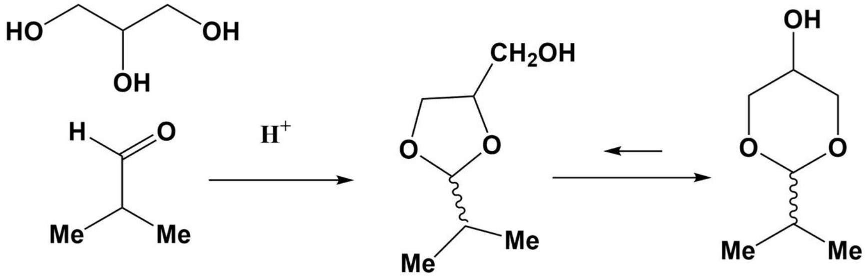

Figure 1.1 – The condensation of glycerol with isobutyraldehyde under acidic conditions [110].

1.2.2 Self-templating and Assembly in Dynamic Combinatorial Networks

Coupling peptide assembly and chemical transformation (peptide oligomerization) in DCNs opens a new door for materials with new structures and functions. In biological systems, replication of biopolymers is usually template-directed. This inspired researchers to amplify the members in the networks with assemblies as the templates; the members bind to the template to copy themselves, and the equilibrium is driven by the affinity between the templates and the members [114].

For peptide assemblies, the ends of the β-sheet amyloid structure may be used to template the formation of the assembly building blocks. Ashkenasy and coworkers ligate the shorter peptide pieces into the full-length building blocks, which then assemble with their templates and elongate the assemblies (Figure 1.2) [115]. As shown in Figure 1.2, the precursors E and N are non-covalently bound to the end of the β-sheet structure for ligation. The ligated E and N become the building block **1** and assemble into the β-sheet structure. Later, Otto and coworkers report self-templating peptide-based macrocycles emerging from a DCN [98], whose hexamers and heptamers assemble into fibers. Their fiber assemblies possess β-sheet structure, and are sensitive to external stimuli such as shearing force. When mechanical force is applied, formation of hexamer and heptamer assemblies is accelerated, which consume the other members in the network through the templating effect.

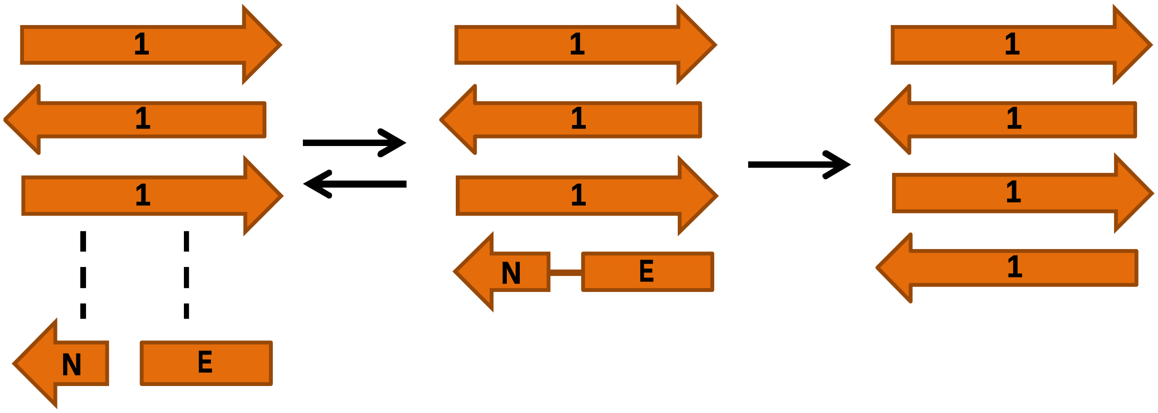

Figure 1.2 – Formation of the building block in a peptide assembly system via self-templating [115]. After the precursors, N and E, are bound to the end of the assembly of molecule 1, they are templated into molecule 1 with a covalent bond.

1.2.3 Functional Dynamic Combinatorial Networks

The reversible and flexible distribution of DCNs makes the networks an attractive tool to find ligands for biomolecular and synthetic receptors for drug discovery [116, 117]. The external targets serve as a template and amplify the desired members in the network. The interaction between the desired members and the targets becomes a "sink," which drives the equilibrium to the desired member [118]. This strategy makes DCN effective in screening not only useful ligands and receptors for different targets, but also catalysts for different reactions. The efficiency of a catalyst depends on its selectivity for the substrate and chemical activity to catalyze the reaction. Several approaches have been made to screen the most suitable catalytic molecules from the candidate pool, based on their selectivity to the target substrates [119, 120]. Catalytic members have been screened in this way for reactions like the Diels—Alder reaction [121] and the acetal hydrolysis reaction [122].

1.3 Objectives

The main goal of this thesis is to study the kinetic of peptide assembly with emergent functions (Figure 1.3). First of all, peptide oligomers capable of assembly are covalently synthesized via DCNs (the 1st dimension in Figure 1.3). Next, the kinetics of peptide assembly are mathematically and experimentally studied to understand the reaction pathway and the resulting morphologies (the 2nd dimension in Figure 1.3). Finally, the emergent catalytic function of the peptide assemblies is analyzed (the 3rd dimension in Figure 1.3). With the following chapters, the construction and the properties of the functional peptide assemblies are demonstrated.

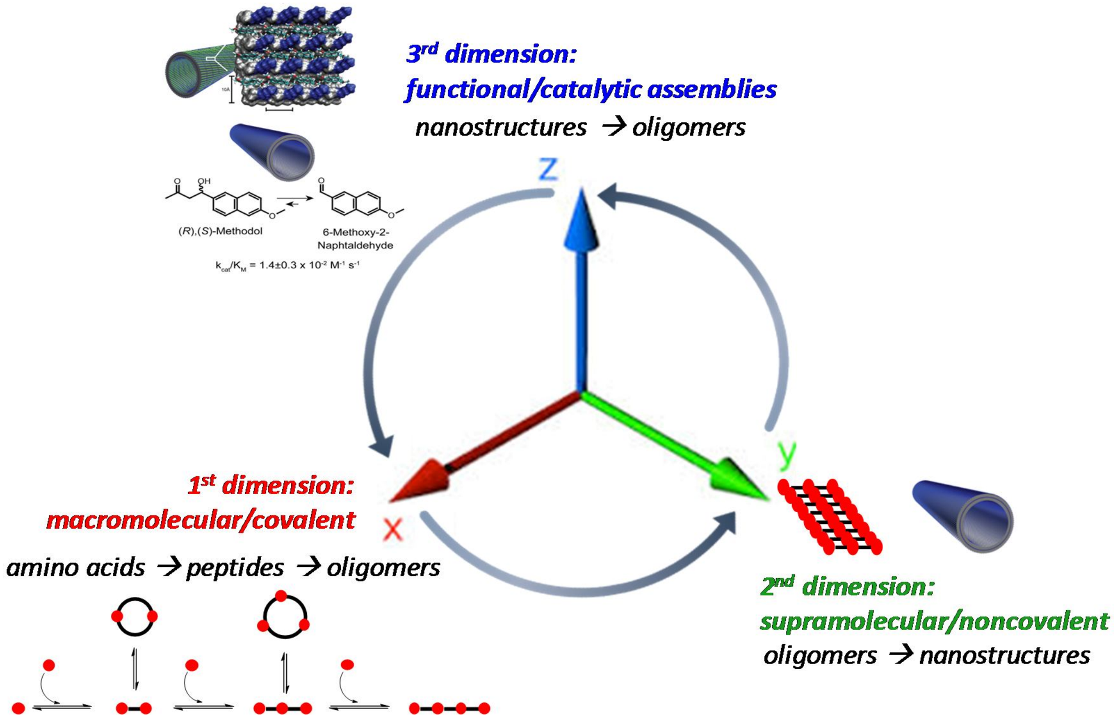

Figure 1.3 — The construction of peptide assembly with emergent catalytic function.

In chapter 2, the kinetics of DCNs are experimentally analyzed and mathematically modeled, which are constructed with chemically modified NF and NFF peptides. The network distribution is affected by both the chemical property of the building blocks and the emergent physical phases. The observed assemblies help to template the formation of peptide oligomers, which yields sequence-defined oligomers with high fidelity.

In chapter 3, a model is constructed to describe the peptide assembly process with two-step nucleation. This model simulates the formation of peptide fibers via the metastable particles. First, a monodisperse model addresses that both the solubility of peptide particle and assembly determines the final state of the system, while the process may be kinetically limited by slow particle growth/dissolution. This monodisperse model is supported with the assembly kinetics of a model peptide, Aβ(16-22), which are rationally explained by the model prediction. Finally, to simulate the transitions in the DCNs, the model is modified for a polydisperse system, and it shows that the steady state of the physical phase distribution and the chemical oligomer concentration distribution may be independent on each other.

In chapter 4, the pH-sensitive Aβ(16-22) assembly is further examined. Aβ(16-22) assembly follows the two-step nucleation, and the fibers nucleate in the particles while grow in the solution. This environmental difference is important for the assembly stability and morphological selection. Here Aβ(16-22) is incubated at both neutral and acidic pH to probe the structural evolution and the possible different reaction pathway.

In chapter 5, the catalytic function of the peptide assemblies is tested. The methodol substrate undergoes retro-aldol reaction with the peptide assemblies, and the experimental results are mathematically analyzed to clarify the reaction mechanism, size of the binding pocket, and the substrate selectivity.

Chapter 6 contains the summary of this work and the recommendation for future studies on the kinetics and structural evolution of peptide assembly.

CHAPTER 2.

KINETICS AND PHASE TRANSITIONS OF THE DYNAMIC COMBINATORIAL NETWORKS

2.1 Introduction

The natural sequence-defined biopolymers, such as enzymes or antibodies, have shown remarkable functions. Template-assisted polymerization is common and fundamental in natural biological systems to synthesize sequence-defined biopolymers, which usually requires catalytic enzymes to drive the reactions, such as DNA replication or protein translation. The necessity of catalytic enzymes complicates those polymer syntheses due to the complex interactions between the substrates and the enzymes. Hence, researchers seek for simpler strategies to synthesize and engineer the sequence-defined polymers for functional materials [123]. However, in several approaches the chain length distribution [124] and the sequence fidelity [125] of the polymers are hard to control. Here, the peptide-based dynamic combinatorial networks (DCNs) are constructed and reported, which show the ability to synthesize monodisperse biopolymers with high fidelity without enzymes. Two networks with reversible acetal linkages are discussed here: *NF*-CHO and *NFF*-CHO networks. Both experimental techniques and mathematical modeling are combined to reveal the chemical and physical properties of the networks. The design of DCNs provides a simpler pathway to generate monodisperse peptide-based polymers, and opens a new way for functional biopolymer synthesis.

2.2 Results and Discussion

2.2.1 Kinetics of the NF-CHO Network

A DCN is generated with the *NF*-CHO monomer. The NF dipeptide is synthesized with standard Boc-peptide synthesis, with the C-terminus reduced to an aldehyde group. The Boc group (*tert*-Butyloxycarbonyl group) on this *NF*-CHO monomer is removed with trifluoroacetic acid (TFA) to activate the monomer, which then may perform polymerization and generate the network via the reversible acetal linkages (Figure 2.1).

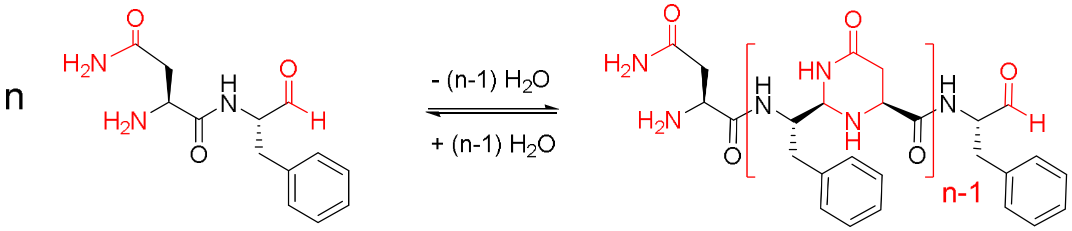

Figure 2.1 – Polymerization of the *NF*-CHO monomers. Image adapted from Ref. [126] with permission.

The *NF*-CHO network is incubated in 40% acetonitrile (ACN) aqueous solution with an initial monomer concentration of 8 mM, and the network species are identified by mass spectrometry (MS) after HPLC separation. Figure 2.2 shows the HPLC analysis of the *NF*-CHO network after two days of incubation. The mobile phase of HPLC (gradient ACN/water) stabilizes the species with quasi-neutral pH, under which the hydrolysis of the acetal linkages is limited [127]. The HPLC peaks have good Gaussian shapes, which suggest that the network species are stable during the HPLC separation without significant hydrolysis. As shown in Figure 2.2, there are five major products in the network identified: monomer, linear dimer, cyclic dimer, linear trimer, and cyclic trimer.

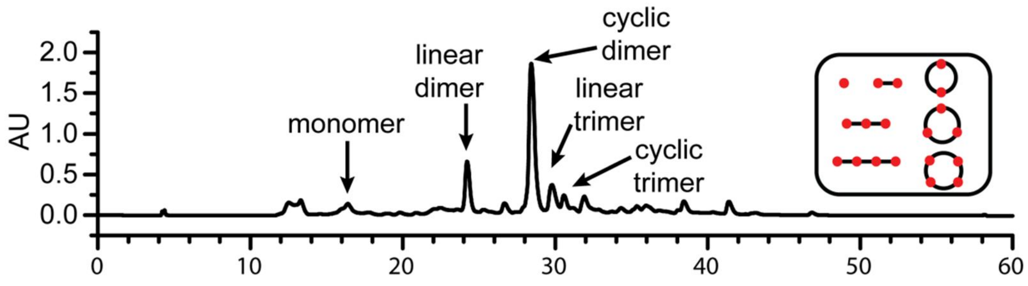

Figure 2.2 – The HPLC spectrum of the *NF*-CHO network after 2 days of incubation. Image adapted from Ref. [126] with permission.

To verify the reaction pathway, a model is constructed to describe the evolution of the network species with time. HPLC and MS sampling provide time-dependent changes over the initial 48 hr of *NF*-CHO network maturation, based on which a kinetic model (Figure 2.3) is constructed. M, D, D_c, T, T_c stand for monomer, linear dimer, cyclic dimer, linear trimer and cyclic trimer, respectively, in Figure 2.3. The model describes the polymerization (forward reaction) with forward rate constants (k_1, k_3, k_5 and k_7), while the backward rate constants are calculated based on the ratio of the forward rate constants and the corresponding equilibrium constants (K_2, K_4, K_6, and K_8).

$$M + M \xrightleftharpoons[k_1/K_2]{k_1} D$$

$$D \xrightleftharpoons[k_3/K_4]{k_3} D_c$$

$$M + D \xrightleftharpoons[k_5/K_6]{k_5} T$$

$$T \xrightleftharpoons[k_7/K_8]{k_7} T_c$$

Figure 2.3 – Kinetic steps for the *NF*-CHO network in Figure 2.2. M, D, D_C, T and T_C represent monomer, linear dimer, cyclic dimer, linear trimer and cyclic trimer, respectively. k_1, k_3, k_5, and k_7 are the forward rate constants and K_2, K_4, K_6, and K_8 are the corresponding equilibrium constants.

The fit of the kinetic model in Figure 2.3 to the *NF*-CHO network concentrations (Figure 2.4a) follows the consumption of monomer up to 12 hours. The calculated curve then underestimates the monomer concentration at 19 hours but overestimates the 24-hour data point. There is also a distinct concentration change for the linear dimer and trimer between 19 and 24 hours as highlighted in Figure 2.4b, which is not captured by the calculated curves. The deviation suggests that a physcial phase transition happens as the network matures, which changes the equilibrium between the species. This physical transition is probed with transmission electron microscope (TEM); as shown in Figure 2.5, the particle phase emerges after 17 hr of incubation, and it is correlated with the deviation between the calculated and experimental data.

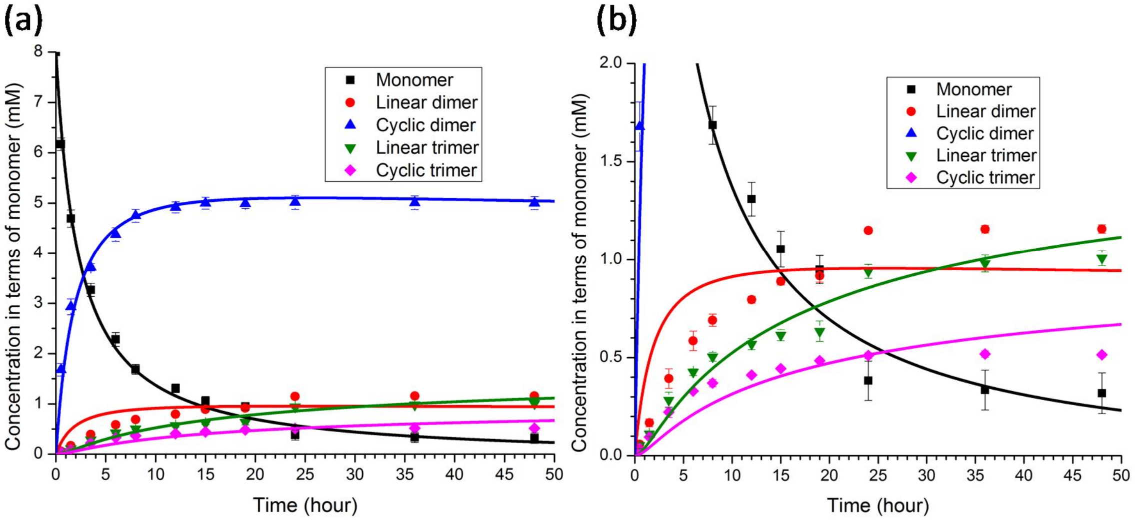

Figure 2.4 – Fit of the *NF*-CHO network concentrations to the model in Figure 2.3. (a) Solid lines are fit using Figure 2.3 to the data. (b) Expansion of the y-axis in the left panel highlights a distinct transition between 19 hr and 24 hr.

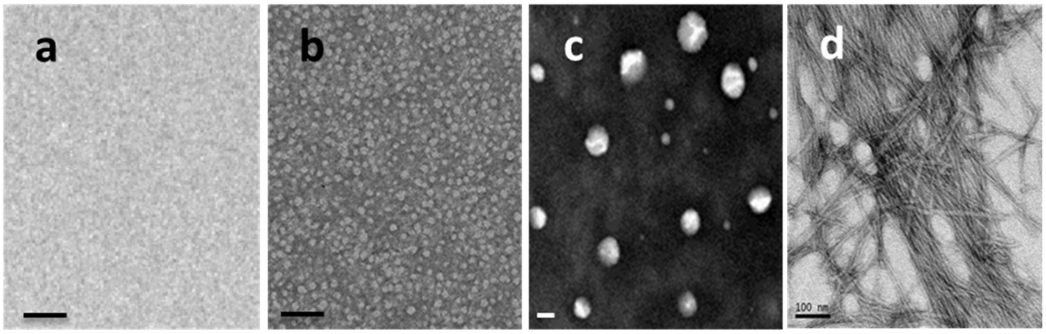

Figure 2.5 – TEM images of the NF-CHO network after (a) 4 hr, (b) 17 hr, (c) 48 hr, and (d) 96 hr of incubation. Image adapted from Ref. [126] with permission. Scale bar = 100 nm.

The distinct changes in the physical state of the network in Figure 2.4 suggest that a single set of parameters is not sufficient to adequately describe the maturation of the network. Thus, the model is expanded into a two-stage model with a boundary at 21.5 hours, halfway between the data points at 19 and 24 hours. This boundary separates the solution stage from the particle stage, which share common reactions but individually have unique rate and equilibrium constants. This two-stage model better describes the evolution of the *NF*-CHO DCN and covers the dimer and trimer formation with the emergent particle stage. Although both intramolecular (cyclization) and intermolecular (oligomerization) reactions occur actively in the solution stage, cyclization appears inaccessible in the particle stage. Amyloid peptide particles are desolvated and favor condensation, which should promote further polymerization between the species. However, the constant concentration of the cyclic dimers in the particle stage suggests some other particle feature limiting cyclization.

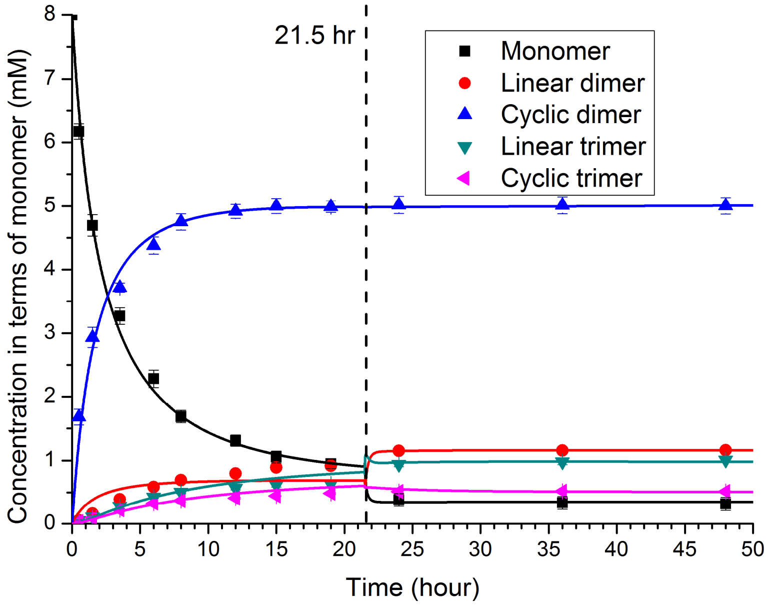

Figure 2.6 – Fit of the *NF*-CHO concentrations to a two-stage model. The boundary at 21.5 hours defines the transition between the solution stage and the particle stage.

To more rigorously evaluate these models and the potential of over-fitting with the increased number of parameters, the corrected Akaike information criterion [128] and the Akaike weight [129] are applied to evaluate the consistency between the models and the experimental data. These criteria provide a measurement of the goodness of fits from the models by penalizing the models with more parameters and awarding the models with less sum of squares error. The corrected Akaike information criterion ($\mathrm{AIC_c}$) is a modified version of the Akaike information criterion (AIC) for small sample sizes, and is calculated according to:

$$\mathrm{AIC_c} = N_d \ln\left(\frac{SSE}{N_d}\right) + 2N_p + \frac{2N_p\left(N_p+1\right)}{N_d - N_p - 1} \tag{2.1}$$

where N_d is the number of data points, N_p is the number of parameters, and *SSE* is the sum of squares error. The Akaike weight reflects the probability that a model describes the data consistently based on AIC$_c$:

$$\text{Akaike weight} = \frac{\exp\left(-\frac{1}{2}\Delta_i\left(\text{AIC}_\text{c}\right)\right)}{\sum_{j=1}^{N_M}\exp\left(-\frac{1}{2}\Delta_j\left(\text{AIC}_\text{c}\right)\right)} \tag{2.2}$$

where N_M is the total number of models for comparison and $\Delta_i\left(\text{AIC}_\text{c}\right)$ is the difference between the AIC$_\text{c}$ of model *i* and the minimum AIC$_\text{c}$ from the model set. As shown in Table 2.1, the single-stage model's *SSE* is larger than the *SSE* for the two-stage model; the Akaike weight of 97% compared to 3% indicates that two-stage model is the more likely description of the experimental data. Hence, the two-stage model is more consistent with the physical phase change and the chemical step change in member concentrations correlated in time.

Table 2.1 – AIC$_\text{c}$ analysis for the fits from the model in Figure 2.3 to the *NF*-CHO network concentrations with single- and two-stage models. N_d is the number of data points, and N_p is the number of parameters.

Model	N_d	SSE	N_p	AIC$_c$	Akaike weight (%)
Single-stage	60	1.41	8	-207	2.94
Two-stage	60	0.81	16	-214	97.1

The evolution of the *NF*-CHO network does not stop in the particle phase. Although the chemical exchange remains in equilibrium between 21-48 hr (Figure 2.6), the particles continue to grow in size by either particle fusion or Ostwald ripening [130, 131]. For the later, large particles grow by consuming the smaller ones given the driving force to lower the total surface area and to minimize the energy in the system. Also, the TEM images show that the particle phase is metastable and transitions into a fiber-rich phase (Figure 2.5c, d). The emergence of the fibers ends the particle stage and shifts the equilibrium between the species again. In this third stage, the fiber stage, the concentration of trimer increases significantly (Figure 2.7), which implies that the linear trimers are the component of the fibers. To identify the composition of the fibers, the fibers are spun down at 16,000 × g for 30 minutes, and then re-suspended in 40% ACN after the supernatant is removed. As the particles are not spun down, HPLC identifies the re-suspended portion as linear trimer assemblies (Figure 2.8), whose peak has the same retention time as in Figure 2.2.

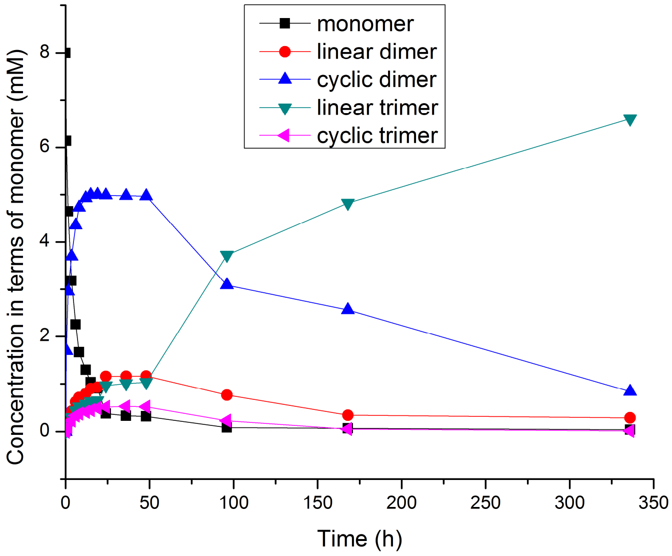

Figure 2.7 – Kinetics of the *NF*-CHO DCN in two weeks. Image adapted from Ref. [126] with permission.

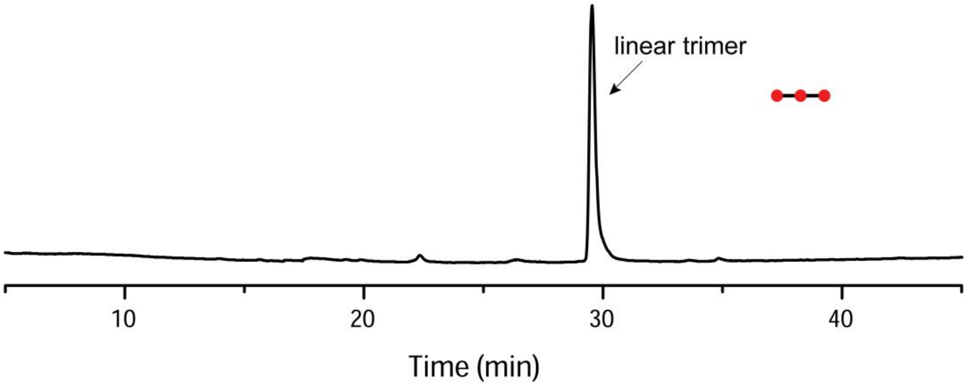

Figure 2.8 – HPLC analysis of the fibers in *NF*-CHO network. Image adapted from Ref. [126] with permission.

2.2.2 Kinetics of the NFF-CHO Network

Similar to the *NF*-CHO monomer, the *NFF*-CHO monomer are synthesized with standard Boc-peptide synthesis followed by reduction with $LiAlH_4$, which then gives the *NFF*-CHO monomer. After the Boc protection group is removed with TFA, the activated *NFF*-CHO monomer is incubated in 40% ACN aqueous solution to generate the network. The species in the network are separated with HPLC under quasi-neutral pH and identified with mass spectrometry. The incubated *NFF*-CHO network is composed of three main species: monomer, linear dimer and linear trimer, as shown in Figure 2.9.

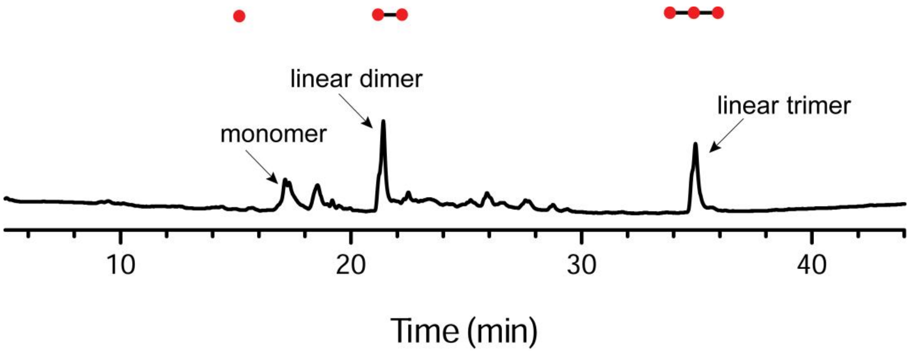

Figure 2.9 – HPLC analysis of the *NFF*-CHO network after two days of incubation. Image adapted from Ref. [126] with permission.

The evolution of the *NFF*-CHO network is followed by HPLC and TEM with time, and a model is constructed based on the observed chemical and physical properties. The emergence of the particle and assembly phases in the *NFF*-CHO network is observed by TEM (Figure 2.10), and the assemblies are characterized as the trimer assemblies (Figure 2.11). The species distribution is followed by HPLC up to 336 hr, the end of the fiber stage. A kinetic model is developed to clarify the reaction mechanism in all three stages (Figure 2.12). The formation of linear dimer and linear trimer is formulated with bond formation (forward) rate constant (k_1 and k_3), while the backward constants are calculated based on the forward constants and the equilibrium constant (K_2 and K_4). For the linear trimer assemblies, a minimal autocatalytic assembly mechanism [66] is applied to the linear trimers. The assembly mechanism contains two steps: the assembly nucleation (k_5) and the autocatalytic growth (k_6). Combined together, the reaction mechanism for the *NFF*-CHO network is summarized in Figure 2.12.

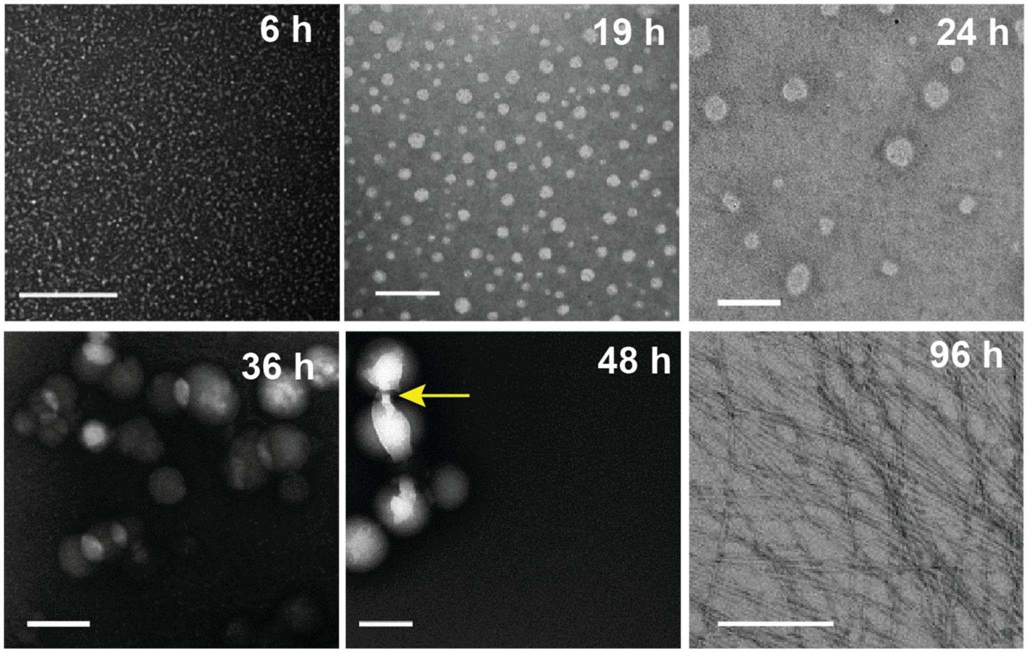

Figure 2.10 – TEM images of the *NFF*-CHO network. Image adapted from Ref. [132] with permission. Scale bar = 200 nm.

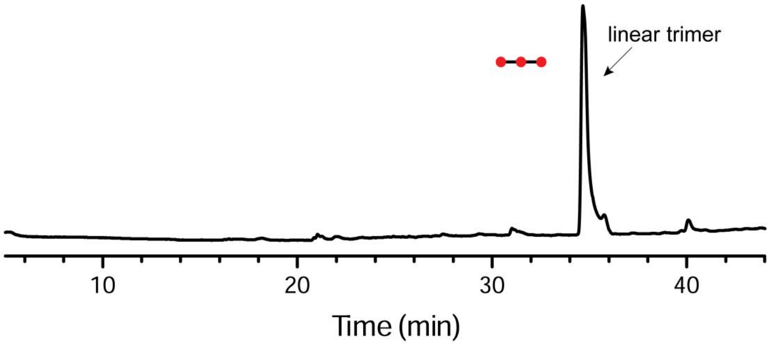

Figure 2.11 – HPLC analysis of the fibers in *NFF*-CHO network [126]. The assemblies are spun down under 16,000 × g and re-suspended after the supernatant is removed. The retention time of the re-suspended assemblies is the same as the linear trimers' in Figure 2.9.

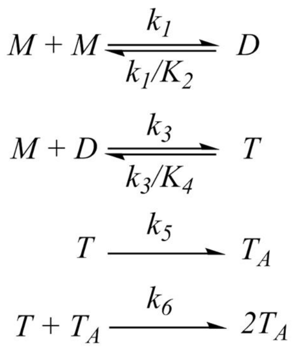

Figure 2.12 – Kinetic steps for the *NFF*-CHO network. *M, D* and *T* represent monomer, linear dimer and linear trimer, respectively. k_1 and k_3 are forward rate constants while K_2 and K_4 are the corresponding equilibrium constants. T_A is an assembled trimer in the fiber phase. $T + T_A$ represents the autocatalysis of unassembled trimer into fibers by pre-formed fibers. k_5 and k_6 are the rate constants for fiber nucleation and assembly.

The reaction mechanism fits the kinetics of the *NFF*-CHO network, which shows stage-wise evolution (Figure 2.13). In the first solution stage, the monomer concentration drops quickly within the first eight hours while the concentration of the linear dimer and linear trimer increases significantly (Figure 2.14). The fast condensation between the building blocks drives the oligomerization quickly, and the accumulation of longer oligomers triggers the formation of the particles, which causes distinct concentration changes. The phase boundary for model fits is assigned as halfway between the data points where distinct changes in concentrations are observed, and by this analysis the particle stage boundary is at 9 hr. This transition is significantly earlier than in the *NF*-CHO network (21.5 hr as shown in Figure 2.6). The relatively fast formation of particles is consistent with the increased hydrophobicity from the additional phenylalanine of the *NFF*-CHO building block compared to *NF*-CHO.

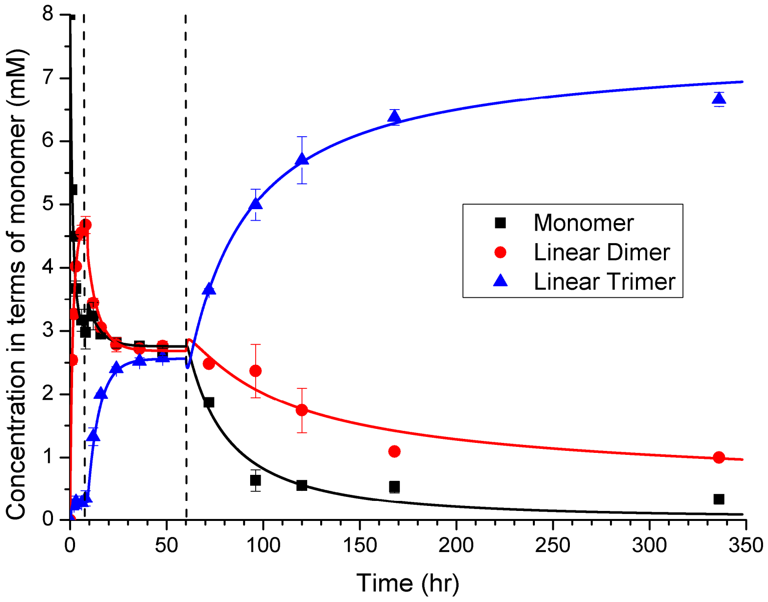

Figure 2.13 – Fits of the model in Figure 2.12 for the *NFF*-CHO network. Three distinct stages of the network are observed over the 350 hrs.

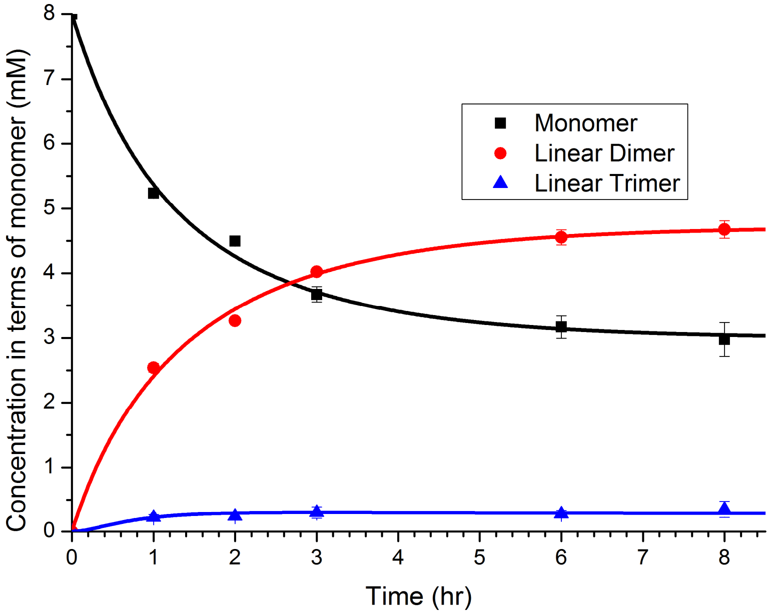

Figure 2.14 – Fits of the model in Figure 2.12 for the *NFF*-CHO network in the solution stage.

The particle stage is the second stage of the *NFF*-CHO network evolution. In the particle stage, the linear trimer accumulates significantly, after which the network member distribution remains remarkably constant for a day (Figure 2.15). Although the ratios of the species concentration do not change, the size of the particles in the network increases between 19 – 48 hr (Figure 2.10). Similar to the *NF*-CHO network, the particles may grow via Ostwald ripening or particle fusion, and later the fibers emerge from the particles (Figure 2.10, 48 and 96 hr).

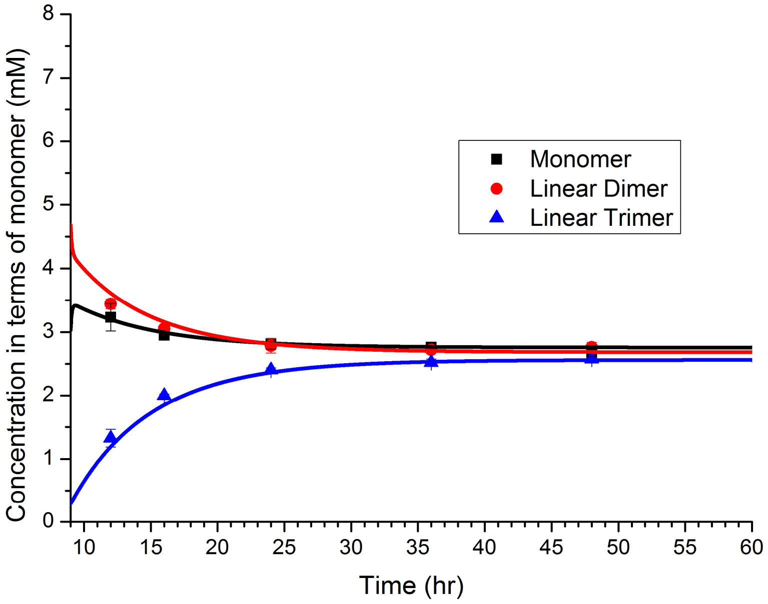

Figure 2.15 – Fits of the model in Figure 2.12 for the *NFF*-CHO network in the particle stage.

At the end of the particle stage, the nucleation of the ordered trimer assembly is triggered inside the particles. Similar to Ostwald's rule of stages [71] and the two-step nucleation [97], the trimers are organized into the ordered fiber structure from a metastable particle phase [72]. The trimer assemblies inside the particles are exposed to the solution phase after they grow long enough by consuming the other molecules in the particles. The exposed trimer assemblies then take over the network by consuming the other trimers in the solution, and create a significant drop of both monomer and linear dimer (Figure 2.13). As more molecules are incorporated into the assembly phase, the concentration of the free molecules in the solution phase decreases, which makes the particle phase dissolve and disappear, as observed in Figure 2.10 at 96 hr.

AIC_c analysis is used to evaluate the validity of the boundaries in the mechanism to address over-fitting in *NFF*-CHO network modeling. The sigmoidal growth of the trimer concentration from solution to particle stages could be the result of self-templating (Figure 2.13). Thus, five models are evaluated here (Table 2.2): 1) a single-stage model (k_1, K_2, k_3, K_4), 2) a single-stage model with the autocatalysis (k_1, K_2, k_3, K_4, k_5, k_6), 3) a two-stage model (k_1, K_2, k_3, K_4 for both solution and particle stages), 4) a two-stage model with autocatalysis in both stages (k_1, K_2, k_3, K_4, k_5, k_6 for both stages), and 5) a two-stage model with autocatalysis in the *particle stage* (k_1, K_2, k_3, K_4 for Solution and k_1, K_2, k_3, K_4, k_5, and k_6 for particle stage). Only data for the first 50 hours are included to evaluate this boundary. AIC_c analysis (Table 2.2) supports the two-stage model, as the Akaike weights of the remaining models are close to zero; the two-stage model without catalysis is almost 99%, corroborating the boundary between the solution and particle stages. The emergence of the particle phase alters the equilibrium between the species, and the kinetic data cannot be described with the same parameter set. Also, AIC_c analysis shows that trimer accumulation is not autocatalytic in both stages; this suggests that the trimer do not self-template in the solution phase, and the particles do not serve as templates to drive the replication of trimer.

Table 2.2 – AIC$_c$ analysis for the model in Figure 2.12 to the kinetic data of the first 50 hours (solution and particle stages) of the *NFF*-CHO DCN. Single-stage includes k_1, K_2, k_3 and K_4 without a boundary. Single-stage with autocatalysis includes k_1, K_2, k_3, K_4, k_5 and k_6 without a boundary. Two-stage includes k_1, K_2, k_3 and K_4 for each of the solution and particle stages. Two-stage with autocatalysis in both stages includes k_1, K_2, k_3, K_4, k_5 and k_6 for each of the Solution and Particle Stages. Finally, two-stage with autocatalysis in the particle stage includes k_1, K_2, k_3 and K_4 for the solution stage and k_1, K_2, k_3, K_4, k_5 and k_6 for the particle stage. N_d is the number of data points and N_p is the number of parameters.

Model	N_d	SSE	N_p	AIC$_c$	Akaike weight (%)
Single-stage with no autocatalysis	30	2.33	4	-67.1	1.43×10^{-12}
Single-stage with autocatalysis	30	2.33	6	-61.0	6.91×10^{-14}
Two-stage with no autocatalysis	30	0.18	8	-131	98.7
Two-stage with autocatalysis in both stages	30	0.18	12	-112	7.50×10^{-3}
Two-stage with autocatalysis in particle stage	30	0.18	10	-122	1.26

The boundary between the particle and fiber stages in the *NFF*-CHO network is also validated by AIC$_c$, together with the autocatalytic nature of the assemblies. The models compared here are: 1) a single-stage model (k_1, K_2, k_3, K_4), 2) a single-stage model with autocatalysis (k_1, K_2, k_3, K_4, k_5, k_6), 3) a two-stage model with autocatalysis in fiber stage (k_1, K_2, k_3, K_4 for the particle stage and k_1, K_2, k_3, K_4, k_5, and k_6 for the fiber stage), 4) a two-stage with nucleation only in fiber stage (k_1, K_2, k_3, K_4 for the particle stage and k_1, K_2, k_3, K_4, and k_5 for the fiber stage), and 5) two-stage with no assembly (k_1, K_2, k_3, K_4 for both the particle and the fiber stage) to fits of data from 9 to 336 hours (Table 2.3). Again both single-stage models (with and without k_5 and k_6) have Akaike weights close to zero, which confirms the different chemistry nature in the particle and the fiber stages. However, the AIC$_c$ analysis does not support the autocatalytic nature (k_6) for the fiber stage, which is common in amyloid self-assembly. The nucleation-only model (in Table 2.3) earns 88.3% of Akaike weight, which is higher than the autocatalytic one (11.6 %). This shows that the resolution of our HPLC data is not sufficient to describe the autocatalytic nature of the *NFF*-CHO trimer assembly. To verify the autocatalytic nature of trimer assembly, fibers of the linear trimers from the *NFF*-CHO network are isolated by centrifugation, and sonicated in an ultrasonic bath to create more active ends. The sonicated seeds then are added into a new network with a 1:10 ratio. As shown in Figure 2.16, the linear trimers of the seeded network accumulate more quickly compared to the unseeded network. This confirms the autocatalytic nature of growth of the trimer assemblies.

Table 2.3 – AIC$_c$ analysis for the model in Figure 2.12 to the kinetic data of 9 to 336 hours (particle and fiber stages) of the *NFF*-CHO DCN. Single-stage includes k_1, K_2, k_3 and K_4 reactions without a boundary. Single-stage with autocatalysis includes k_1, K_2, k_3, K_4, k_5 and k_6 without a boundary. Two-stage with autocatalysis in the fiber stage includes k_1, K_2, k_3 and K_4 for the particle stage and k_1, K_2, k_3, K_4, k_5 and k_6 for the fiber stage. Two-stage with nucleation only in fiber stage includes k_1, K_2, k_3 and K_4 for the particle stage and k_1, K_2, k_3, K_4 and k_5 for the fiber stage. Two-stage with no assembly includes for both the particle and the fiber stage.

Model	N_d	SSE	N_p	AIC$_c$	Akaike weight (%)
Single-stage with no autocatalysis	30	8.48	4	-28.3	4.60×10^{-15}
Single-stage with autocatalysis	30	6.21	6	-31.6	2.38×10^{-14}
Two-stage with autocatalysis in fiber stage	30	0.52	10	-90.0	11.6
Two-stage with nucleation only in fiber stage	30	0.53	9	-94.0	88.3
Two-stage with no assembly	30	6.57	8	-22.7	2.79×10^{-16}

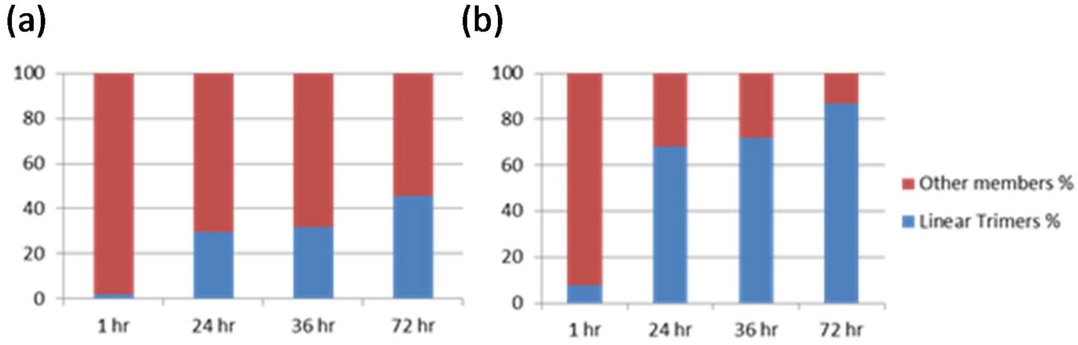

Figure 2.16 – Autocatalytic nature of the trimer assembly in NFF-CHO by seeding the network. (a) Growth of linear trimer and other network members in *NFF*-CHO Network without seeding. (b) Growth of linear trimer and other network members in *NFF*-CHO Network with the seeds (1:10) of trimer assemblies. Image adapted from Ref. [132] with permission.

2.2.3 Compare the NF- and NFF-CHO networks

The evolution pathway of *NF*- and *NFF*-CHO networks is similar, as they both undergo two same phase transitions. The emergence of the particle phase is triggered by the accumulation of the longer oligomers, which reach their solubility and precipitate. The formation of the particle phase is similar to the liquid-liquid phase separation in protein or peptide aggregation [133, 134]. The condensed metastable particle phase helps to overcome the initial desolvation, and provides an environment for the more thermodynamically favored fiber phase to nucleation. This “two-step nucleation” process has been observed in other systems, such as crystallization of proteins [70, 78-83] and minerals [87, 88]. However, the major difference between the DCN phase transitions and the other two-step nucleation system is the interplay between the physical and chemical properties of the networks.

The network evolution is complicated as the physical and chemical distributions are affecting each other. As shown in Figure 2.17, after the monomers oligomerize into longer species, the first physical phase transition (particles) is triggered due to the accumulation of the insolvable oligomers. As suggested by the modeling results, the new physical phase changes the chemical nature of the network, and drives the redistribution between the species. Finally, as the second phase transition, the emergent fiber phase contributes to the second chemical redistribution between the species. Both *NF*- and *NFF*-CHO networks evolve in this process, and in both networks the linear trimers are selected to assemble, perform self-template, and predominate in the network with low dispersity and high fidelity.

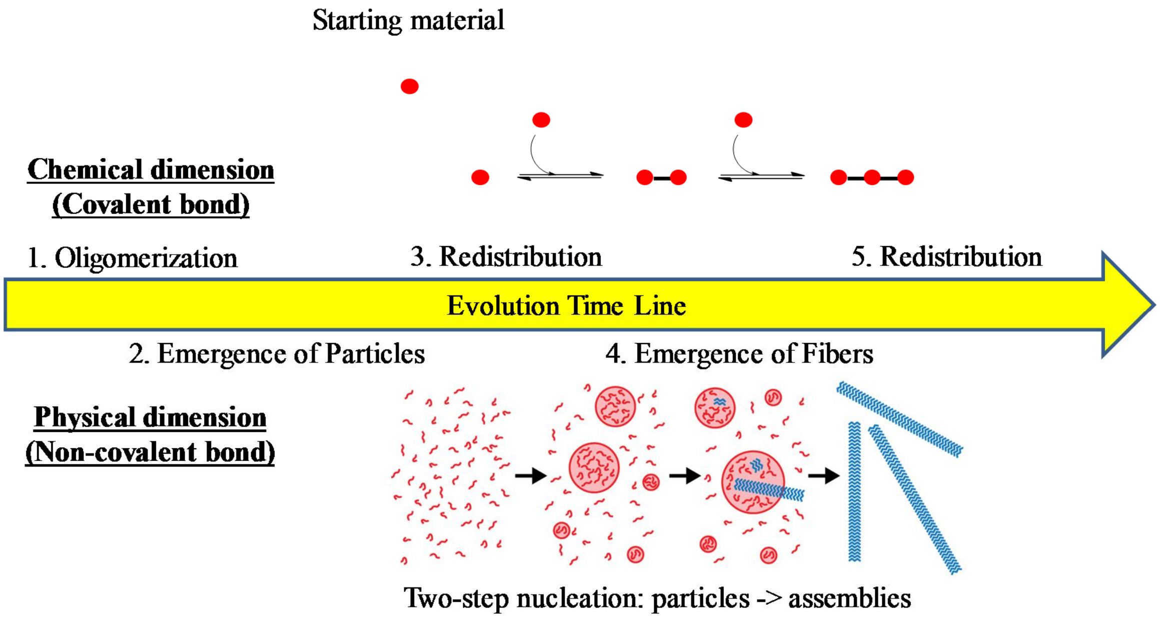

Figure 2.17 – The evolution of the chemical and physical dimentions in DCNs.

The apparent difference of *NFF*- from the *NF*-CHO network is the absence of the cyclic molecules. In the *NF*-CHO network, cyclic dimers and trimers are observed since the solution stage, but the reversible cyclization is hindered in the particle stage, based on the constant cyclic dimer concentration in Figure 2.6. As the particles may disfavor the reversible cyclization of the oligomers, the early appearance of the particles in the *NFF*-CHO network may limit the cyclic products. Due to the extra phenylalanine, the *NFF*-CHO oligomers are more hydrophobic and nucleate particles more quickly. The time window for *NFF*-CHO oligomers to cyclize is shorter, and thus the cyclic molecules are not detected via HPLC.

Both *NF*-CHO and *NFF*-CHO networks follow complex kinetic mechanisms, and they require distinct kinetic constants associated with different stages timed to physical phase transitions. Additional constraints on particle number, size distribution, growth kinetics, and internal structure will be important to a more detailed mechanistic model to cover both the physical and chemical transitions. However, at this point the chemical and physical distributions combine to create dynamics consistent with elementary mass action kinetics.

2.3 Conclusion

In this chapter, the construction and analyses of two dynamic combinatorial networks (DCNs) are discussed, which are built from the *NF*- and *NFF*-CHO monomers individually. Without enzymes driving the oligomerization, these networks show ability to synthesize sequence-defined oligomers with low dispersity via self-templating.

With the experimental results and model discrimination combined, it is shown that the evolution of both networks is stage-wise, and the chemical and physical properties of the networks are essential to derive the final state. As the complex interaction between the physical and the chemical dimensions complicates the design of new networks, the entire process may be simpler to design and engineer if the two-step nucleation process is better understood.

2.4 Methods

2.4.1 Materials

All commercially available chemicals are obtained from Sigma-Aldrich, AnaSpec, Nova Biochem, and Acros organics. HPLC grade acetonitrile and water are obtained from Sigma-Aldrich and/or Fisher Scientific, and distilled deionized water for sample preparation is obtained from EMD chemicals Inc.

2.4.2 Synthesis of NF-CHO

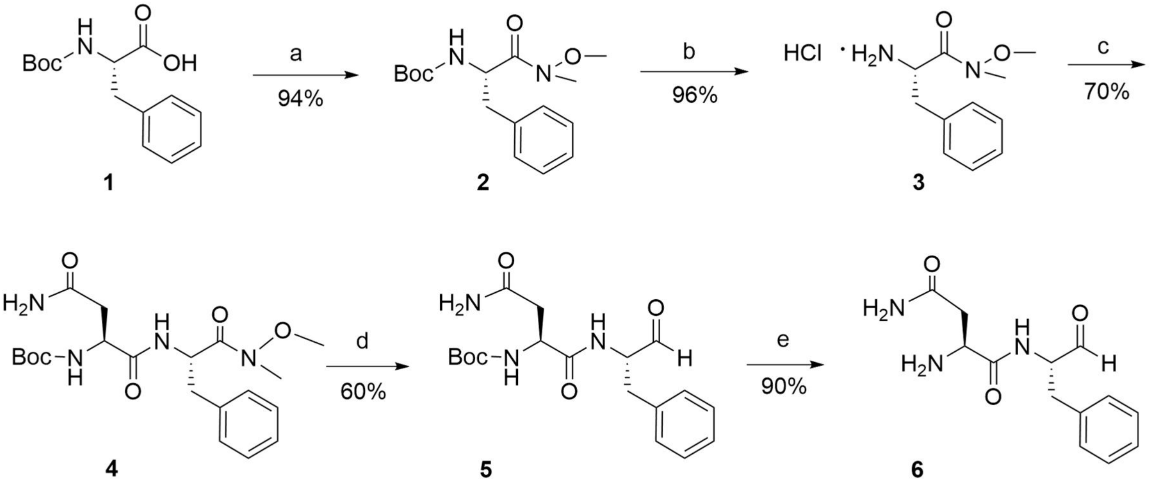

Figure 2.18 – Reaction pathway to synthesize the *NF*-CHO monomer.

The *NF*-CHO monomers are synthesized following Figure 2.18. In step a, N-Boc-L-Phe (molecule **1**) is dissolved in dichloromethane and reacted with 1,1-carbonyldiimidazole and N,O-dimethylhydroyxlamine hydrochloride to obtain Boc-Phe-weinreb (molecule **2**).

The t-butoxycarbonyl protecting group (Boc group) is then removed by HCl (step b) to result molecule **3**, which is then coupled with Boc-Asn (step c) to gain molecule **4**. After the Boc-Asn-Phe-weinreb molecule **4** is reduced (step d), the Boc group of the reduced product **5** from step d is removed to access the free monomer **6**, *NF*-CHO monomer. The detailed steps for synthesizing the *NF*-CHO molecule are indicated in Ref. [132].

2.4.3 Synthesis of NFF-CHO

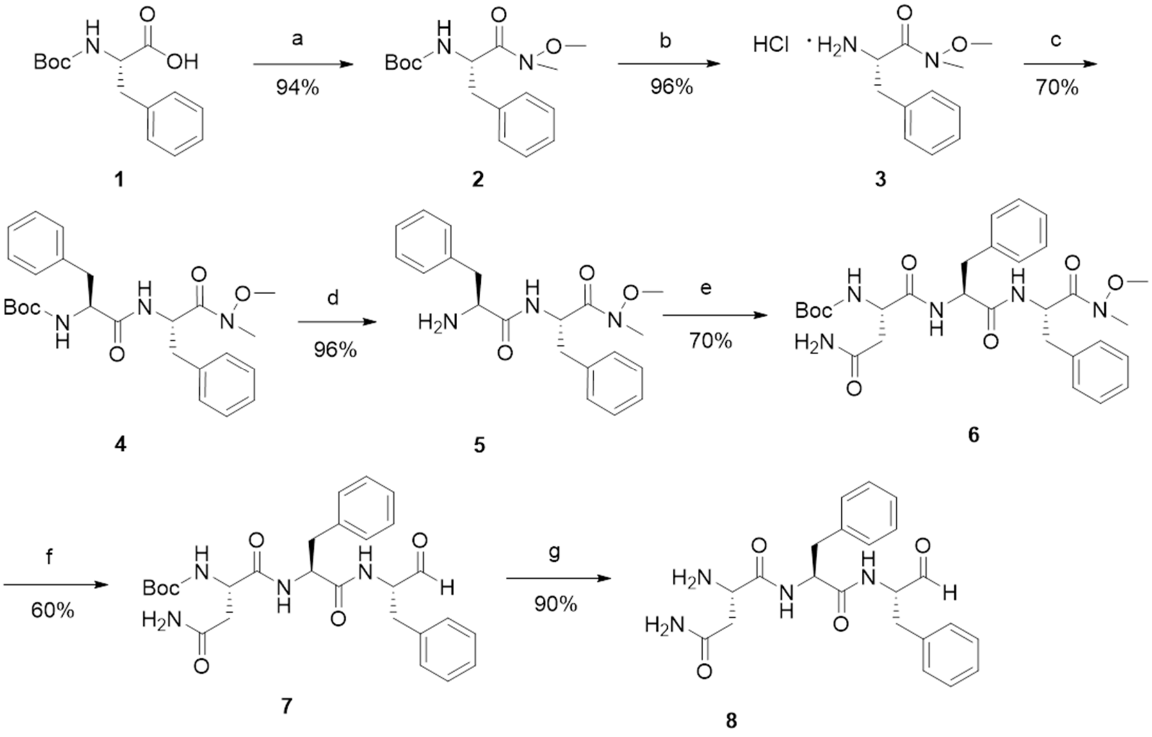

Figure 2.19 – Reaction pathway to synthesize the *NFF*-CHO monomer.

The *NFF*-CHO monomer is synthesized by following Figure 2.19. To synthesize the *NFF*-CHO monomer, N-Boc-L-Phe (molecule **1**) is reacted with 1,1-carbonyldiimidazole and N,O-dimethylhydroyxlamine hydrochloride to obtain Boc-Phe-weinreb (molecule **2**). The t-butoxycarbonyl protecting group (Boc group) is then removed by HCl (step b), and the resulting product **3** is coupled with Boc-Asn (step c) to gain molecule **4**. In step d, the Boc group on Asn is removed with HCl, and the resulting molecule **5** is coupled with Boc-Asn (step e) for molecule **6,** Boc-Asn-Phe-Phe-weinreb amide. After the reduction of the C-terminus of molecule **6**, the Boc group at the N-terminus of molecule **7** is removed to deprotect and activate the free monomer NFF-CHO, molecule **8**. The detailed synthesis procedure for NFF-CHO is in Ref. [132].

2.4.4 Preparation of the Dynamic Combinatorial Network Solution

After final deprotection by trifluoroacetic acid (TFA) of the network monomer, Boc-*NF*-CHO or Boc-*NFF*-CHO, TFA is removed and the monomer is dried under vacuum under room temperature. To start the network incubation, the dried monomer is then dissolved in 40% acetonitrile in water with N_2 protection; dissolution of the monomer is assisted by vortexing and bath sonication until solution becomes clear.

2.4.5 HPLC and LC-MS Analyses of the DCNs

The species of the DCNs are separated and quantified with HPLC. Separation of the DCN members is performed with Waters Delta 600 with reversed-phase column (Kromasil 100-5C18, 4.6 × 250mm) with the mobile phase flow rate equal to 1.0 mL/min. The mobile phase is from 10% to 90% acetonitrile, with the gradient of 2% acetonitrile/min. UV absorbance is monitored at both 258nm (for Phenyl ring side chain absorption) and 222nm (for amide bond absorption). LC-MS analyses are performed on Waters Synapt G2 MS/Acquity UPLC system to separate and characterize the species.

2.4.6 Transmission Electron Microscopy

The morphologies of the macrostructures of the DCNs are observed with transmission electron microscopy (TEM). The DCN solution is loaded onto the copper grid for 3 min, and the excess solution is wicked away with filter paper. The resulting grid is negatively stained with 5% uranyl acetate in water for three minutes before the stain solution is removed with filter paper. The grid is then stored in a vacuum desiccator overnight to remove the remaining liquid. TEM images are recorded with a Hitachi H-7500 transmission electron microscope.

2.4.7 Model Construction and Discrimination

The kinetic models are constructed to describe the evolution of the DCN distribution. The oligomerization reactions are assumed to be reversible; the bond forming reactions are expressed in its reaction constant and the reverse reactions are expressed with the combination of both the bond forming reaction and the equilibrium constants. Assembly kinetics are modeled with the two-step process proposed by Finke [66], which includes an assembly nucleation step and an autocatalytic assembly growth step. The models are fit to the member concentrations by minimizing the sum of square error (SSE) between the calculated and experimental values [135]. A Latin Hypercube sampling [136] provides the initial guesses for the parameters and the optimization procedure is carried out with the *fmincon* function in MATLAB 2012a (The Mathworks, Inc). Model discrimination is achieved with the corrected Akaike information criterion (AIC_c) [129] to verify the consistency of the models and the experimental data.

CHAPTER 3.

A KINETIC MODEL FOR TWO-STEP NUCLEATION OF PEPTIDE ASSEMBLY

3.1 Introduction

In Chapter 2, it is demonstrated that the chemical distribution in the DCNs is affected by the emergent physical phases. After a solution stage, both the *NF*-CHO and the *NFF*-CHO trimer assemblies dominate the system via a metastable particle phase. In this chapter, this complex process is mathematically modeled. First, the peptide assembly for a monodisperse system is modeled with a two-step nucleation pathway, and the monodisperse model shows the capability to rationally explain the kinetics of Aβ(16-22) assembly in various acetonitrile-water solvent mixtures. Finally, this monodisperse model is expanded to a polydisperse model with peptide oligomerization, and both the final chemical distribution and the physical distribution are discussed.

3.2 Results and Discussion

3.2.1 Development of a peptide assembly model with two-step nucleation

For a monodisperse system, the peptide self-assembly under two-step nucleation starts with the formation of the metastable particle phase (Figure 3.1a and Figure 3.1b). This particle nucleation is assumed to be analogous to nucleation of droplets in vapor or solution, which follows the classical nucleation theory [72, 137, 138]. The particle nucleation rate (R_1) is calculated as

$$R_1 = p_1 S \exp\left(-\frac{p_2}{\ln^2 S}\right) \tag{3.1}$$

in which p_1 and p_2 are peptide concentration-independent constants. S is the supersaturation for particles, defined as the ratio of peptide concentration (C) over the solubility of particles (C_{1*}) such that $S = C / C_{1*}$. The solubility of particles is estimated with the Flory-Huggins solution theory [138-140]:

$$C_{1*} = \frac{\rho_{mon} e^{-\chi}}{W_{mon}\left(1 - e^{-\chi}\right)} \tag{3.2}$$

where ρ_{mon} is the peptide density, W_{mon} is the peptide molecular weight and χ is the Flory-Huggins parameter. The typical density of protein and molecular weight of peptide Ac-KLVFFAE-NH_2 are used for ρ_{mon} and W_{mon}, respectively (Table 3.1). The miscibility of the peptide and the solvent is critically tied to the Flory-Huggins parameter (χ). As χ increases, the peptide solubility for particles decreases, and the peptides then have a stronger propensity to nucleate particles. Figure 3.2a shows the simulated Aβ(16-22) peptide solubility as a function of the Flory-Huggins constant (χ), and Figure 3.2(b)-(d) show the development of the particle phase given different χ. The particle nucleus size and the particle growth mechanism are also required to calculate the phase distributions, as shown in Figure 3.2(b)-(d); these mechanism and will be discussed in the following.

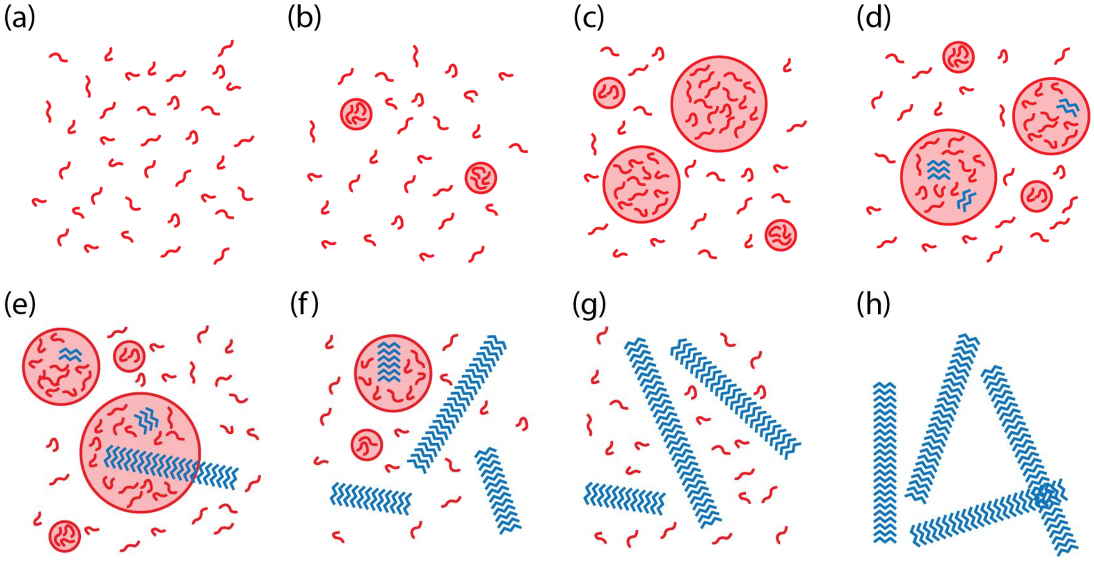

Figure 3.1 – The peptide assembly mechanism under two-step nucleation. (a) Initially, peptides are dissolved in the solution, and (b) the metastable peptide particles nucleate if the peptides are not completely soluble. (c) The particles grow when the solution is supersaturated for particles. (d) Later, the assemblies nucleate inside the particles, and (e) extend into the solution after growing into a critical size by consuming the peptides inside the particles. (f) The assemblies propagate by consuming the free peptides in the solution phase, which decreases the free peptide concentration. (g) The particles start to dissolve when the solution becomes undersaturated for particles. (h) If assembly dissolution is negligible, after all free peptides are depleted, the assemblies become the only species remaining in the solution.

Table 3.1 – Summary of values for the mathematical model in equations (3.1) to (3.10).

Parameter	Description	Units	Value
p_1	Prefactor for particle nucleation rate	L^{-1} sec^{-1}	10^{14}
p_2	Constant in the exponential part of the particle nucleation rate (R_1)	–	10^{0}
ρ_{mon}	Density of peptide	g cm^{-3}	1.4
χ	Flory-Huggins parameter	–	7.5–12
W_{mon}	Molecular weight of peptides	g $mole^{-1}$	894
γ	Interfacial tension between peptide and solvent	g s^{-2}	15
k_b	Boltzmann constant	J K^{-1}	1.38×10^{-23}
T	Temperature	K	300
N_A	Avogadro's number	$mole^{-1}$	6.023×10^{23}
k_{g0}	Kinetic constant for particle growth	$mole^{-1}$ sec^{-1}	10^{0}–10^{4}
k_n	Assembly nucleation constant	sec^{-1}	10^{-7}
k_{AG}	Assembly growth constant inside the particle, (R_4)	sec^{-1}	10^{-2}
ρ_{par}	Particle density	g cm^{-3}	1.4
k_g	Assembly growth constant in the solution phase, (R_6)	$mole^{-1}$ sec^{-1}	10^{2}
k_{bp}	Prefactor of Assembly dissolution	sec^{-1}	10^{10}
ΔG	Binding energy between peptides in the assembly	J	$10^{-19.1}$–$10^{-18.8}$
C_0	Initial peptide concentration	mM	0.5–1
N	Sample size	–	10^{5}

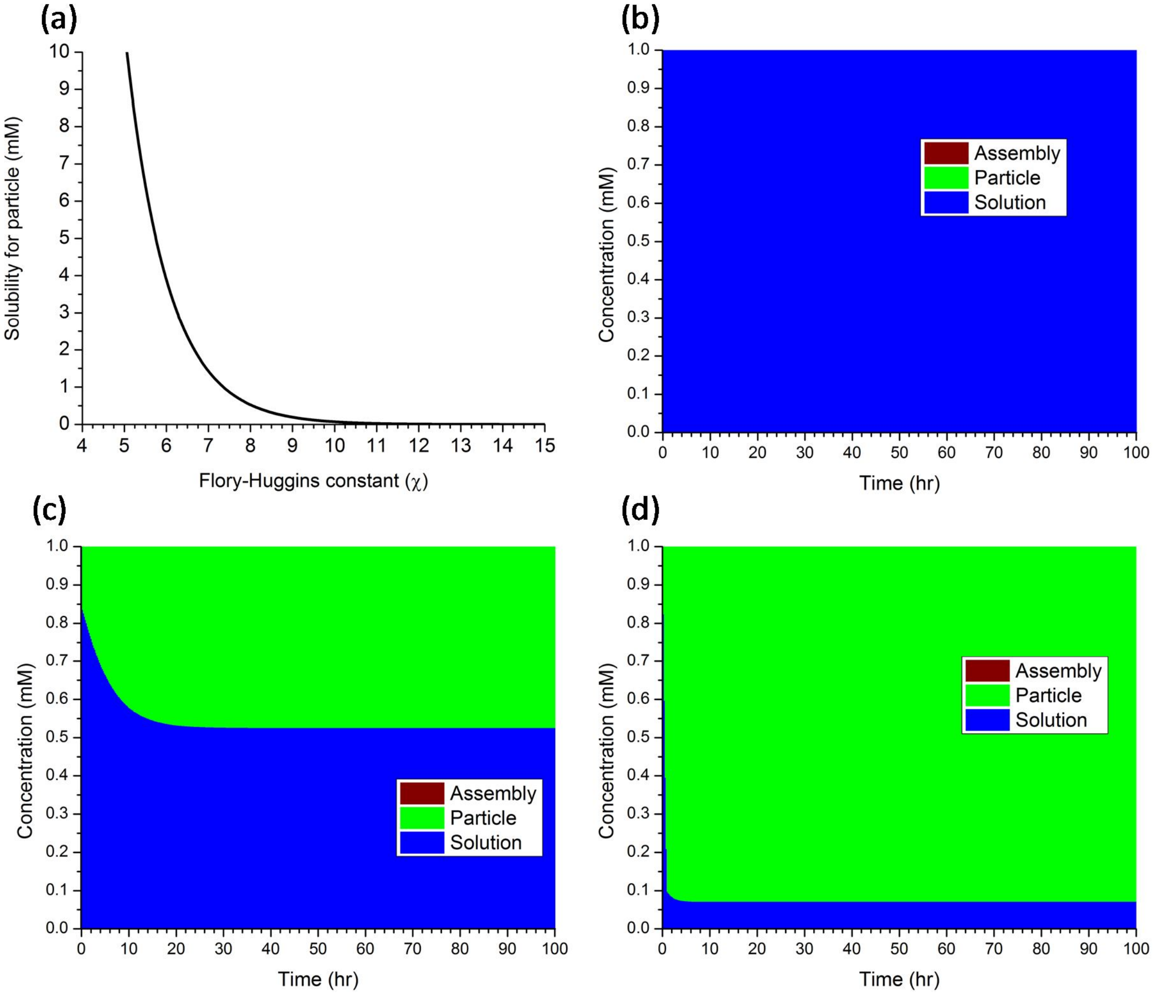

Figure 3.2 – The peptide solubility for particles as a function of the Flory-Huggins constant (χ). (a) The peptide becomes less soluble as χ increases. (b) Simulation of the particle phase development with $\chi = 6$. (c) Simulation of the particle phase development with $\chi = 8$. (d) Simulation of the particle phase development with $\chi = 10$. The parameters used for these simulation are taken from Table 3.1 with $k_n = 0$, $k_{g0} = 10$ mole^{-1} sec^{-1} and $C_0 = 1$ mM.

A stable particle nucleates with a critical number of peptides to overcome the interfacial energy with the bulk free energy change. This critical number (n^*) is then the particle nucleus size, which is determined by the Gibbs-Thompson equation [141]:

$$n^{*} = \frac{32\pi\gamma^{3}}{3\left(k_{b}T\right)^{3}\ln^{3}S}\left(\frac{W_{mon}}{N_{A}\rho_{mon}}\right)^{2} \tag{3.3}$$

where γ is the interfacial tension between peptide and solvent, k_b is Boltzmann's constant, N_A is Avogadro's number, and T is the temperature, which is set as a constant in this study. Whenever a particle is formed, n^* peptides are removed from the solution phase and form a new particle. Equation (3.3) suggests that, under a supersaturation S, a particle remains stable against dissolution as long as its size is at least n^*; hence, the stability of a particle depends on its size, and so does the peptide solubility of that particle [130, 142, 143]. However, here the peptide solubility of all particles is assumed to be size-independent [144-146] for simplicity.

The particles may grow or dissolve (Figure 3.1c), which depends on the solubility of particles (C_{1*}). When particle growth is bulk diffusion-limited, the net particle growth rate for an individual particle with x peptides may be expressed as [143, 145]:

$$r_{2}\left(x\right) = k_{g0}x^{1/3}\left(C - C_{1*}\right) \tag{3.4a}$$

and the overall particle growth rate for all the particles with x peptides is:

$$R_{2}\left(x\right) = k_{g0}x^{1/3}\left(C - C_{1*}\right)P\left(x\right) \tag{3.4b}$$

where k_{g0} is the particle growth constant and $P(x)$ is the total number of particle with x peptides. This net expression results in either particle growth or dissolution based on the peptide concentration in the solution phase, C. When $C > C_{1*}$, solution is supersaturated for particle and particles grow according to R_2 until $C = C_{1*}$. On the other hand, when $C < C_{1*}$, the solution is undersaturated and the particles dissolve until $C = C_{1*}$. Thus for all particles, both their growth and dissolution stop when $C = C_{1*}$ and they never happen simultaneously.

Assemblies nucleate inside the metastable particles (Figure 3.1d). As the assembly nucleation rate (R_3) is a function of particle mass [97, 147], R_3 for the particles with x peptides is formulated as:

$$R_3(x) = k_n P(x) x \tag{3.5}$$

where k_n is the assembly nucleation constant. The assembly nucleus size is assumed to be two, the minimal number to form assemblies with an intermolecular bonding [65]. As the peptide concentration in the particles is fixed, it is not explicitly represented in equation (3.5). Once the reaction happens, two of the peptides in the particle become an assembly, and the assembly grows by consuming the other peptides in the same particle. The net assembly growth rate inside particles with x free peptides (R_4) is expressed as [90]:

$$R_4(x) = k_{AG} A_p(x) \tag{3.6}$$

where k_{AG} is the growth constant and $A_p(x)$ is the total number of assemblies in particles with x peptides. With R_3 and R_4, multiple nuclei may nucleate and grow within the same particle, as observed previously [148]. This makes the model more flexible compared to other studies [72, 97, 147], which are limited to one nucleus per particle at most.

Once the assemblies grow long enough, their sizes (L_A) become greater than the particle diameter (L_P), and thus they are exposed to the solution phase (Figure 3.1e). By assuming that the assembly propagates at the two ends, L_A for an assembly containing y peptides is then calculated as $0.5(y-1)$ nm, where 0.5 nm is the typical spacing between peptides in the assemblies [149]. The particle diameter is calculated by assuming that the particle density (ρ_{par}) is similar to the typical protein density (Table 3.1). The diameter for a particle with x peptides is calculated as:

$$L_P(x) = 2\left(\frac{xW_{mon}}{N_A \rho_{par}} \frac{3}{4\pi}\right)^{\frac{1}{3}} \tag{3.7}$$

Before the sizes of the assemblies reach the critical value L_P, the assemblies stay in the particles. Once an assembly with y peptides emerges into the solution, those y peptides are then removed from the particle phase and added into the assembly phase (Figure 3.1f, Figure 3.1g, and Figure 3.1h). The following growth of assembly in the solution phase is formulated as a function of the remaining peptides in the solution phase, and the corresponding growth rate (R_5) for assemblies with y peptides is:

$$R_5(y) = 2k_g A(y) C \tag{3.8}$$

where k_g is the assembly growth constant in solution phase, and $A(y)$ is the concentration of assemblies with y peptides in the solution phase. The leading “2” in R_5 stands for the two active ends per assembly. The peptide dissociation may occur at the assembly ends due to instability. The assembly dissolution rate (R_6) is formulated as:

$$R_6\left(y\right)=k_{bp}\exp\left(\frac{-\Delta G}{k_bT}\right)B\left(y\right)A\left(y\right) \tag{3.9}$$

where k_{bp} is the prefactor of the rate constant, ΔG is the binding energy[150] between the assembled peptides, and $B(y)$ is the number of breakable bonds at the assembly ends. If $y = 2$, there is only one bond breakable and thus $B(y) = 1$; if $y > 2$, both the bonds at the ends are breakable and $B(y) = 2$. The dissolution rate constant is expressed in an Arrhenius-type instead of a single constant to keep it flexible enough to predict an assembly system with peptide oligomerization [98], whose binding energy is proportional to the oligomer length [150]. When the assembly formation inside the particle is not significant, the concentration of the assembled peptides changes only when the assembly grow or dissolve. Thus, the assembly phase reaches a steady state if the assembly growth rate (R_5) and the assembly dissolution rate (R_6) are equal; given $R_5 = R_6$, the solubility of assemblies (C_{2*}) may be calculated with the assumption that for all assemblies, $y > 2$:

$$C_{2*}=\frac{k_{bp}\exp\left(\frac{-\Delta G}{k_bT}\right)}{k_g} \tag{3.10}$$

Figure 3.3 shows the solubility for assembly as a function of the stacking energy between the assembled peptides (ΔG). As ΔG increases, the peptide solubility for assemblies decreases significantly (Figure 3.3a); the assemblies become stable and start to accumulate by consuming the other phases (Figure 3.3(b-d)).

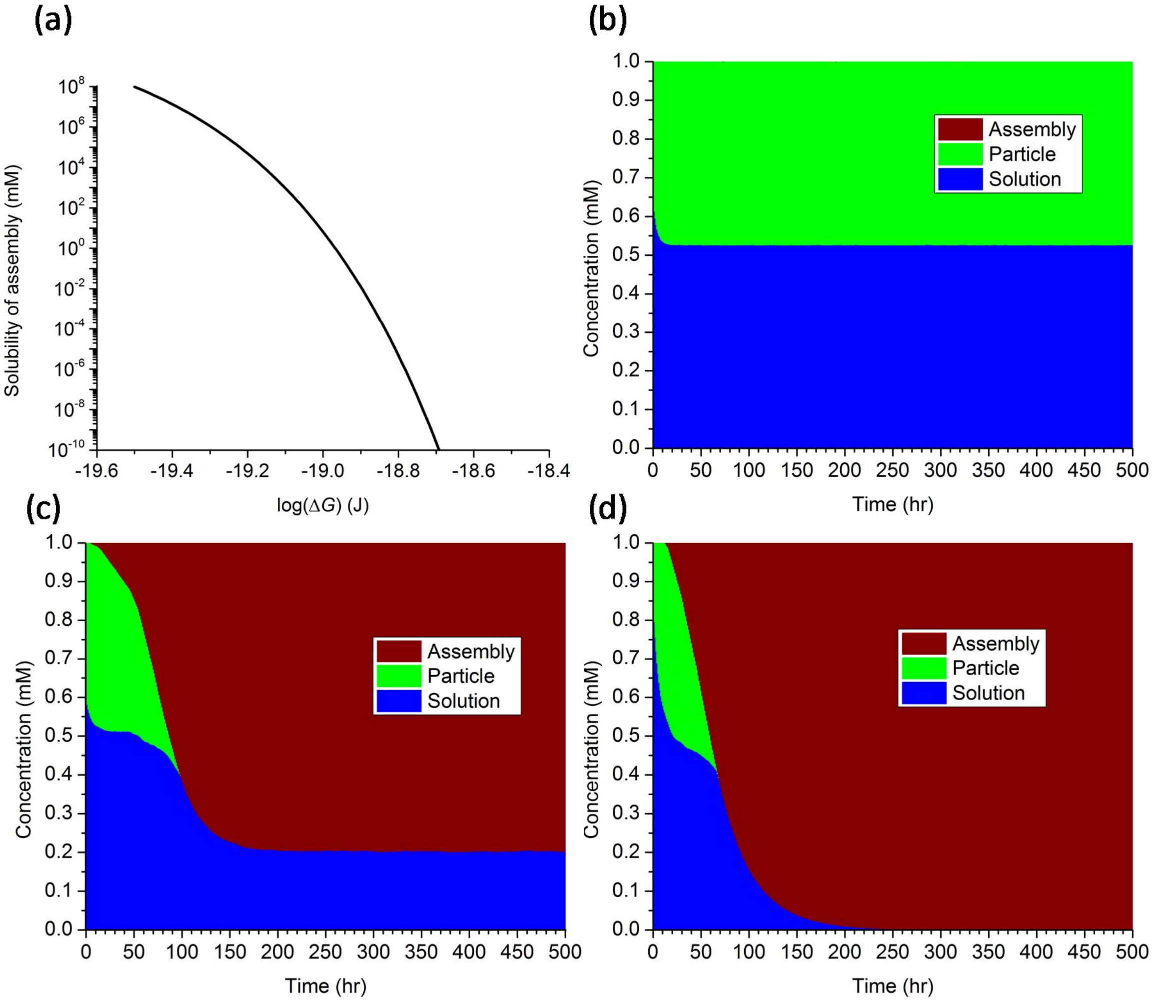

Figure 3.3 – The peptide solubility for assembly as a function of the assembly stacking energy (Δ*G*). (a) The peptide becomes less soluble as (Δ*G*) increases. (b) Simulation of the assembly phase development with $\Delta G = 10^{-20}$ J. (c) Simulation of the assembly phase development with $\Delta G = 1.14 \times 10^{-19}$ J. (d) Simulation of the assembly phase development with $\Delta G = 10^{-17}$ J. The parameters used for these simulation are taken from Table 3.1 with $\chi = 8$, $k_{g0} = 10$ mole^{-1} sec^{-1}, and $C_0 = 1$ mM.

3.2.2 Simulation of the Peptide Assembly Model with Two-step Nucleation

The developed model is executed with the Monte Carlo algorithm, and three parameters are selected to simulate the transition from the particle phase to the assembly phase. The particle growth constants (k_{g0}) and Flory-Huggins parameters (χ) are varied to investigate the transition of the metastable particle phase, while the stability of assemblies and the assembly dissolution rate are varied as a function of stacking energy of the peptides (ΔG). To verify the threshold supersaturation for particle nucleation, the lower limit of χ is selected to make the peptide solubility close to the initial peptide concentration of 1 mM, while the upper limit is selected to make the peptide extremely insoluble (Figure 3.2a). The upper and lower limits for ΔG are selected to span the range from very soluble to extremely insoluble, under a 1 mM peptide concentration (Figure 3.3a). By changing k_{g0}, the contribution of the net particle growth rate may be investigated without altering the thermodynamic properties of the system. 200 parameter sets are generated based on these three variables with a Latin hypercube sampling [136], with parameter ranges given in Table 3.1. The results presented here are simulated up to 100 hr as a practical experimental time scale, and the behaviors of the model under these parameters are classified in Table 3.2. For the details about the criteria categorizing these results, see Section 3.4.2 in Methods.

Table 3.2 – Distribution of simulation results from the general model with 200 parameter sets.

Case	Number of cases	Percentage (%)
1. Pure solution phase	7	3.5
2. Pure assembly phase	52	26
3. Solution-assembly equilibrium	16	8
4. Solution-particle equilibrium	103	51.5
5. Three-phase equilibrium	5	2.5
6. Slow kinetics	17	8.5

For some parameter sets, the simulated results show no particle formation within the reaction time (Figure 3.1). There are seven parameter sets (3.5%) that remain in this pure solution phase (Case 1). For all parameter sets, at the initial time the solution is supersaturated since the free peptide concentration C is greater than the solubility of particles C_{1*}. Thus, the supersaturation for particles, defined as $S = C/C_{1*}$, is greater than one. However, the peptides will not form particles within the reaction time unless the solubility C_{1*} is below a threshold value [90], and this threshold value controls the particle nucleation kinetics. As shown in Figure 3.4, within the reaction time of 100 hr, the particles do not nucleate unless the initial supersaturation is greater than 1.35. The threshold value for particle nucleation is then around 0.74 mM, based on this critical supersaturation of 1.35 and the initial peptide concentration of 1 mM. A supersaturation between 1.0 and 1.35 is then defined as the metastable zone for particle formation. Particles nucleate within the reaction time only when the supersaturation for particles is greater than this threshold value; otherwise no particles are observed within the reaction time.

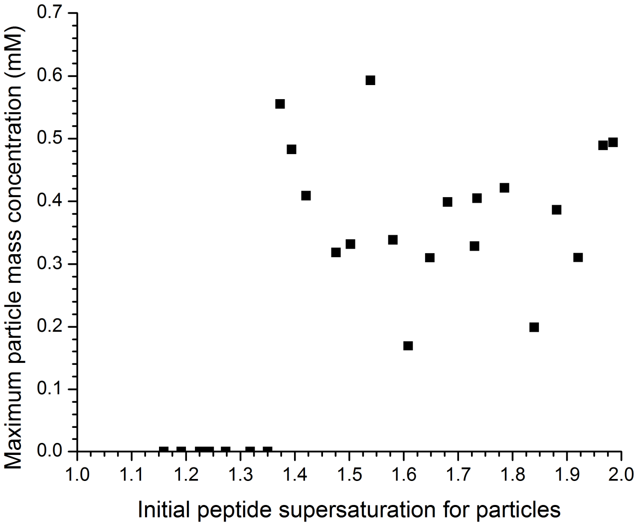

Figure 3.4 – Threshold supersaturation to trigger particle nucleation. The particles nucleate within the simulation time only when the initial supersaturation is greater than 1.35; the corresponding threshold value for particle nucleation is 0.74 mM. The maximum particle mass concentration for each parameter set is defined as the maximum peptide concentration in the particle phase within the reaction time.

Under two-step nucleation, the metastable particles transition as the assembly phase matures. As illustrated in Figure 3.5 for a single parameter set, all peptides are consumed by the assembly phase and this result is categorized as the pure assembly phase (Case 2 in Table 3.2). As shown in Figure 3.6, Case 2 on average has the highest ΔG compared to the other cases. The assemblies in the solution phase do not consume the peptides inside the particle phase directly. Rather, once the solution is undersaturated for particles, the particles start to dissolve and release the peptides into the solution. The assemblies then consume the released peptides, and become the dominating phase at equilibrium.

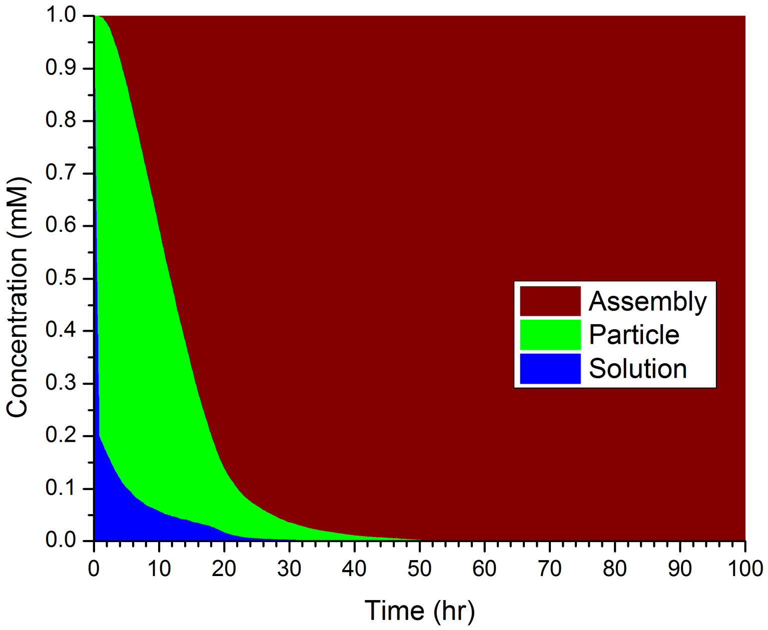

Figure 3.5 – The assembly phase dominates after the solution and the particle phase vanish. The particles nucleate since the peptide solution is saturated and vanish after the peptides are depleted because of the consumption from the assembly phase. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 16.6 mole^{-1} sec^{-1}, χ = 9.27, ΔG = 1.44 × 10^{-19} J, and C_0 = 1 mM.

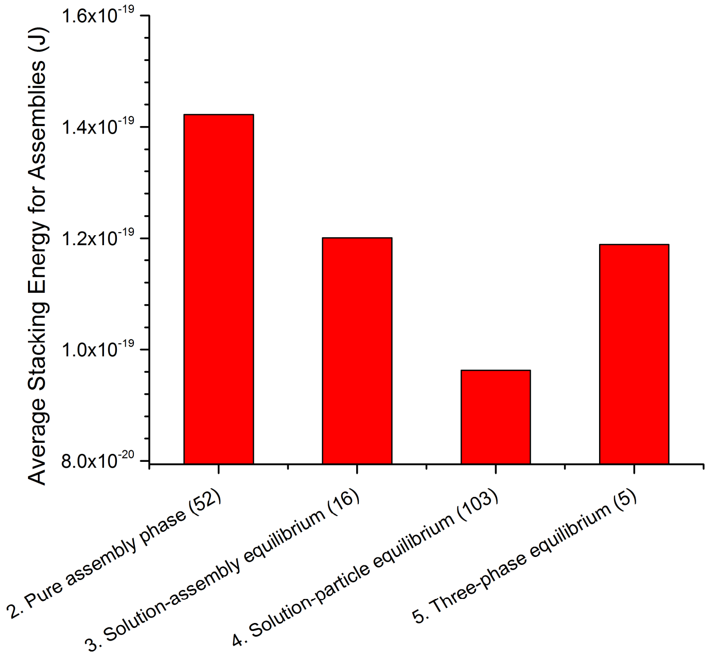

Figure 3.6 – Average stacking energy (Δ*G*) for Case 2 (pure assembly phase), Case 3 (solution-assembly equilibrium), Case 4 (solution-particle equilibrium) and Case 5 (three-phase equilibrium).

If the free peptides are soluble for assemblies due to a weaker stacking energy, Δ*G,* the free peptides may coexist and equilibrate with the assemblies, as illustrated in Figure 3.1g and Figure 3.7. This solution-assembly equilibrium is classified as Case 3 in Table 3.2. The final peptide concentration, determined by the solubility of assemblies, is not sufficient to sustain the particles. The metastable particle phase dissolves as in Case 2, and the difference between Case 2 and Case 3 is the reversibility of assembly growth. In Case 2 the growth of assemblies is nearly irreversible, while Case 3 is obtained when the assemblies exhibit significant reversibility.

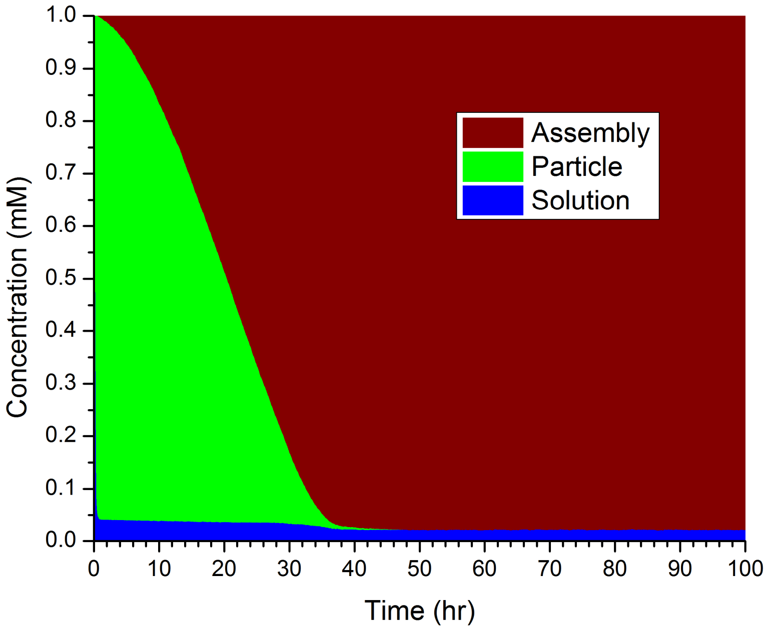

Figure 3.7 – The system reaches equilibrium between the solution and the assembly phases. The particle phase vanishes after the assembly phase becomes dominant, while assemblies are still partially soluble. The final state is the equilibrium between the solution and the assembly phases. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 128.3 mole^{-1} sec^{-1}, χ = 10.25, ΔG = 1.23 × 10^{-19} J, and C_0 = 1 mM.

When the stacking energy of assemblies is not strong enough, the assemblies cannot grow in the solution. The remaining solution and particle phases reach equilibrium with each other, as shown in Figure 3.1c, Figure 3.1d and Figure 3.8. This solution-particle equilibrium occurs in over half of the parameter sets (Case 4 in Table 3.2), and is characterized by the absence of the assembly phase. As shown in Figure 3.1, on average Case 4 has the weakest stacking energy (ΔG) compared to the other cases with an assembly phase, and this solution-particle equilibrium has been experimentally observed [21]. Kinetically limiting the assembly nucleation rate (k_n) may also result in achieving a stationary state between the solution and particle phases [72]. Under this condition, the particle phase develops and coexists with the solution phase because the assembly nucleation rate is extremely slow, and the system remains kinetically trapped.

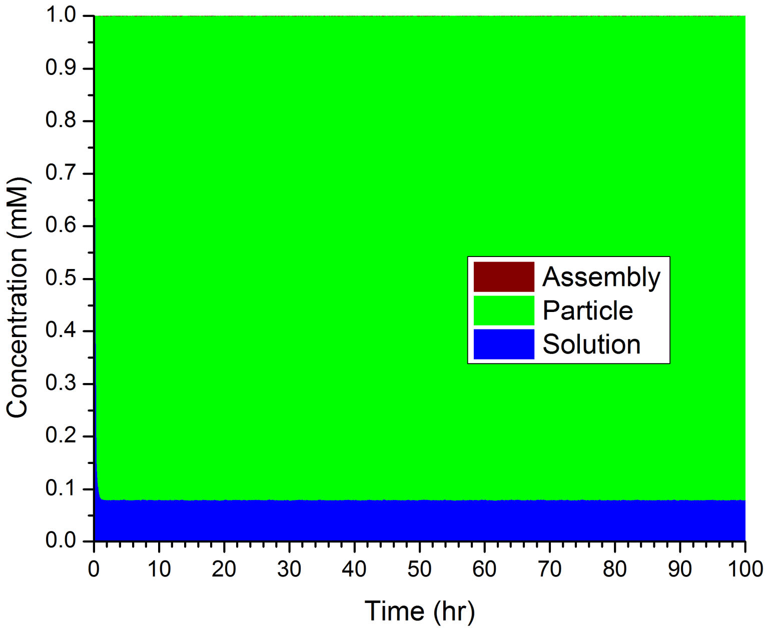

Figure 3.8 – Equilibrium between particle and solution phases is achieved when the assemblies are not stable. The assemblies may not be stable enough to maintain a positive growth rate. The remaining phases are the particle and the solution phases. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 128.3 mole^{-1} sec^{-1}, χ = 10.25, ΔG = 1.23 × 10^{-19} J, and C_0 = 1 mM.

When both the assembly and particle phases are observed and the distribution between phases reaches a stationary state, the relatively rare three-phase equilibrium can be observed (Figure 3.1f and Figure 3.9). Both particle and assembly supersaturations are crucial to achieving the coexistence of all three phases (Figure 3.10). When the particle phase has the higher solubility, it is eliminated and the result is classified as either Case 2 or 3. When the assembly phase has the higher solubility, the result is classified as the solution-particle equilibrium (Case 4). To obtain the rare three-phase equilibrium (Case 5), the solubilities for particles and assemblies need to be high enough to maintain the solution phase, and close enough to each other to stabilize both phases. Figure 3.10 suggests that the final equilibrium states can be rationally predicted as a function of the solubilities for particles and assemblies, as they are influenced by the Flory-Huggins parameter (χ) and the stacking energy of the assemblies (ΔG), respectively.

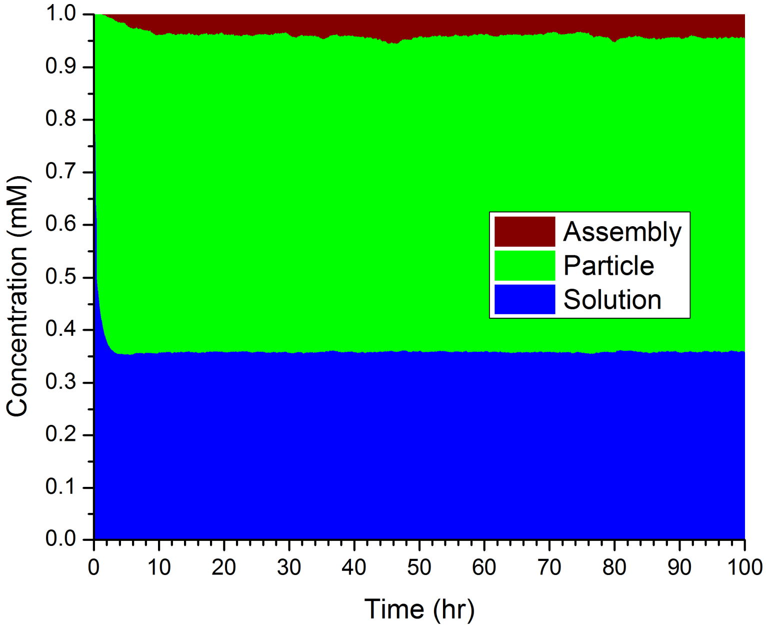

Figure 3.9 – The rare 3-phase equilibrium is achieved when both particle and assembly phases exist. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 149.8 mole^{-1} sec^{-1}, χ = 8.40, $\Delta G = 1.12 \times 10^{-19}$ J, and C_0 = 1 mM.

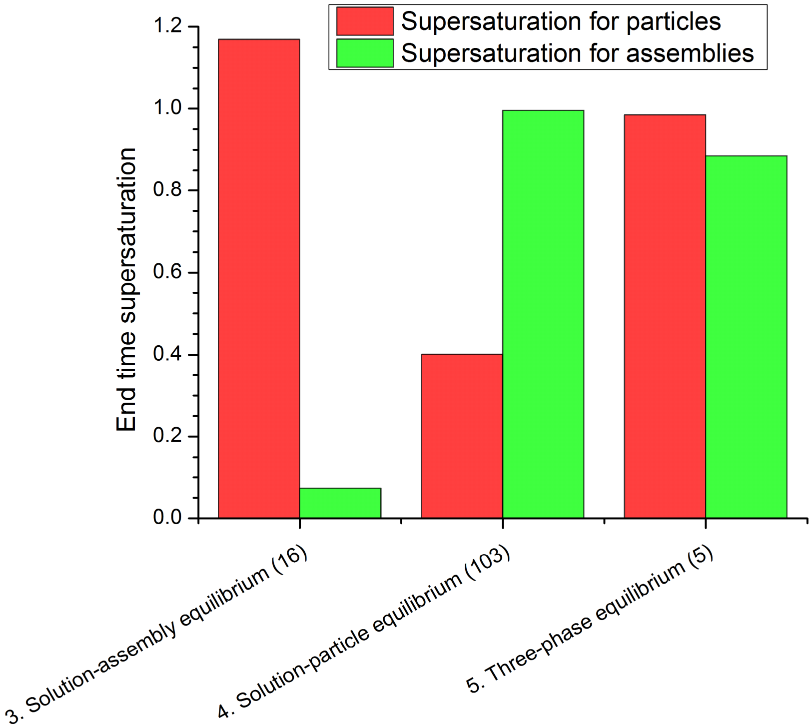

Figure 3.10 – Average final supersaturations for particles and assemblies determine the final phase distribution. The numbers in the parentheses are the numbers of results in each case.

The final states of a 1.0 mM peptide solution are simulated with different solubilities for particles and assemblies (Table 3.3). For a 1 mM peptide solution, both solubilities are in the range from 0.005 mM (least soluble) to 1.1 mM (totally soluble). Figure 3.11 shows the prediction of the solubility difference between the particle and assembly phases based on Table 3.3, with equation (3.2) and equation (3.10) used to derive the particle (C_1^*) and assembly (C_2^*) solubilities, individually. The diagonal of Figure 3.11 represents the equal solubility between the two phases; the region above the diagonal represents the assembly-rich results while the region below the diagonal represents the particle-rich results. Although Figure 3.11 predicts the final thermodynamic states, it does not take the kinetic information into consideration.

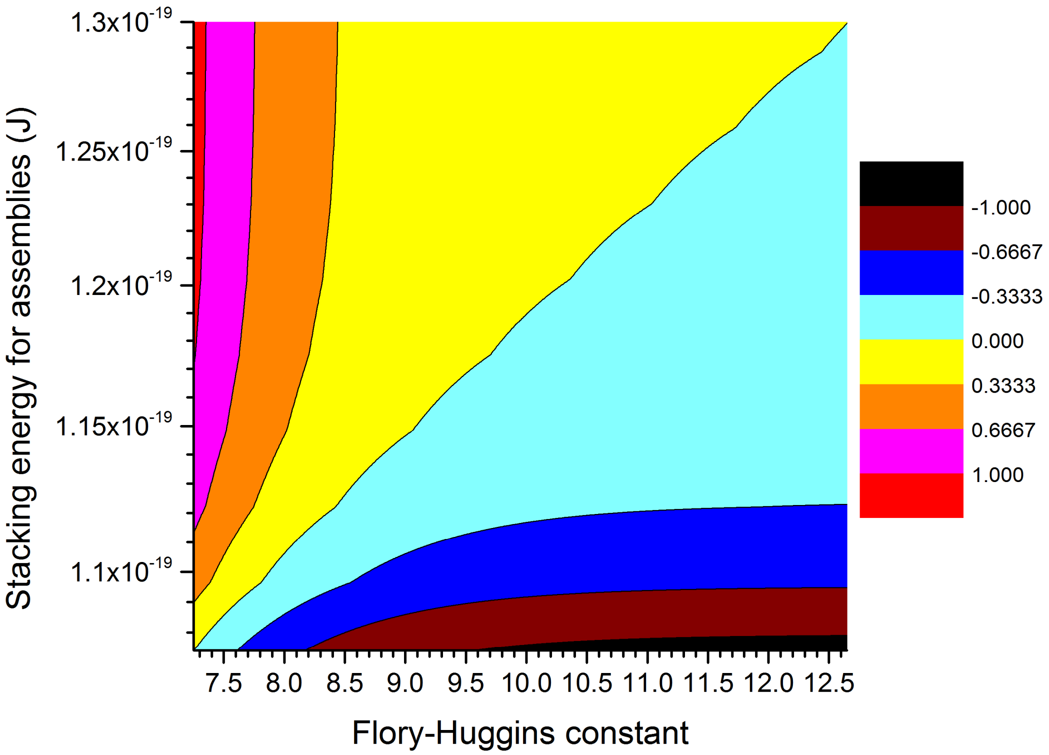

Figure 3.11 – The solubility difference between the particles and the assemblies as a function of the Flory-Huggins constant (χ) and the stacking energy (ΔG). The ranges of the χ and ΔG values in Table 3.3 and the parameter values in Table 3.1 are used. The difference is calculated as the solubility difference between particles and assemblies: $C_{1*} - C_{2*}$.

Table 3.3 – Flory-Huggins parameter (χ) and stacking energy of assembly (ΔG) used in Figure 3.12. The values in the parentheses indicate the corresponding solubility for particles and assemblies given the Flory-Huggins parameter (χ) and stacking energy (ΔG), respectively.

χ (and solubility for particles, mM)	ΔG (J) (and solubility for assemblies, mM)
12.65 (0.005)	1.297×10^{-19} (0.005)
9.66 (0.1)	1.173×10^{-19} (0.1)
8.97 (0.2)	1.144×10^{-19} (0.2)
8.56 (0.3)	1.128×10^{-19} (0.3)
8.27 (0.4)	1.116×10^{-19} (0.4)
8.04 (0.5)	1.107×10^{-19} (0.5)
7.86 (0.6)	1.099×10^{-19} (0.6)
7.71 (0.7)	1.093×10^{-19} (0.7)
7.57 (0.8)	1.087×10^{-19} (0.8)
7.46 (0.9)	1.082×10^{-19} (0.9)
7.35 (1.0)	1.078×10^{-19} (1.0)
7.26 (1.1)	1.074×10^{-19} (1.1)

Figure 3.12 shows the kinetic simulation with the entire two-step nucleation model, from equation (3.1) to equation (3.9), as a function of the Flory-Huggins constant (χ) and the stacking energy (ΔG). As shown in Figure 3.12, no particle forms when the solubility of particles is equal to or greater than 0.8 mM, consistent with the threshold value (0.74 mM) obtained from Figure 3.4; this pure solution phase is not predicted in Figure 3.11, which does not include the kinetic information. When the solubility of particles is equal to or less than 0.7 mM, different phase equilibria are observed. The assemblies survive when they are less soluble than the particles, and vice versa. The rare three-phase equilibrium is reproduced with equal solubilities for particles and assemblies, as predicted. However, when both the solubilities of particles and assemblies are as low

as 0.1 mM, the transition to the three-phase equilibrium becomes slow and does not reach a steady state within the reaction time. Such slow kinetics, which are classified as Case 6, will be further discussed below.

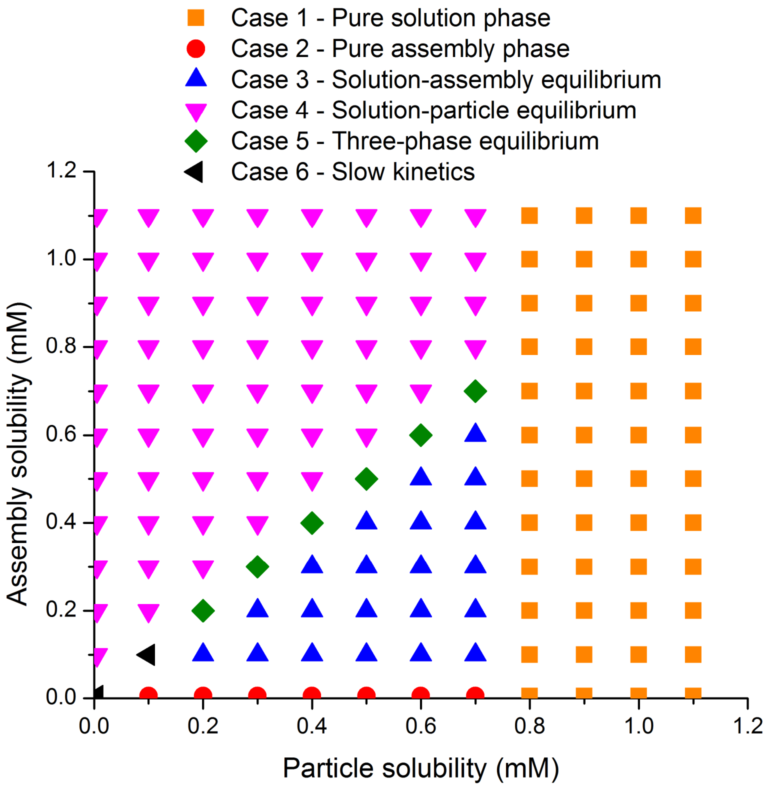

Figure 3.12 – Final states predicted by the model for 1.0 mM peptide solution as a function of solubilities for particles and assemblies. The parameter sets are listed in Table 3.1 and Table 3.3 with $k_{g0} = 10^4$ mole^{-1} sec^{-1}, and $C_0 = 1$ mM, except $k_n = 0.5 \times 10^{-7}$ sec^{-1}. k_n is decreased from 10^{-7} sec^{-1} in Table 3.1 to ensure that the solubility of the assemblies is not affected by the assembly nucleation in the particles, and thus three-phase equilibrium may be achieved within the reaction time.

Not all parameter sets achieve their final thermodynamic equilibria by the end of the reaction time. Although most of the parameter sets do predict a final distribution within the reaction time of 100 hr, 17 out of 200 parameter sets exhibit slow kinetics (Case 6), and the unfinished reactions are attributed to slow particle dissolution (Figure 3.13). Although the assembly phase is already growing outside of the particles, the relatively stable particle phase dissolves slowly even though the free peptide concentration is close to zero. On average, the parameter sets contributing to slow kinetics have the lowest particle growth constant (k_{g0}), as shown in Figure 3.14. Particle dissolution, which shares the same rate equation with particle growth (equation (3.4a) and (3.4b)), delays the entire process and thus the process remains unfinished after 100 hr. To verify the importance of this kinetic constant without changing the final thermodynamic distribution, the simulation in Figure 3.13 is repeated with k_{g0} increased to its upper limit indicated in Table 3.1. With fast particle dissolution, the peptide self-assembly process is finished within the reaction time (Figure 3.15). Taken together, these results demonstrate that the final states in peptide assembly via 2SN are controlled principally by the Flory-Huggins parameter and the peptide stacking energy. The system may be slowed by a small particle growth constant controlling both the particle growth and dissolution rates, such that assembly growth is delayed due to the scarce peptide resource in solution.

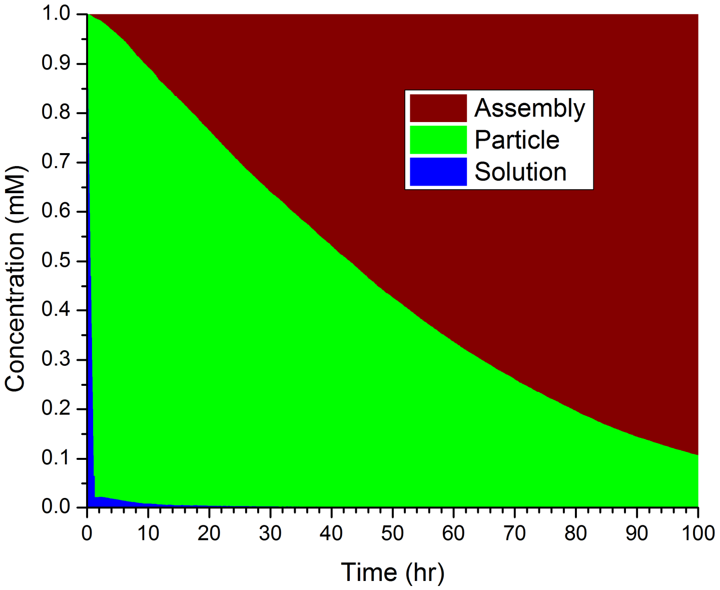

Figure 3.13 – The slow particle dissolution makes the kinetics slow. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 2.16 mole^{-1} sec^{-1}, χ = 10.80, ΔG = 1.51 × 10^{-19} J, and C_0 = 1 mM.

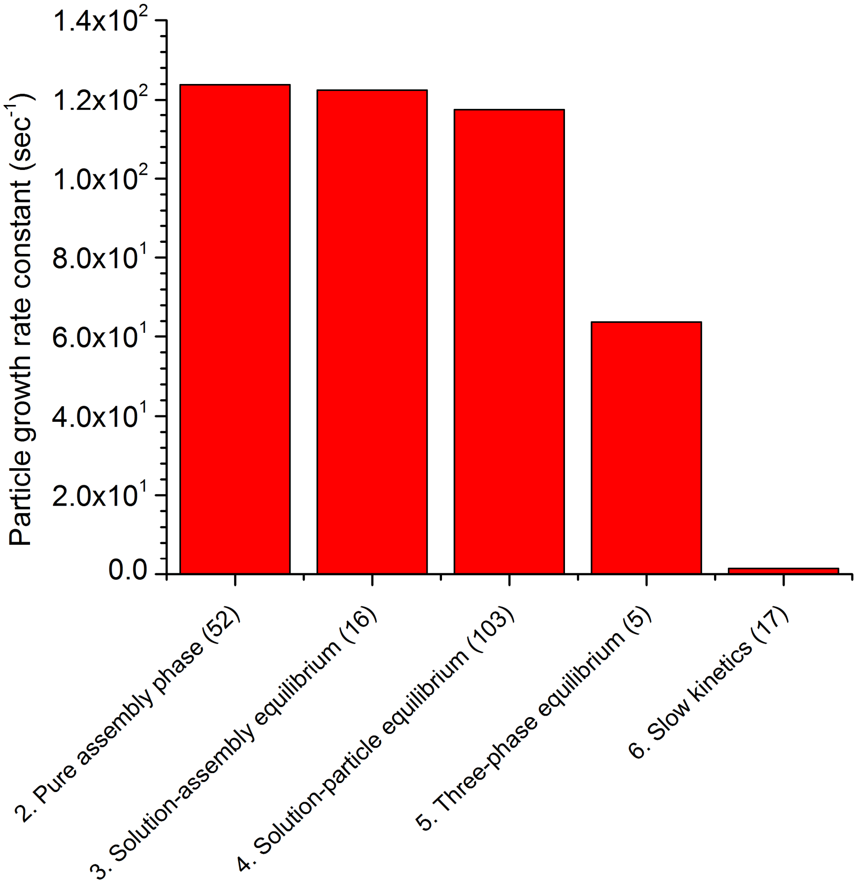

Figure 3.14 – Average particle growth constant (k_{g0}) for Case 2 (pure assembly phase), Case 3 (solution-assembly equilibrium), Case 4 (solution-particle equilibrium), Case 5 (three-phase equilibrium) and Case 6 (slow kinetics). On average, Case 6 has the lowest particle growth rate constant. The numbers in the parentheses are the number of results in those cases.

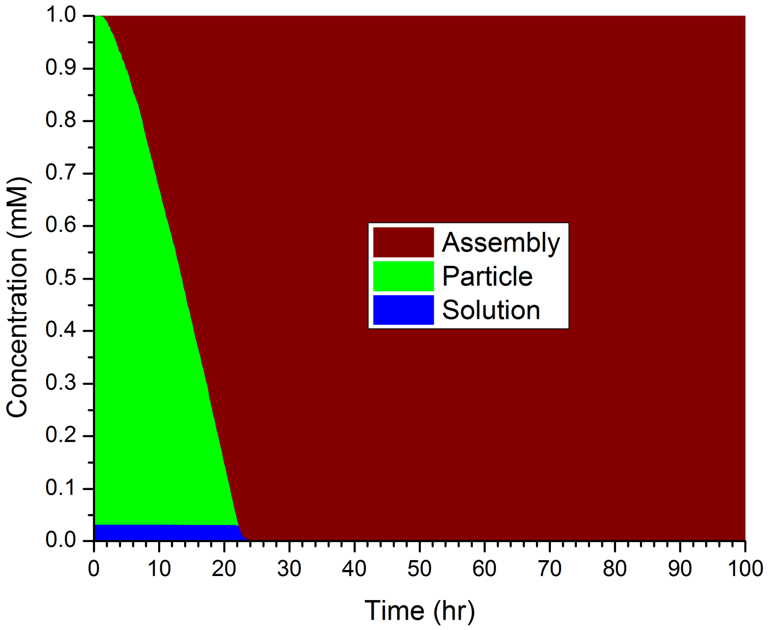

Figure 3.15 – The slow kinetics in Figure 3.13 reach the final state within the reaction time given the increased particle growth constant. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 10^4 mole^{-1} sec^{-1}, χ = 10.80, ΔG = 1.51×10^{-19} J, and C_0 = 1 mM.

3.2.3 The Model Predicts Aβ(16-22) Assembly

The structurally defined Aβ(16-22) assemblies provide an opportunity to validate the model's utility, and accordingly we incubated the purified Aβ(16-22) peptide in different aqueous acetonitrile (ACN) mixtures to follow assembly maturation. We initially predicted the kinetics of Aβ(16-22) assembly would be slower in more hydrophobic environments, where the assemblies are more soluble [21]. Here ACN is used to increase the solvent hydrophobicity and stabilize the free Aβ(16-22) peptide. A series of topography images of Aβ(16-22) solutions in 40% ACN are measured by transmission electron microscopy (TEM) over time. As shown in Figure 3.16a after one hour of incubation, 0.5 mM of Aβ(16-22) is above the threshold value for particle nucleation in 40% ACN. The particles are evenly distributed on the EM copper grid with homogeneous particle sizes. In the TEM image of the sample incubated for five hours (Figure 3.16b), the fiber phase appears together with larger particles, while the density of particles decreases significantly. Such dynamic particle size variation shows the metastable phase to be distinct from amphiphilic micelles, whose sizes are determined by the monomer structure [151, 152]. Although Figure 3.16(a)-(c) show that the particle number decreases and particle size increases, these images on dried grids are not sufficient to accurately quantify the overall mass change of the particle phase. After 48 hours of incubation, the metastable particle phase is depleted and the stable fiber phase dominates the system (Figure 3.16c). This series of images indicates that Aβ(16-22) assembles into fibers with the formation and the subsequent dissolution of the metastable particle phase under these conditions.

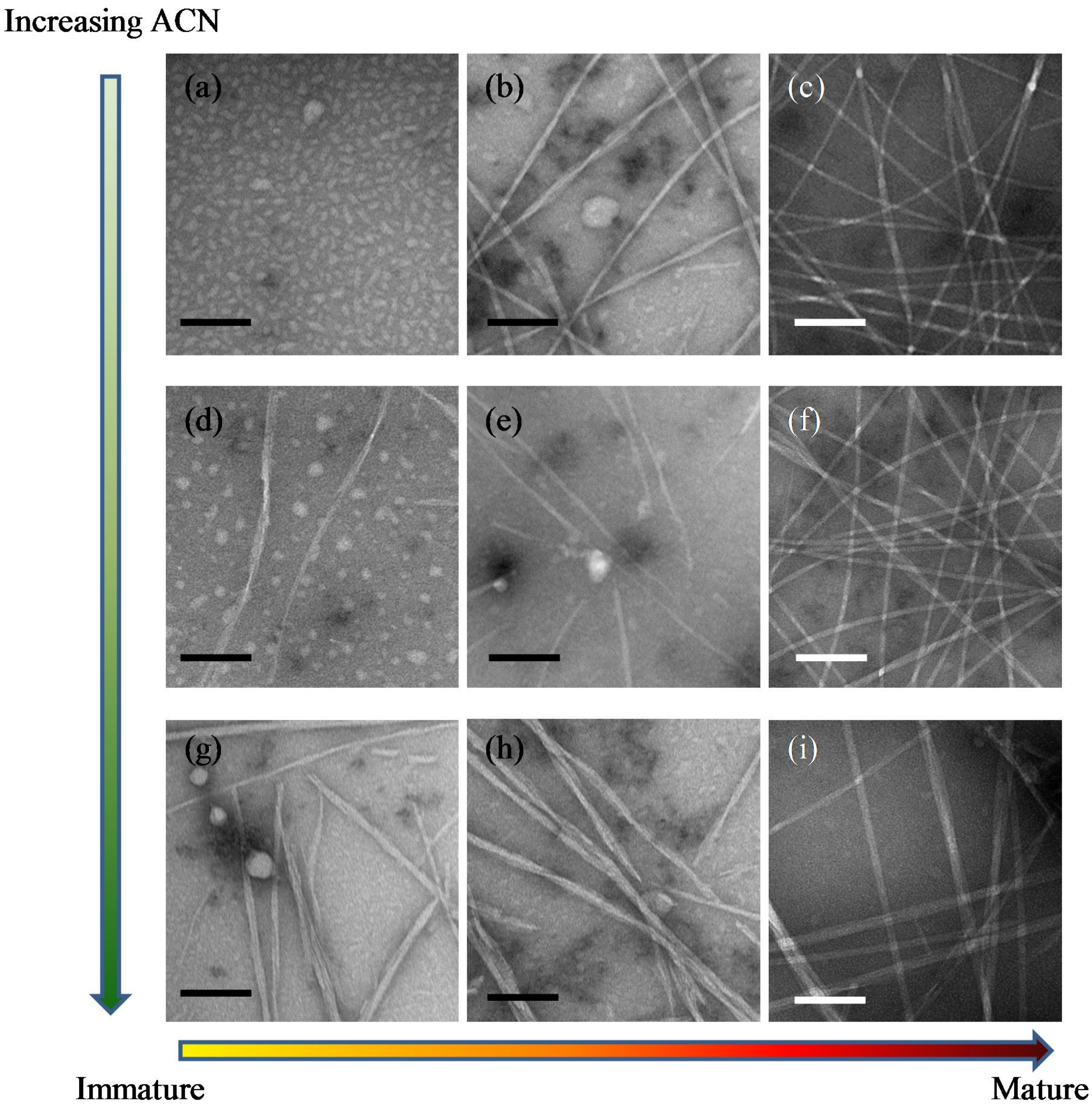

Figure 3.16 – TEM images of 0.5 mM Aβ(16-22) solution in (a-c) 40%, (d-f) 60%, and (g-i) 80% acetonitrile in water. Images are taken after incubation for (a, d, and g) 1 hr, (b, e, and h) 5 hr, and (e, f, and i) 48 hr. Scale bar = 100 nm.

Incubation of Aβ(16-22) in 60% and 80% ACN exhibits particle and assembly phases that are kinetically distinct from 40% ACN. Counter to our original hypothesis, the fiber maturation kinetics are now faster—fibers appear in 60% (Figure 3.16d) and 80% (Figure 3.16g) conditions after 1 hour of incubation, in contrast to the pure particle phase in 40% ACN (Figure 3.16a). The observed particle size at 1 hour also increases with increasing ACN concentration (Figure 3.16a, Figure 3.16d, and Figure 3.16g). Through these EM images, the properties of the dynamic particle phase, the particle size and the particle number, are certainly tunable with different solvent compositions. The final fiber assemblies remain unchanged in all three solvents (Figure 3.16c, Figure 3.16f, and Figure 3.16i) and mature through the same intermediate particle transitions (Figure 3.16b, Figure 3.16e, and Figure 3.16h) [63]. These results are consistent with the morphological evolution of Aβ(16-22) assemblies occurring under a 2SN process.

In addition to the TEM images, circular dichroism (CD) analyses provide time dependent evaluation of the assembly progress. While Thioflavin T (ThT) and Congo red (CR) are typically used to probe peptide assemblies and protein aggregation, they do not bind to Aβ(16-22) assemblies in organic solvent [153, 154], and thus CD provides a valuable option. As shown in Figure 3.17a and Figure 3.18, the ellipticity at 217 nm [51] indicates that β-sheets form immediately. By this analysis, Aβ(16-22) assembles faster in 80% ACN than in 40% ACN (Figure 3.17a), and all assemblies reach equilibrium by 48 hr regardless of the solvent composition, consistent with the TEM analyses (Figure 3.16e, Figure 3.16f, and Figure 3.16i).

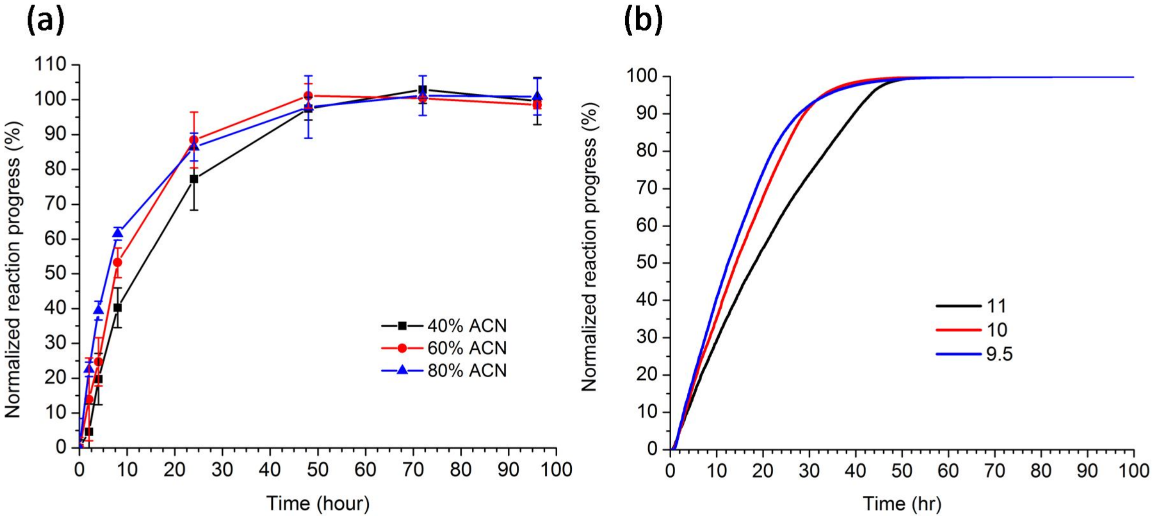

Figure 3.17 – Kinetics of Aβ(16-22) self-assembly (a) measured experimentally as a function of acetonitrile concentration and (b) simulated as a function the Flory-Huggins parameter.(a) The CD intensity at 217 nm is normalized to the average value of the last three time points, when the signatures are stable and the kinetics reaches equilibrium. (b) As χ decreases, Aβ(16-22) assembles faster. Table 3.1 is used for the parameter set with k_{g0} = 200 mole^{-1} sec^{-1}, k_g = 5 mole^{-1} sec^{-1}, $G = 10^{-18.5}$ J, $k_n = 5 \times 10^{-7}$ sec^{-1} and C_0 = 0.5 mM.

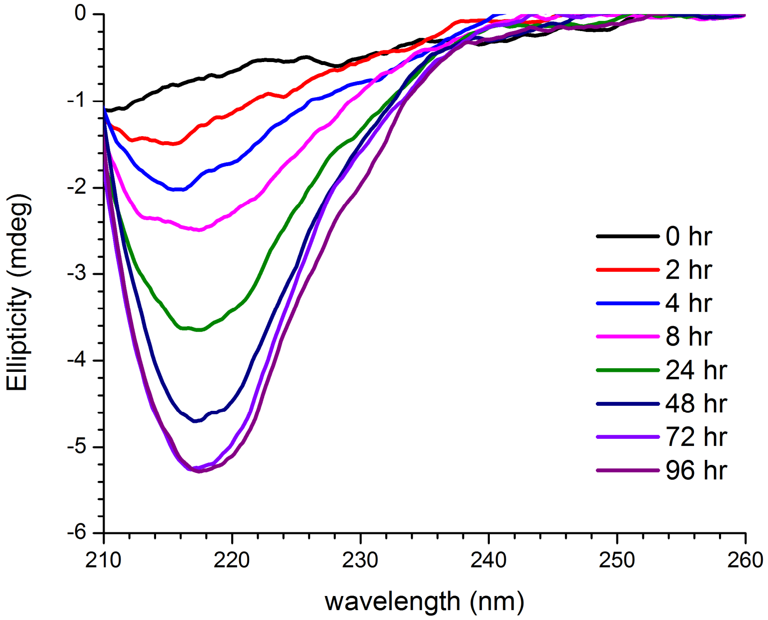

Figure 3.18 – CD spectrum of 0.5 mM Aβ(16-22) in 40% ACN as a function of time. The negative signature at 217 nm develops as the β-sheet content increases, and is used as a probe to calculate the progress of assembly.

FT-IR analyses of the final products verify amide stretches at 1625 and 1690 cm^{-1} (Figure 3.19), consistent with anti-parallel β-sheet amyloid fibers [155] under all conditions. Both TEM images (Figure 3.16) and IR spectra (Figure 3.19) then confirm that the structure of Aβ(16-22) fibers are not ACN concentration-dependent, allowing the Aβ(16-22) assembly kinetics under different ACN concentrations to be compared.

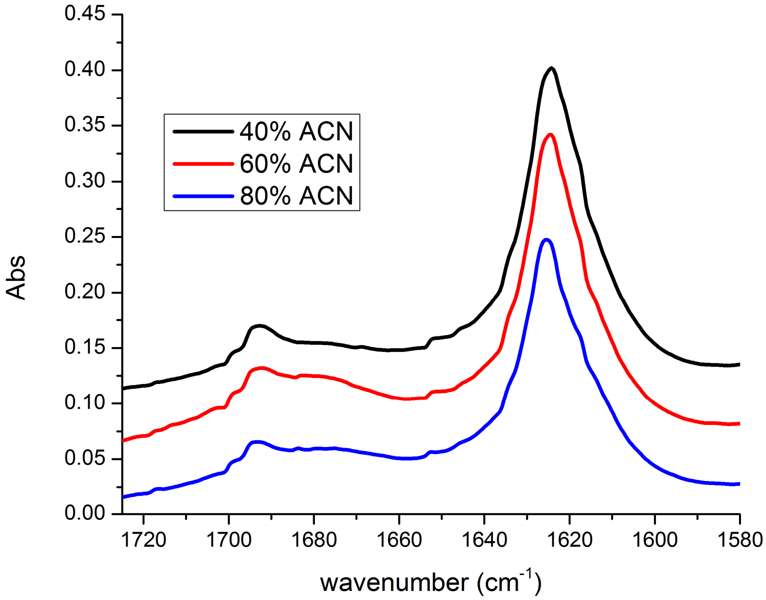

Figure 3.19 – FTIR spectra of Aβ(16-22) assemblies in 40%, 60% and 80%. The overlapping peaks suggest that the normal mode of peptide stacking is not ACN concentration-dependent. All spectra show a peak at 1625 cm^{-1} and another at 1690 cm^{-1}, which are the typical signatures for the anti-parallel β-sheet structure.

The experimentally observed kinetics contradict our original hypothesis that the assembly rate should be slower in more hydrophobic surroundings due to a reduced driving force. While the number of particles decreases and the particle size increases with increasing ACN concentrations (Figure 3.16a, Figure 3.16d, Figure 3.16g, and Figure 3.20), the rate of assembly increases (Figure 3.17a) but the final equilibrium position is not impacted (Figure 3.21). These results are consistent with two key model parameters describing the effect of varying ACN concentration. The model shows that the particle nucleation and particle growth rates are related to supersaturation C_{1*} and thus the Flory-Huggins parameter (χ), while the assembly growth rate depends directly on the rate constant (k_g). The assembly growth constant k_g is selected instead of the stacking energy (ΔG) because the constant final equilibria suggest that stacking energy remains sufficiently large with different ACN concentrations (Figure 3.21). Accordingly, simulations with varied Flory-Huggins parameters (χ) and assembly growth constants (k_g) are specifically compared with the experimental data.

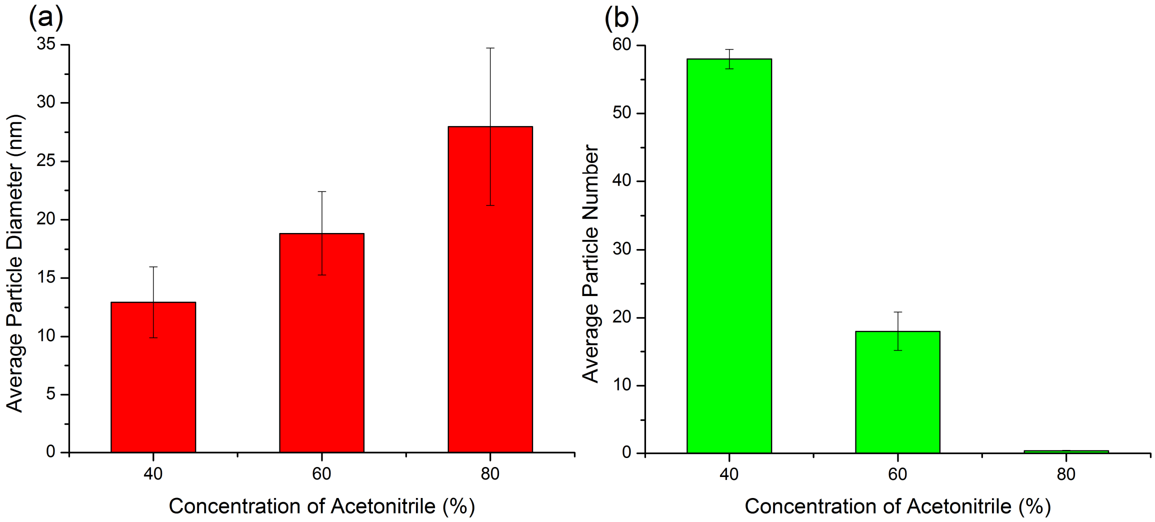

Figure 3.20 – The average particle diameter and concentration of Aβ(16-22) after one-hour of incubation. (a) The average particle diameters and (b) average particle numbers after the Aβ(16-22) solution is incubated for one hour as a function of acetonitrile concentration.Both the particle diameters and numbers are obtained from the average of multiple regions (400 × 400 nm^2) of the TEM images.

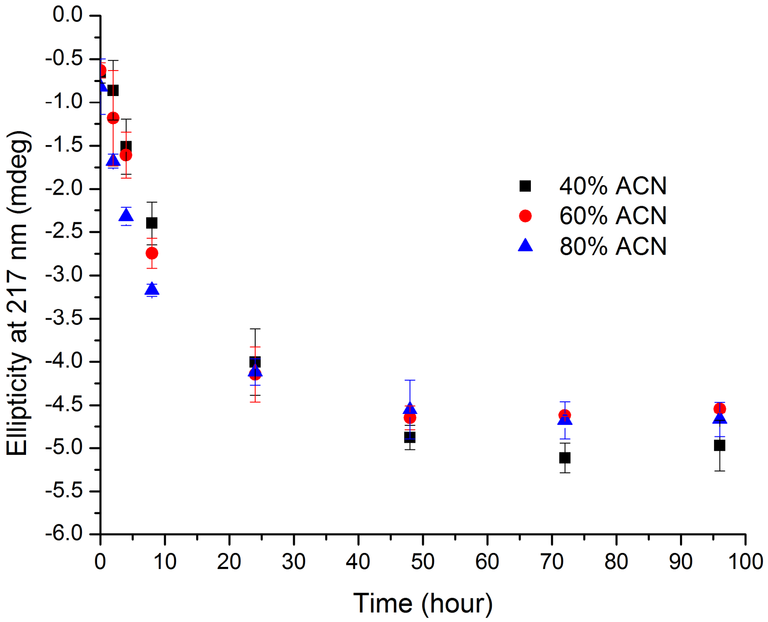

Figure 3.21 – Ellipticity at 217 nm of Aβ(16-22) self-assembly in 40%, 60% and 80% as a function of time. The end time intensity for different solvents is similar, which suggests that the final equilibrium does not depend on the acetonitrile concentration.

The trends of the Aβ(16-22) kinetics are simulated as a function of the Flory-Huggins parameter (χ), and the peptides assemble more quickly as χ decreases (Figure 3.17b). As defined in the model and shown in Figure 3.22, when χ increases, peptide solubility of particles (C_{1*}) decreases, resulting in fewer free peptides in solution available for assembly. By comparing Figure 3.17a and Figure 3.17b, χ and the ACN concentration appear correlated inversely, such that increasing the ACN concentration decreases χ in the system. This observation is consistent with our original hypothesis, since higher solubility is predicted with a lower χ. However, the higher solubility of particles has the effect of speeding up rather than slowing down final assembly.

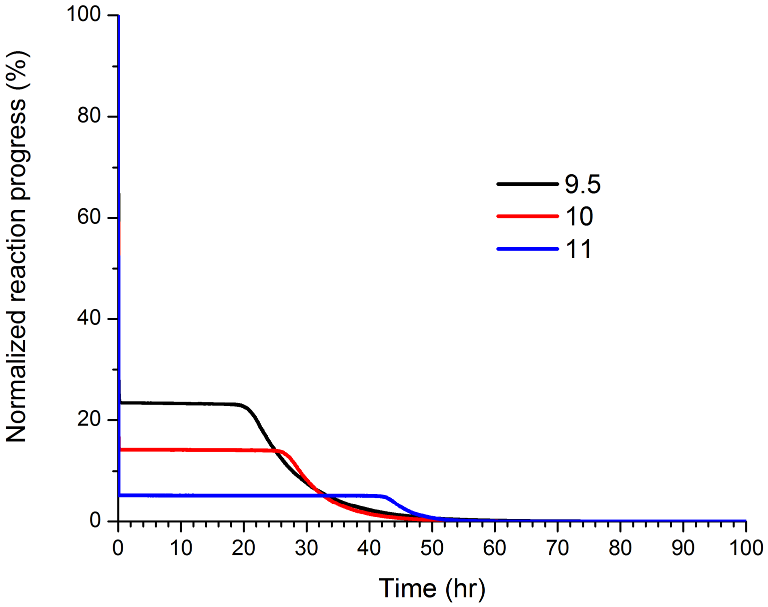

Figure 3.22 – Simulation results of free Aβ(16-22) concentration as a function of Flory-Huggins parameter (χ). As χ increases, the peptide solubility of particle and thus the free peptide concentration decrease. Table 3.1 is used for the parameter set with k_{g0} = 200 mole^{-1} sec^{-1}, k_g = 5 mole^{-1} sec^{-1}, $\Delta G = 10^{-18.5}$ J, $k_n = 5 \times 10^{-7}$ sec^{-1} and C_0 = 0.5 mM.

The observed inverse relationship between ACN concentration and Flory Huggins parameter (χ), as seen in the assembly fraction kinetics, can be further investigated by comparing particle size. The TEM images at one hour show that the particle size increases with increasing ACN concentration (see Figure 3.16a, Figure 3.16d, Figure 3.16g, and Figure 3.20), while the observed particle density decreases. The corresponding model predictions show that particle size increases as χ decreases (Figure 3.23a), with a corresponding decrease in particle density (Figure 3.23b). Thus, the inverse relationship between ACN concentration and Flory Huggins parameter χ is further supported by the comparison of model predictions and experimental data for particle size, as well as assembly fraction.

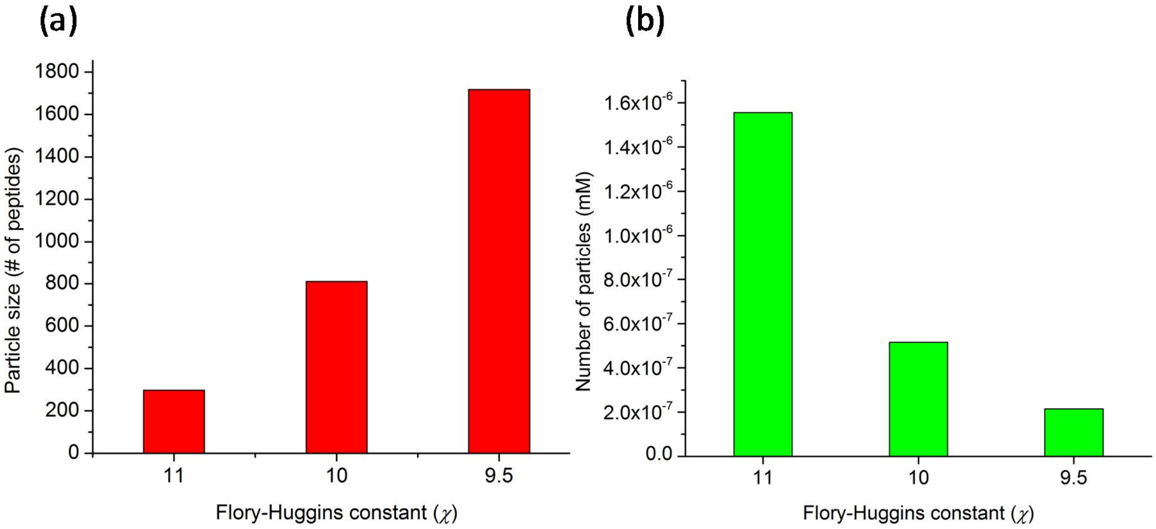

Figure 3.23 – Simulation results of particle size distribution after one hour of reaction as a function of the Flory-Huggins parameter (χ). (a) Particle number concentration and (b) particle concentration are acquired with Table 3.1 as the parameter set with k_{g0} = 200 mole^{-1} sec^{-1}, k_g = 5 mole^{-1} sec^{-1}, $G = 10^{-18.5}$ J, $k_n = 5 \times 10^{-7}$ sec^{-1} and C_0 = 0.5 mM.

The Flory-Huggins parameter is not the only parameter that may affect the assembly kinetics. The assembly growth constant (k_g) increases the rate directly (Figure 3.24), and the simulated results are consistent with the experimental kinetics (Figure 3.17a). However, k_g changes the assembly growth rate only and has no effect on early particle phase evolution, as shown in Figure 3.25. Neither the number of particles nor the average particle size changes with k_g after one hour of incubation. Hence, although k_g may increase with increasing ACN concentrations and accelerate assembly kinetics, it alone is not sufficient to explain the changes in the particle phase in different solvents, as shown in Figure 3.16 and Figure 3.20.

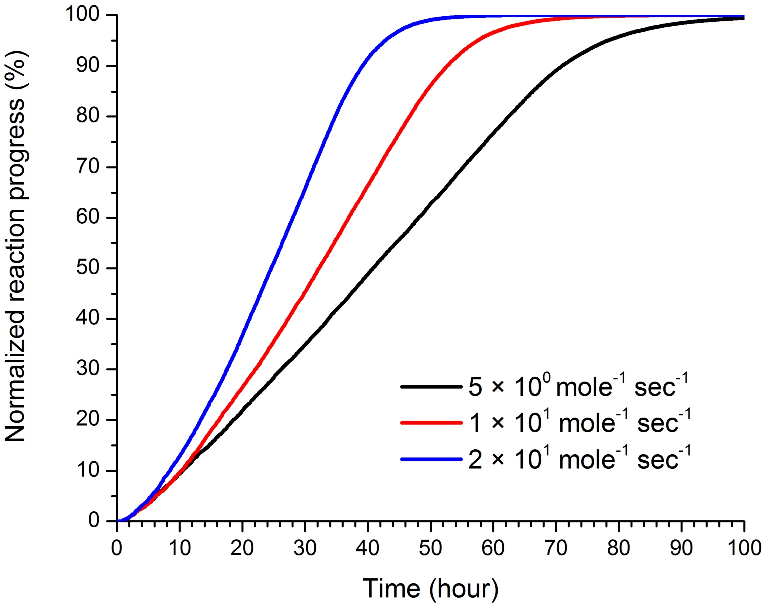

Figure 3.24 – Simulation results of 1 mM Aβ(16-22) self-assembly as a function of the assembly growth constant (k_g). As k_g increases, KLVFFAE assemble faster. The parameters used for this simulation are taken from Table 3.1 with k_{g0} = 200 mole^{-1} sec^{-1}, χ = 10, ΔG = $10^{-18.5}$ J, and C_0 = 0.5 mM.

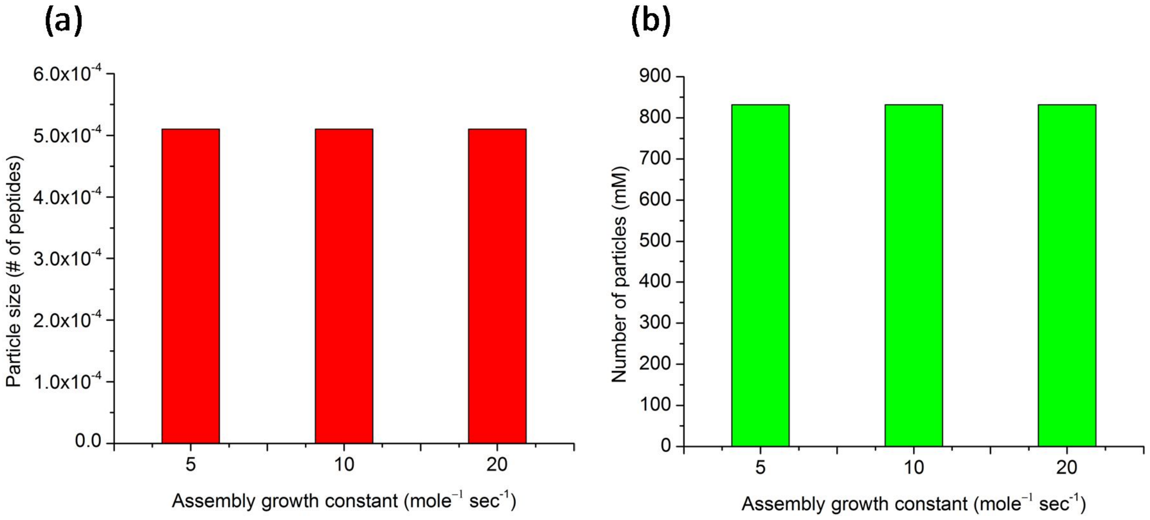

Figure 3.25 – Simulation results of particle size distribution after one hour of reaction as a function of assembly growth constant (k_g). (a) Particle number concentration and (b) particle concentration are acquired with Table 3.1 as the parameter set with k_{g0} = 200 mole^{-1} sec^{-1}, χ = 10, ΔG = $10^{-18.5}$ J, and C_0 = 0.5 mM.

Originally ACN was expected to slow down the assembly kinetics of Aβ(16-22) by increasing the hydrophobicity of the solvent, increasing the solubility of the free peptides, and decreasing the driving force for assembly. However, based on the experimental and calculated results together, ACN does appear to increase the solubility of particles, but does not significantly affect the more stable assemblies. As the solubility of particles increases with the hydrophobicity of the solvent, more peptides become available in the solution phase for assembly growth, which makes the overall kinetics faster.

3.2.4 The Two-Step Nucleation Model for a Polydisperse System

The model with two-step nucleation proposed in Section 3.2.1 is extended for polydisperse systems, in which the peptides may also undergo oligomerization, as discussed in Chapter 2. This extension makes the monodisperse model built in Section 3.2.1 capable of simulating a polydisperse system, while the original monodisperse condition becomes a special case when the oligomerization does not happen.

For the oligomerization, the bond-forming rate constant and the equilibrium constant are assumed to be independent of the oligomer length, and the oligomerization mechanism is assumed to be the same as, for example, the polymerization of citric acid [156]:

$$L_i + L_m \underset{k_1/K_1}{\overset{k_1}{\rightleftharpoons}} L_{i+m} + W \tag{3.11}$$

where L_i and L_m are the *i*- and *m*-th species. Two additional model parameters are added, where k_1 is the forward (condensation) constant and K_1 is the equilibrium constant. When K_1 increases, the oligomers grow longer and the average oligomer length (*j*) increases as well (Figure 3.26). As the reaction is performed is a water-rich environment, the water concentration (*W*) is assumed to be constant.

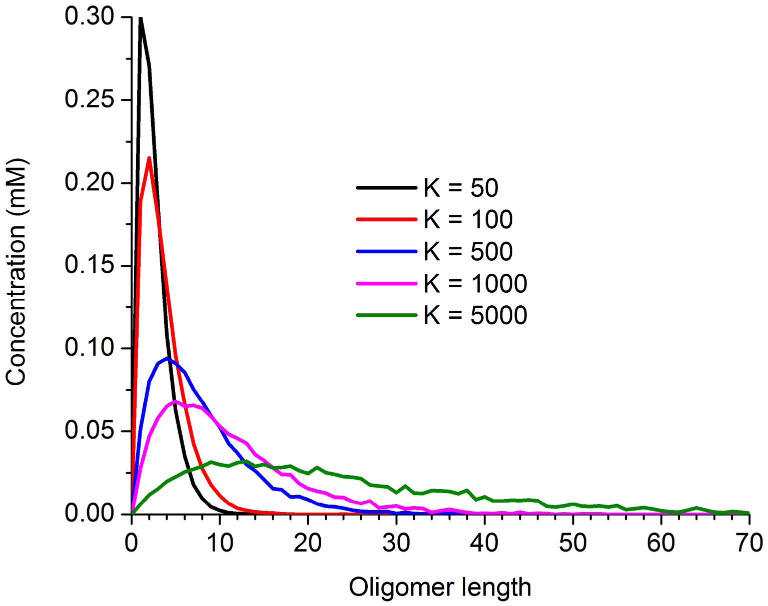

Figure 3.26 – The oligomer distribution as a function of the oligomerization equilibrium constant (K_1).

Similar to the monodisperse model in Section 3.2.1, supersaturation for particles (*S*) is required for calculating the nucleation, growth, and dissolution of the particles. Here, the entire species distribution is considered when determining the supersaturation of the solution, and thus the total number concentration (*C*) for species is required and is expressed as:

$$C = \sum_i L_i \tag{3.12}$$

The solubility (C_{1*}) for the entire system is derived with Flory-Huggins theory as a function of the average oligomer length (j) [138]:

$$C_{1*} = \frac{\rho_{mon} e^{-j\left(1-j-\frac{1}{j}\right)\chi}}{jW_{mon}\left(1-e^{-{}^{-j\left(1-j-\frac{1}{j}\right)\chi}}\right)} \tag{3.13}$$

where ρ_{mon} is the monomer density, W_{mon} is the monomer molecular weight, and χ is the Flory-Huggins parameter. Given the solubility (C_{1*}) and the species concentration (C), the supersaturation of the solution (S) may be calculated as $S = C / C_{1*}$. Similar to the particle nucleation in the dispersion polymerization system [157], here the particle nucleation is assumed to be following the classical nucleation theory [72, 137, 138], with rate equation expressed the same as equation (3.1). The particle nucleus size is again determined by Gibbs-Thompson equation, the same as equation (3.2). Whenever a particle nucleates, n^* oligomers will be randomly selected in the simulation, and they are remove from the solution phase and assemble into the new particle.

The particle then grows if the solution is still supersaturated ($C > C_{1*}$) or dissolves if the solution becomes undersaturated ($C < C_{1*}$). For a bulk diffusion-limited system, the rates of particle growth and dissolution may be combined into a net particle growth rate as:

$$R_{p1} = k_{g0} x^{1/3} \left(C - C_{1*} \right) P\left(x \right) \tag{3.14}$$

where k_{g0} is the particle growth constant and $P(x)$ is the total number of particle with size x. As defined previously for the monodisperse system, x is the mass of the particle in terms of the number of monomers. In other works, the particle growth constant for the polydisperse systems is related to the diffusivity of each individual oligomer [158, 159], which is a function of the oligomer chain length. However in those cases, the supersaturation for the individual oligomer species would be used to calculate their individual particle growth rate, rather than the supersaturation based on the entire oligomer distribution. Hence, to be consistent with the previous assumption for particle nucleation, the supersaturation (C_{1*}) based on the overall oligomer distribution is used, and all oligomers share the same particle growth constant (k_{g0}), as shown in equation (3.14). Once the particle growth happens in the simulation, a random oligomer in the solution will be added into a chosen particle; the probably of a particle to be chosen is proportional to the particle size, x.

The assembly nucleation for a polydisperse system is more complicated than for a monodisperse system, as the numbers of different polymers is particle-wise. Hence, assembly nucleation rate for the i-th polymer in all particles is formulated as:

$$R_{p2} = \sum_{q} k_n P_q \left(i \right) \tag{3.15}$$

where k_n is the assembly nucleation constant, $P_q(i)$ is the number of the i-th polymer in a specific particle q. The assumption of the minimal size of assembly is again made here, and if $P_q(i) < 2$, that i-th polymer will not nucleate as the minimal size is two, as defined in Section 3.2.1. The assembly nucleus in the particle grows by consuming the other polymers with the same length, and is expressed as:

$$R_{p3} = \sum_{q} k_{AG} A_q(i) P_q(i) \tag{3.16}$$

where k_{AG} is the growth constant and $A_q(i)$ is the total number of i-th assemblies in the particle q. The assemblies are exposed to the solution phase and are considered as leaving the particles when their lengths are greater than the particle diameters (L_p), which is determined by equation (3.7).

Finally, the assembly growth (equation (3.17)) and dissolution (equation 3.18) rates for the i-th polymer assemblies are expressed as the following:

$$R_{p4} = \sum_{y} 2k_g A_i(y) L_i \tag{3.17}$$

$$R_{p5} = \sum_{y} k_{bp} \exp\left(\frac{-i\Delta G}{k_b T}\right) B_i(y) A_i(y) \tag{3.18}$$

where k_g is the assembly growth constant and k_{bp} is the assembly dissolution constant in the solution. $A_i(y)$ is the number of assemblies with y i-th polymers (assembly length = y), and L_i is the concentration of the unassembled i-th polymer in solution. ΔG is the binding energy [150] between two monomers in the monomer assemblies. Here the binding energy between assembled polymers is assumed to be a linear function of the length the polymers, and thus for an *i*-th mer the binding energy is $i\Delta G$ [150, 160]. $B_i(y)$ is the number of breakable bonds at the assembly ends. If $y = 2$, there is only one bond breakable and thus $B = 1$; if $y > 2$, both the bonds at the ends are breakable and $B = 2$.

3.2.5 Simulation Results from the Polydisperse Model

To simulate the phase transitions in the DCNs and to understand the interplay between the oligomer distribution (chemical distribution) and different physical phases (physical distribution), the peptide assembly model is extended with peptide oligomerization. The resulting model becomes capable of simulating a polydisperse system, which is similar to the polydisperse DCNs discussed in Chapter 2. Three key thermodynamic parameters are selected to probe the transitions and are sampled with Latin hypercube sampling: the oligomerization equilibrium constant (K_1), the Flory-Huggins constant (χ), and the stacking energy for monomer (1-mer) assemblies (ΔG). The oligomerization equilibrium constant, K_1, is selected to test how the system would evolve with different oligomer bond strength and average oligomer chain length (j). Here the oligomerization rate constant, k_1 in equation 3.11, is set in the fast limit to ensure that the oligomerization kinetics will not limit the change in chemical distribution, and the system reaches its thermodynamic state quickly. For the other two parameters, both χ and ΔG have shown their impact on the final state of the monodisperse model; thus, they are selected again to investigate their impact on both the chemical and the physical distribution of the polydisperse system. The 500 parameter sets are constructed based upon these constants with Latin hypercube samplings, and the remaining parameters and constants are summarized in the following table:

Table 3.4 – Summary of values for the polydisperse peptide assembly model.

Parameter	Description	Units	Value
k_1	Oligomerization rate constant	mole^{-1} sec^{-1}	2×10^{-4}
K_1	Oligomerization equilibrium constant	–	5×10^{4} - 100×10^{4}
p_1	Prefactor for particle nucleation rate	L^{-1} sec^{-1}	10^{13}
p_2	Constant in the exponential part of the particle nucleation rate	–	10^{0}
ρ_{mon}	Density of peptide	g cm^{-3}	1.4
χ	Flory-Huggins parameter	–	4 - 8
W_{mon}	Molecular weight of peptides	g mole^{-1}	410
γ	Interfacial tension between peptide and solvent	g s^{-2}	15
k_b	Boltzmann constant	J K^{-1}	1.38×10^{-23}
T	Temperature	K	300
N_A	Avogadro's number	mole^{-1}	6.023×10^{23}
k_{g0}	Kinetic constant for particle growth	mole^{-1} sec^{-1}	10^{3}
k_n	Assembly nucleation constant	sec^{-1}	10^{-8}
k_{AG}	Assembly growth constant inside the particle	sec^{-1}	10^{-3}
ρ_{par}	Particle density	g cm^{-3}	1.4
k_g	Assembly growth constant in the solution phase	mole^{-1} sec^{-1}	10^{2}
k_{bp}	Prefactor of Assembly dissolution	sec^{-1}	10^{10}
ΔG	Binding energy between peptides in the assembly	J	$10^{-19.7}$–$10^{-18.5}$
C_0	Initial peptide concentration	mM	1
N	Sample size	–	2×10^{5}

The same criteria used to classify the monodisperse model are used again here. Although these criteria do not characterize the chemical distribution of the results, they classify the final phase distribution (physical distribution) on the remaining physical phases. Here, the 500 parameter sets are sufficient to sample all six categories (Table 3.5), with Figure 3.27 showing a representative example for each case.

Table 3.5 – Distribution of simulation results from the polydisperse model with 500 parameter sets.

Case	Number of cases	Percentage (%)
1. Pure solution phase	4	0.8
2. Pure assembly phase	211	42.2
3. Solution-assembly equilibrium	31	6.2
4. Solution-particle equilibrium	18	3.6
5. Three-phase equilibrium	17	3.4
6. Slow kinetics	219	43.8

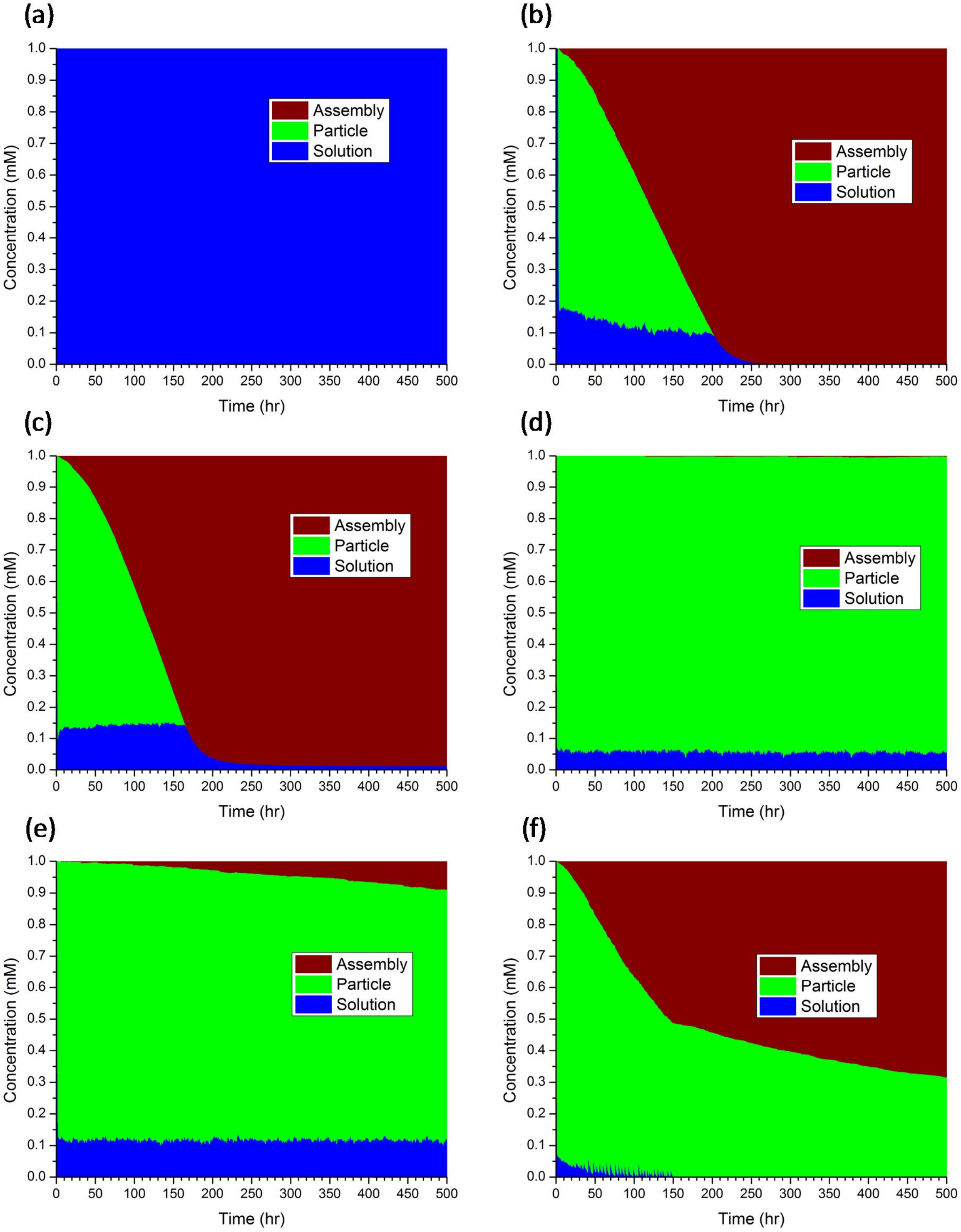

Figure 3.27 – Examples of the simulation results from the polydisperse peptide assembly model of (a) pure solution phase, (b) pure assembly phase, (c) solution-assembly equilibrium, (d) solution-particle equilibrium, (e) three-phase equilibrium, and (f) slow kinetics. The parameters in Table 3.4 are used with: (a) K_1 = 5.1 × 10^4, χ

= 4.63, and $\Delta G = 10^{-18.87}$ J, (b) $K_1 = 9.6 \times 10^5$, $\chi = 4.11$, and $\Delta G = 10^{-18.75}$ J, (c) $K_1 = 3.6 \times 10^5$, $\chi = 6.18$, and $\Delta G = 10^{-19.16}$ J, (d) $K_1 = 6.7 \times 10^5$, $\chi = 7.04$, and $\Delta G = 10^{-19.64}$ J, (e) $K_1 = 5.4 \times 10^5$, $\chi = 5.77$, and $\Delta G = 10^{-19.6}$ J, (f) $K_1 = 6.3 \times 10^5$, $\chi = 5.71$, and $\Delta G = 10^{-19.61}$ J

For a polydisperse system, the particle nucleation will not occur if the oligomers are soluble, due to short average oligomer length (j) or soluble building blocks with a low Flory-Huggins constant, as shown in Figure 3.28. For Case 1, the pure solution phase, the species undergo oligomerization but do not trigger the particle nucleation, as shown in Figure 3.27a and Figure 3.29; the species distribution (chemical distribution) reaches equilibrium within the first few hours, and all the oligomer concentrations remain constant without the perturbation of the other physical phases. On average, the Case 1 results have the both the lowest oligomerization equilibrium constant (K_1) (Figure 3.30) and the lowest Flory-Huggins constant (χ) (Figure 3.31); both of these parameters are critically tied the solubility for particles, as shown in equation (3.13) and Figure 3.28. If the 1-mer is less soluble and thus χ is high enough, even the 1-mers only are sufficient to nucleate the particle phase, as shown with the monodisperse model. If the oligomers grow steadily and thus K_1 is large enough, particle nucleation would be triggered due to the long chain length.

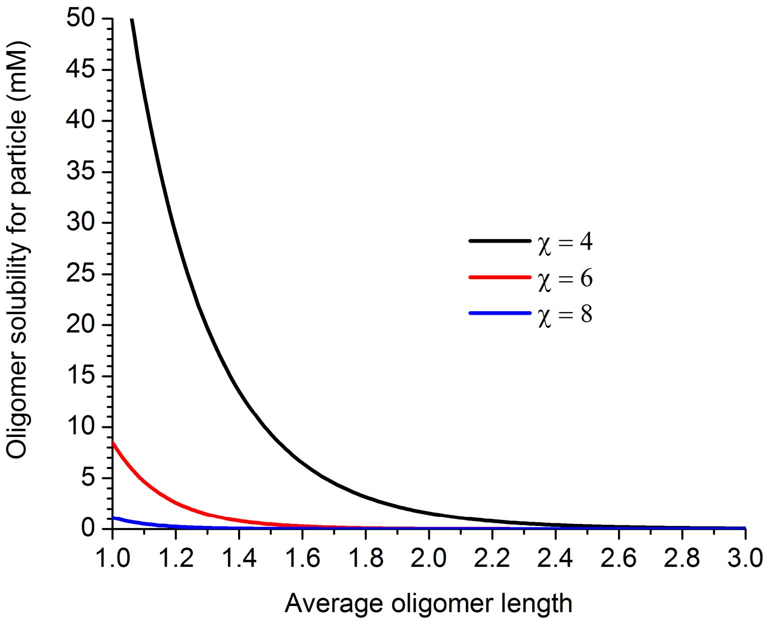

Figure 3.28 – Oligomer solubility for particle as a function of the average oligomer length and the Flory-Huggins constant (χ). The constants in Table 3.4 are used as the peptide properties and equation (3.19) is used to derive the solubility.

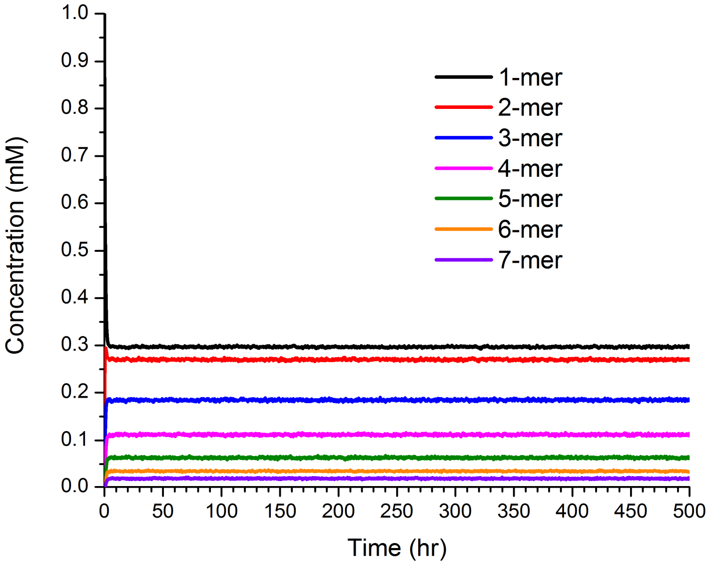

Figure 3.29 – Simulation results of the oligomer distribution when particle phase is not nucleated. The parameters used for this simulation are taken from Table 3.1 with K_1 = 5.1 × 10^4, χ = 4.63, and $\Delta G = 10^{-18.87}$ J.

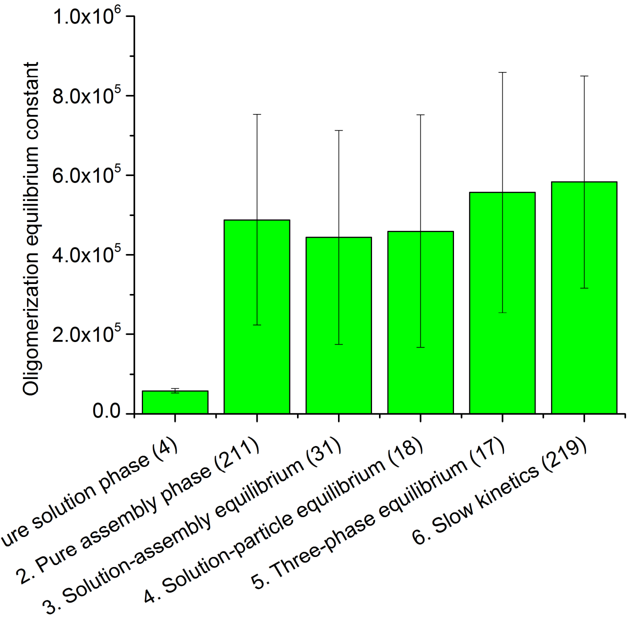

Figure 3.30 – Average oligomer equilibrium constant (K_1) from the simulation results of the polydisperse model. The numbers in the parentheses are the number of results in each individual case.

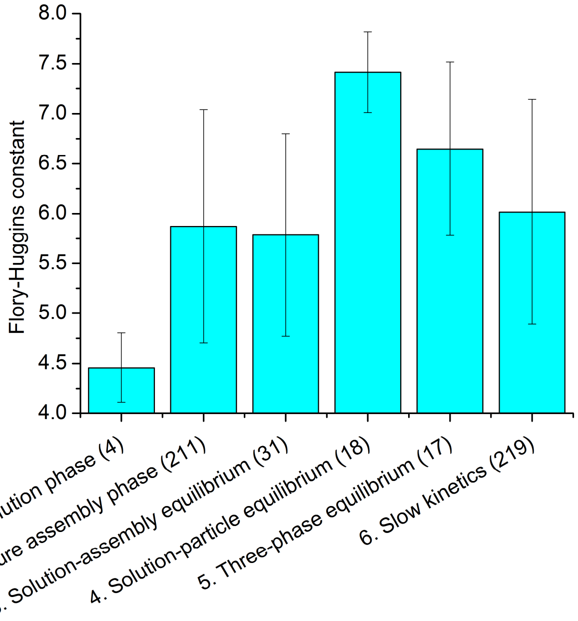

Figure 3.31 – Average Flory-Huggins constant (χ) from the simulation results of the polydisperse model. The numbers in the parentheses are the number of results in each individual case.

Although a smaller K_1 is necessary to keep the oligomers soluble and thus to result in a pure solution phase, in other cases K_1 does not impact the final physical distributions significantly. All the other cases have similar average K_1 with similar large deviation, as shown in Figure 3.30. After particles nucleate in a solution, the phase distribution between the solution, particle and assembly phases is affected more by the Flory-Huggins constant χ (Figure 3.31) and the monomer assembly stacking energy ΔG (Figure 3.32). For example, when the assemblies are stable given a strong ΔG, the assembly dissolution becomes negligible and thus the assembly growth is almost irreversible. The stable assemblies then accumulate while all the other phases diminish, which results in the pure assembly phase (Case 2).

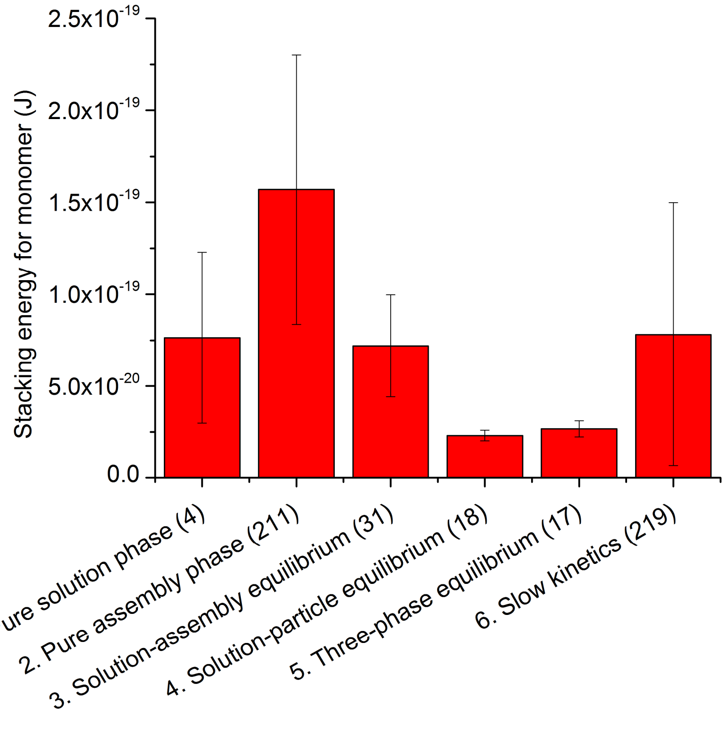

Figure 3.32 – Average stacking energy of monomer assembly (ΔG) from the simulation results of the polydisperse model. The numbers in the parentheses are the number of results in each individual case.

The solution reaches the pure assembly phase (Case 2) when the oligomer assemblies become predominant while all the other phases disappear. As shown in Figure 3.32, Case 2 on average has the highest stacking energy for monomer assembly (ΔG). The strong stacking energy (ΔG) makes the assemblies stable, and the assembly phase is dominated by the assemblies of shorter oligomers, which have the highest assembly mass. Here these assemblies with the highest mass are referred as the major products. Out of 211 results, the major product in Case 2 is either the 1-mer assembly (108 results) or the 2-mer assembly (103 results). Two representative examples are shown in Figure 3.33 for the 1-mer and 2-mer assemblies as the major product. Other than these major products, multiple types of oligomer assemblies are observed in each individual simulation, as shown in Figure 3.33. Unlike the DCNs discussed in Chapter 2, which both yield 3-mer assemblies only, the coexistence of multiple types of assemblies suggests that there are other conditions to be considered to generate a predominant assembly. Also, in the Case 2 simulation results, there is no 3-mer assembly observed as the major product. The lack of 3-mer as the major product may be connected with the stable 1-mer and 2-mer assemblies, which consumed the possible resource (free 1-mer and 2-mer) to make the 3-mer.

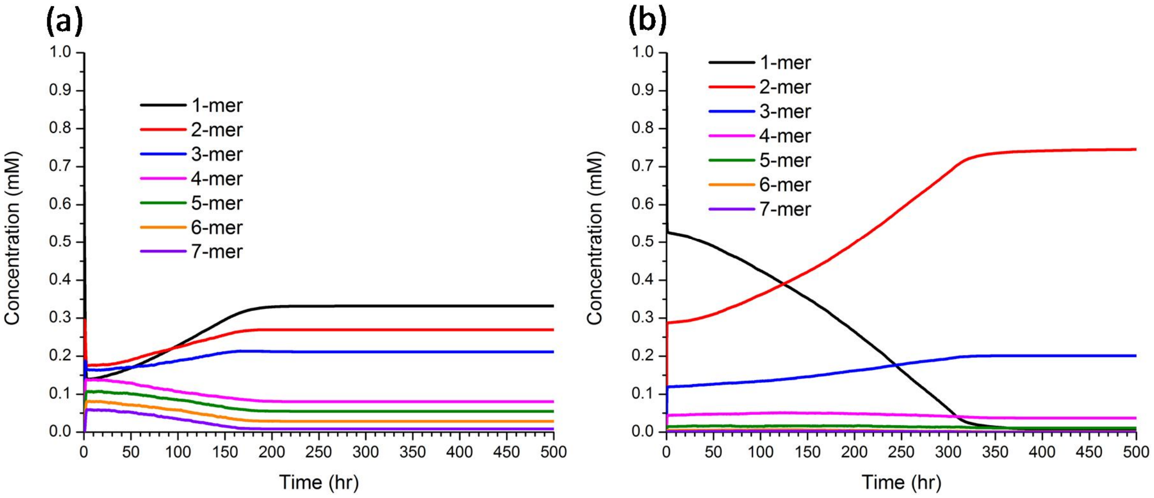

Figure 3.33 – Simulation results of the oligomer distribution for the pure assembly phase. (a) The major product is 1-mer assembly. (b) The major product is 2-mer assembly. The parameters used for this simulation are taken from Table 3.1 with (a) K_1 = 6.3 × 10^5, χ = 4.16, and ΔG = $10^{-18.65}$ J, and (b) K_1 = 7.1 × 10^5, χ = 7.94, and ΔG = $10^{-19.06}$ J.

Although 3-mer assembly is not the major product in any Case 2 results, it is selected in some of the Case 3 results, the solution-assembly equilibrium (Figure 3.27c). As shown in Figure 3.34 as a representative example, the 3-mer assembly is found as the major product when the solution and the assembly phase reach equilibrium (Figure 3.34a), while other minor assembly products still present as well (Figure 3.34b and Figure 3.34c).

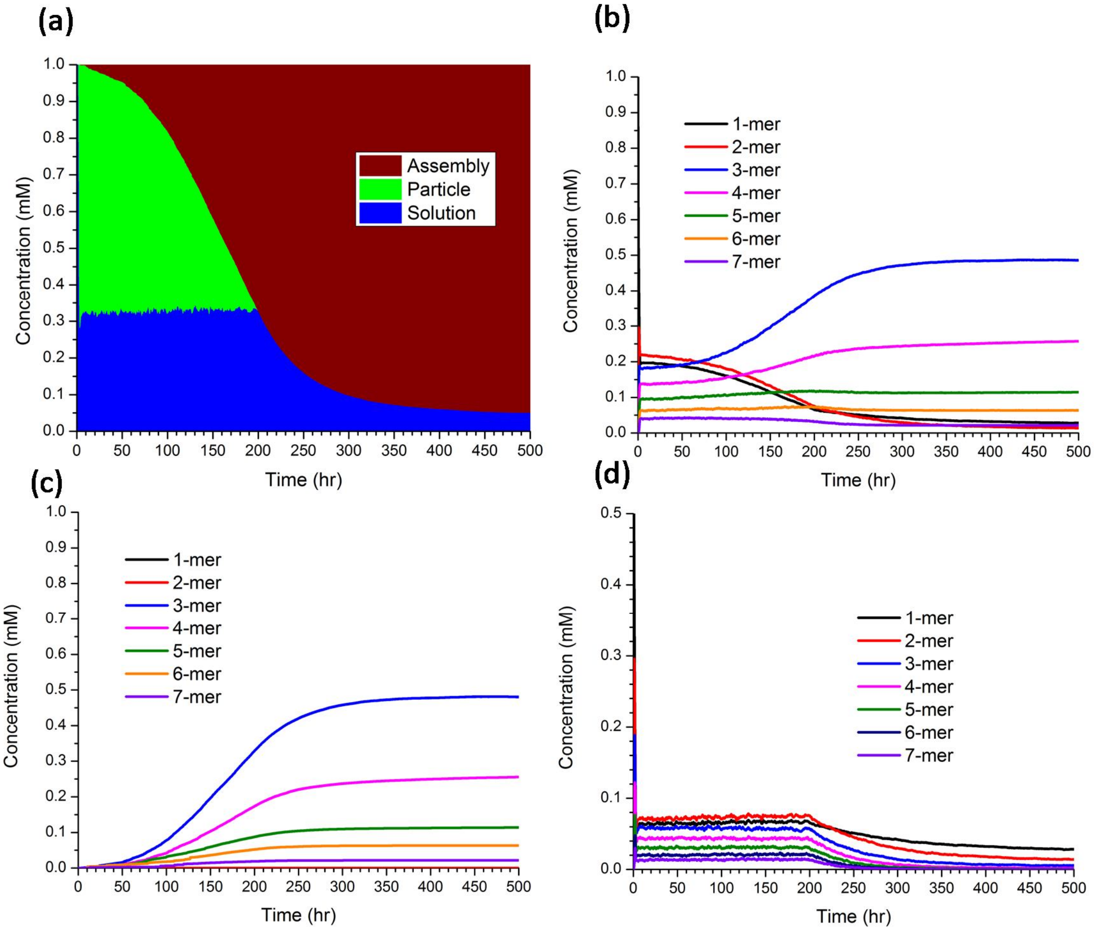

Figure 3.34 – A representative simulation result for the solution-assembly equilibrium. The parameters used for this simulation are taken from Table 3.1 with $K_1 = 2.9 \times 10^5$, $\chi = 4.67$, and $\Delta G = 10^{-19.06}$ J. (a) The distribution of molecules in the solution, particle, and assembly phases. (b) The overall oligomer distribution, (c) the oligomer distribution in the assembly phase, and (d) the oligomer distribution in the solution phase.

For Case 3, some results show solution-assembly equilibrium at the end time, but their chemical distributions are not stationary. In Figure 3.35 as a representative example, the overall oligomer concentration distribution has not stabilized after 500 hr (Figure 3.35b), as the 1-mer assembly is dissolving (Figure 3.35c and Figure 3.35d). The dissolved 1-mer molecules do not accumulate in the solution phase, as the solution phase concentration does not change (Figure 3.35d). Rather, the 1-mer undergoes oligomerization and then is consumed by other assemblies. This suggests that, for a polydisperse system, the equilibrium for the physical distribution (phase distribution) may not guarantee the equilibrium in the chemical distribution (the oligomer concentrations).

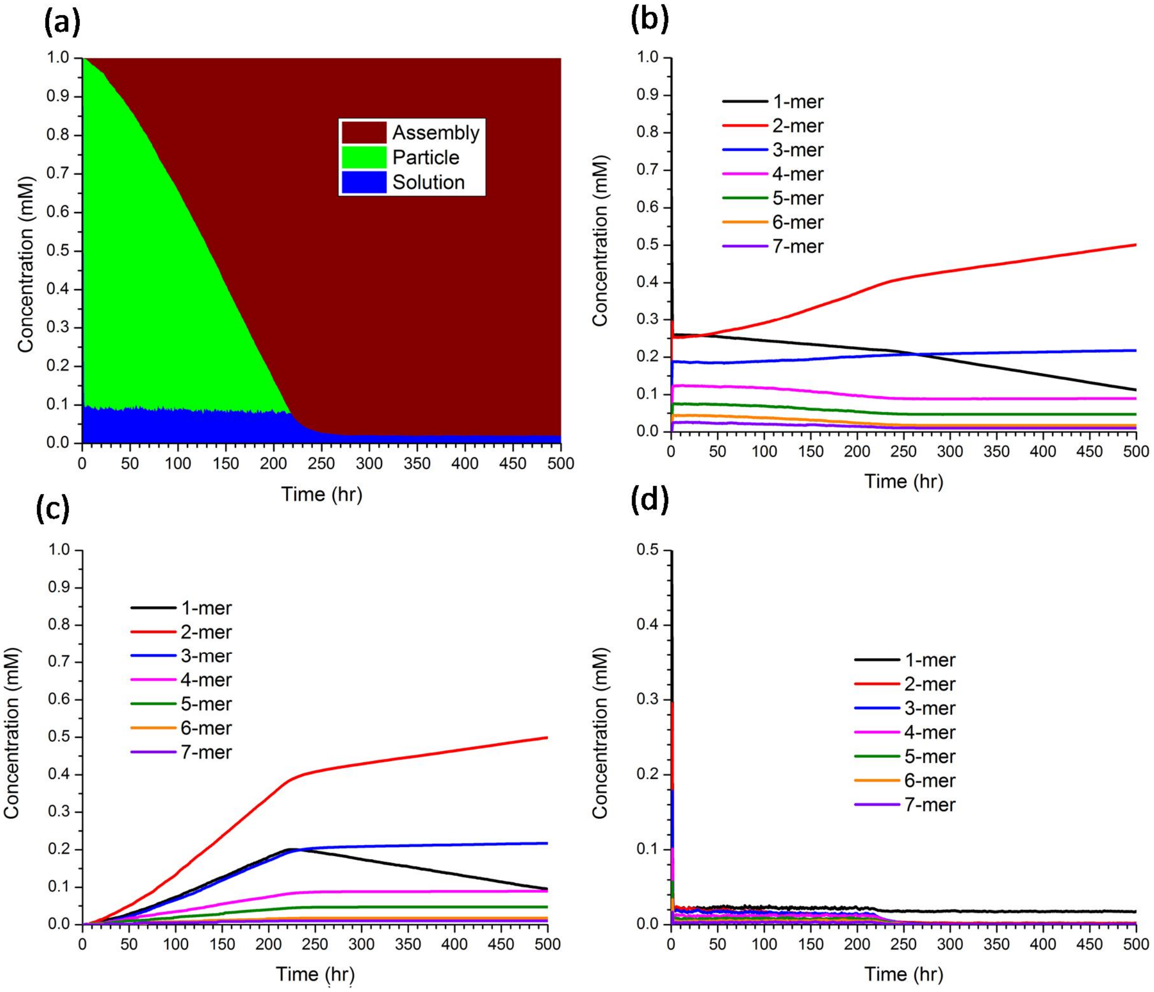

Figure 3.35 – A representative simulation result of solution-assembly equilibrium with stationary phase distribution but active oligomer exchange. The parameters used for this simulation are taken from Table 3.1 with $K_1 = 7.9 \times 10^5$, $\chi = 5.6$, and $\Delta G = 10^{-18.90}$ J. (a) The distribution of molecules in the solution, particle, and assembly phases. (b) The overall oligomer distribution, (c) the oligomer distribution in the assembly phase, and (d) the oligomer distribution in the solution phase.

For the monodisperse model, when the assemblies are not stable, they dissolve upon exposure to the solution, and remaining the solution phase and the particle phase reach equilibrium (Case 4). For this solution-particle equilibrium, the results on average have the highest Flory-Huggins constant (χ) to maintain the particle phase (Figure 3.31), while they have the lowest stacking energy for the assemblies to destabilize the assemblies (Figure 3.32).

Case 5 is the three-phase equilibrium, the coexistence of all solution, particle and assembly phases. Similar to the changing chemical distribution in Case 3, some results of Case 5 have not reached the steady state for the oligomer distribution (Figure 3.36a and Figure 3.36b). The assemblies of the shorter oligomers (1-mer, 2-mer or 3-mer) dissolve (Figure 3.36c), while the longer oligomers grow slowly due to the scarce free long oligomers (Figure 3.36d); oligomer breakage and incorporation into the particles may cause a low concentration for the longer oligomers, and thus delay the assembly growth. The slow assembly growth and the persistent particle and solution phases, together result in a three-phase coexistence.

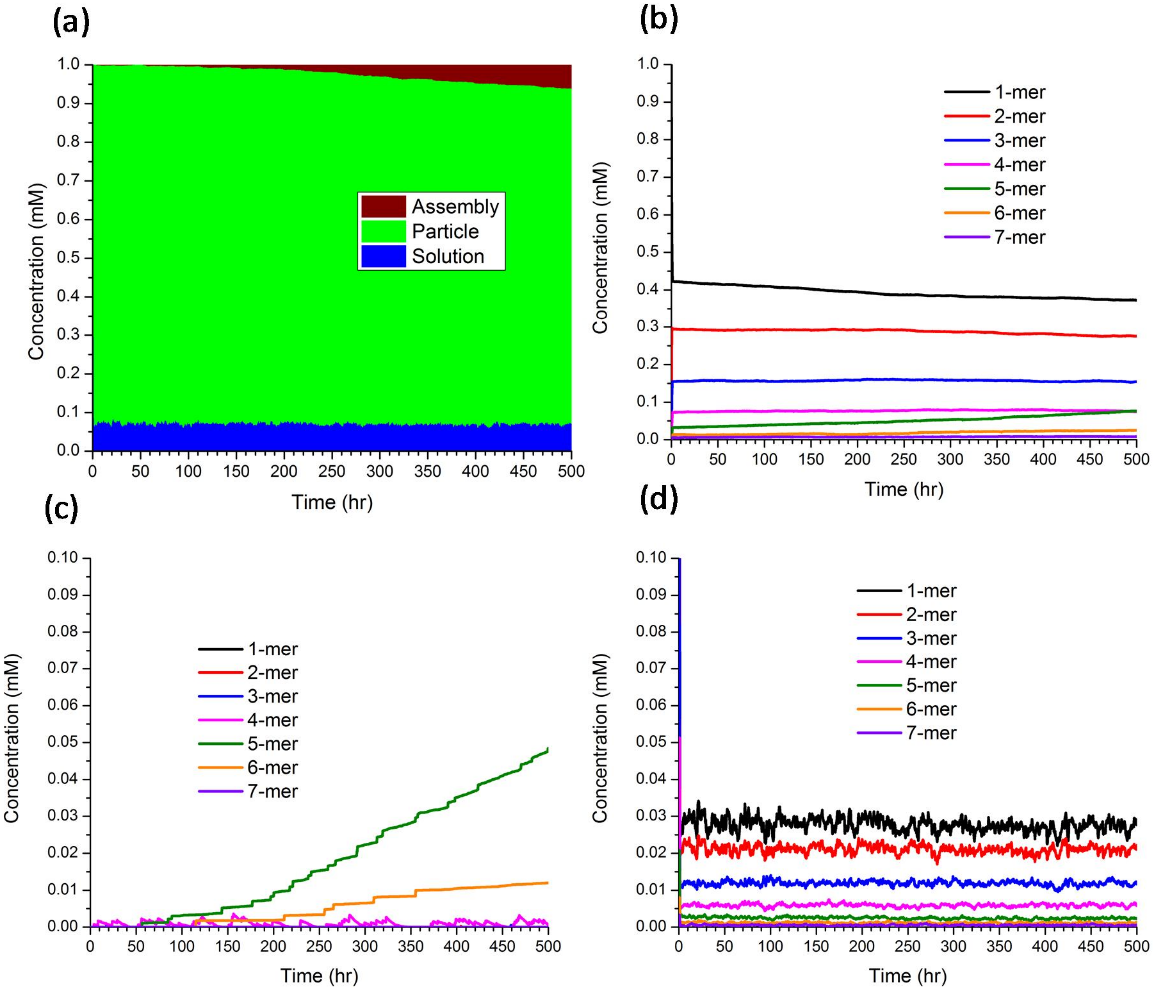

Figure 3.36 – A representative simulation result of three-phase equilibrium. The parameters used for this simulation are taken from Table 3.1 with $K_1 = 4.8 \times 10^5$, $\chi = 7.1$, and $\Delta G = 10^{-19.49}$ J. (a) The physical distribution of molecules in the solution, particle, and assembly phases. (b) The overall oligomer distribution, (c) the oligomer distribution in the assembly phase, and (d) the oligomer distribution in the solution phase.

The above discussion suggests that the steady state of the physical distribution may be achieved without the constant chemical distribution. However, for the slow kinetics (Case 6), some results show the opposite: the stationary chemical distribution with a changing physical distribution. Case 6 is classified as slow kinetics because its physical distribution is still changing after the reaction time (Figure 3.37). However, the physical distribution may continue to change without further chemical exchange, as shown in Figure 3.37a and Figure 3.37b as a representative example. The dissolved oligomers from the particles do not undergo reversible oligomerization, but are incorporated into the assemblies directly.

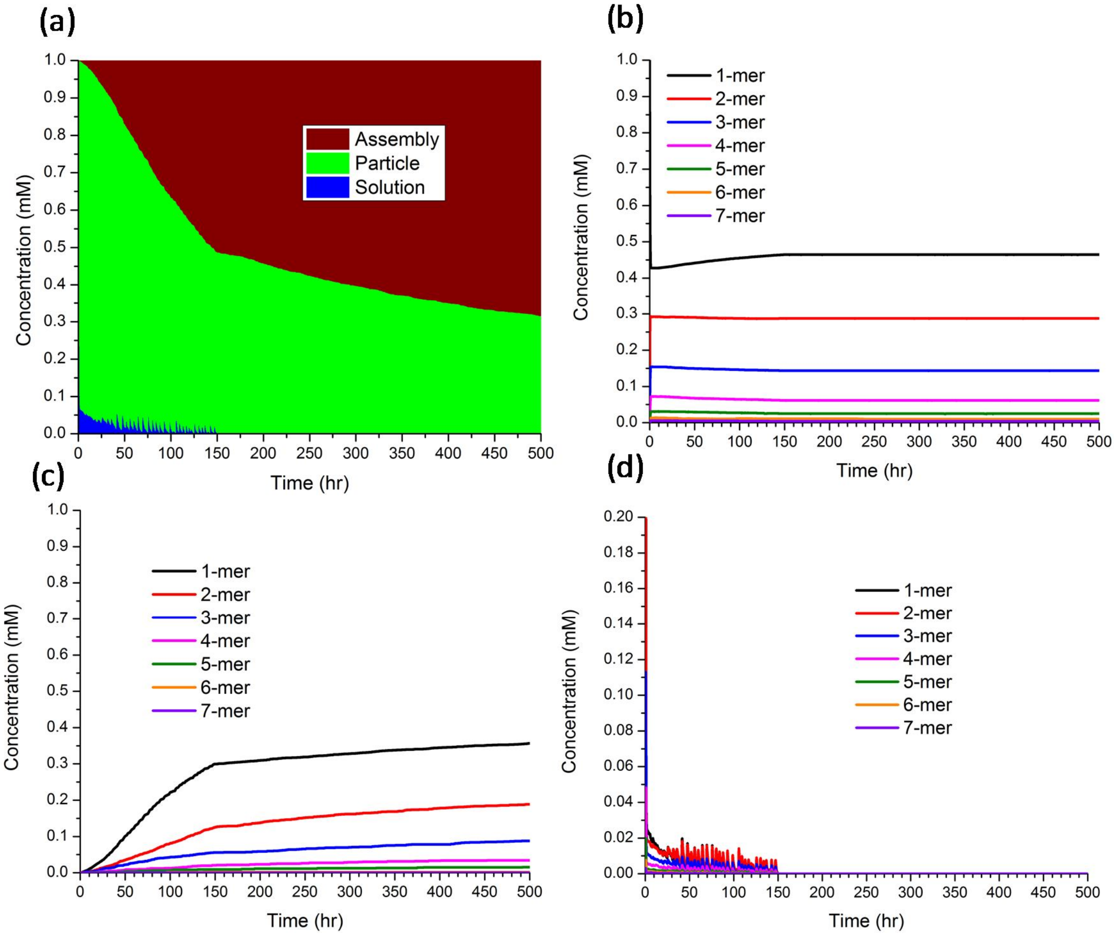

Figure 3.37 – A representative simulation result of slow kinetics with constant chemical distribution. The parameters used for this simulation are taken from Table 3.1 with $K_1 = 6.3 \times 10^5$, $\chi = 7.07$, and $\Delta G = 10^{-18.50}$ J. (a) The physical distribution of molecules in the solution, particle, and assembly phases. (b) The overall oligomer distribution, (c) the oligomer distribution in the assembly phase, and (d) the oligomer distribution in the solution phase.

The above results suggest that, when the peptides are capable of both chemical exchange and physical transitions, the equilibrium of the chemical and the physical distribution can be independent of each other (Figure 3.35 and Figure 3.37). The final state of the chemical distribution is summarized in Table 3.6 in terms of different physical states. The chemical distribution may achieve equilibrium when the assembly phase is stationary (Case 2 and Case 4), but the constant physical distribution may not guarantee a stationary chemical distribution (Case 3, 5, and 6).

Table 3.6 – The equilibrium of the chemical distribution as a function of the final physical distribution. Case 2 is pure assembly phase, Case 3 is solution-assembly equilibrium, Case 4 is solution-particle equilibrium, Case 5 is three-phase equilibrium, and Case 6 is slow kinetics.

Chemical distribution	Case 2	Case 3	Case 4	Case 5	Case 6
In equilibrium	211	8	18	15	40
Not in equilibrium	0	23	0	2	179
Total	211	31	18	17	219

Also, for all of the results with assembly phase observed, they all contain multiple types of assemblies, which contradict with the experimental observation about the *NF*-CHO and *NFF*-CHO DCNs. For both DCNs, only 3-mer (trimer) assemblies are observed, which suggests that the stacking energy of the oligomer assemblies may not be linear as assumed in equation (3.18) [161]. To reproduce the homogeneous 3-mer assemblies observed in the DCNs, the stacking energy (ΔG) for the 3-mer assembly is assumed to be stronger than that for the other assemblies, and thus only 3-mer assembly is stable enough to accumulate. As shown in Figure 3.38, the replication of 3-mer is achieved with the stable 3-mer assemblies, and the longer oligomers (4-mer and 5-mer) are absence after the 3-mer starts to accumulate, similar to the experimental results.

However, although the simulation shows that the system may be dominated by the 3-mer assemblies, why the 3-mer assemblies are more stable than the others remains to be clarified. One possibility is that

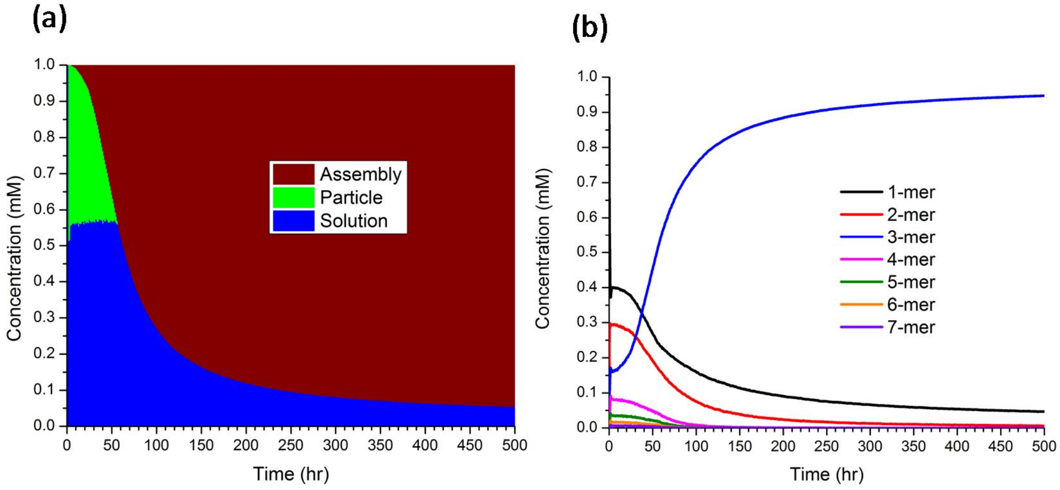

Figure 3.38 – Simulation results of 3-mer assemblies as the dominating product. The parameters used for this simulation are taken from Table 3.1 with K_1 = 5 × 10^5, χ = 6, ΔG = $10^{-18.00}$ J for 3-mer and ΔG = $10^{-20.00}$ J for the rest oligomers. (a) The physical distribution of molecules in the solution, particle, and assembly phases. (b) The overall oligomer distribution,

3.3 Conclusion

A peptide assembly model using two-step nucleation is constructed to describe the maturation of assemblies via a metastable particle phase. The model predicts different phase transitions and equilibria are accessible by varying the Flory-Huggins parameter (χ), the particle growth constant (k_{g0}), and the peptide stacking energy in the assemblies (ΔG). The model shows that the solubilities for particles and assemblies, which are significantly influenced by χ and ΔG, determine the final thermodynamic states of the system. When the solubility of the particles is higher than that of assemblies, the particle phase dissolves while the assembly phase either dominates the system or exists in equilibrium with free peptides. When the solubility of the assemblies is higher, the particle phase equilibrates with the solution phase and no assemblies are apparent. The final state of the system can therefore be rationally predicted given the values of the solubility of the particles and assemblies, but slow particle dissolution can delay maturation of the assembly phase.

This model was validated with Aβ(16-22), the nucleation core of Amyloid β peptide of Alzheimer's disease, by evaluating assembly kinetics in solvents with different acetonitrile concentrations. TEM analyses show that the particle number increases and particle size decreases with decreasing acetonitrile concentration. Both TEM and CD data confirm that the kinetics of Aβ(16-22) assembly are faster with increasing acetonitrile, while the solvent composition does not significantly impact the final equilibrium. The model predicts the more hydrophobic acetonitrile stabilizes the free Aβ(16-22) peptides and increases its solubility of particles, which leaves more peptides for assembly growth and thus drives the reactions faster. These results can now be used to evaluate the assembly of more complex systems, such as recently published dynamic chemical networks that juxtapose chemical polymerization with physical phase transitions to create functional assemblies.

Finally, this model is extended for a polydisperse system with peptides able to perform oligomerization. The same six categories are considered, with 500 different parameter sets, which contain different values of oligomerization equilibrium constant (K_1), Flory-Huggins constant (χ), and the stacking energy for the monomer assembly (ΔG). The simulation results show that, for the polydisperse system, the steady states of the chemical and the physical distribution may be independent. Neither is a necessity for the other, and each could be achieved without the other. Also, none of the 500 results returns a single type dominating assembly. For the results with assemblies, they all contain multiple types of assemblies, and this is not consistent with the previous experimental observation for the DCNs. However, extra restriction on the stability of the assemblies may force the trimer assembly to dominate, as observed in the experimental system.

3.4 Methods

3.4.1 Monte Carlo Simulation Algorithm

The kinetic Monte Carlo algorithm from Wang et al.[162] is used to follow the formation and size evolution of each individual particle and assembly, together with the peptide concentration in the solution phase. Rather than simulating the entire solution, only a portion of it is followed, and the simulated sample volume (*V*) is calculated as:

$$V = \frac{N}{C_0 N_A} \tag{3.11}$$

where N is the number of peptides to be observed and C_0 is the initial peptide concentration. The algorithm selects reactions to occur based on their individual stochastic rates, which are used to calculate the probabilities of the reactions. The stochastic rate (W_i) for the *i*-th reaction is calculated as:

$$W_i = \frac{R_i}{V^{\alpha(i)}} \tag{3.12}$$

where $\alpha(i)$ is determined as a function of the reaction order. For zeroth- and first-order reactions (R_1, R_3, R_4, and R_6) $\alpha(i) = 0$, while for the second order reactions (R_2 and R_5) $\alpha(i) = 1$. The individual probability (P_i) for the i-th reaction is then calculated based on the total stochastic rate:

$$P_i = \frac{W_i}{\sum_i W_i} \tag{3.13}$$

Then, the next reaction, μ, is selected by a random number (q_1) uniformly distributed between 0 and 1:

$$\sum_{i=1}^{\mu-1} P_i < q_1 < \sum_{i=1}^{\mu} P_i \tag{3.14}$$

Finally, the time interval τ for the next reaction to happen is determined by:

$$\tau = \frac{1}{\sum_i W_i} \ln\left(\frac{1}{q_2}\right) \tag{3.15}$$

where q_2 is another random number uniformly distributed between 0 and 1. Both q_1 and q_2 are sampled with the **`rand`** function in MATLAB with the current time as the seed. The execution of this Monte Carlo algorithm in MATLAB 2012a (The Mathworks, Inc) follows Wang's procedure [162], and the parameters and variables for equation (3.1) - (3.11) are listed in Table 3.1.

3.4.2 Criteria to Classify the Simulation Results.

The results from the Monte Carlo simulations are classified into different cases based on the remaining phases at the end. A set of criteria is established to categorize the results into different cases (Table 3.7). At the end time, a phase is considered as existing if it has a concentration greater than 1% of the total concentration, in order to ignore the misleading fluctuation of concentration associated with stochastic fluctuations. The concentration change for the last 5% of the reaction time (five hours) is used to define if multiple phases reach equilibrium with each other. The finite difference method is utilized to calculate the slopes of the concentration change with 0.05 hours as the interval:

$$\text{slope} = \frac{B_{j+1} - B_j}{\Delta t} \tag{3.16}$$

where B_{j+1} and B_j are the peptide concentration in a phase at time t_{j+1} and t_j, respectively. Here $\Delta t = t_{j+1} - t_j = 0.05\ \text{hour}$.

The **`ttest`** function in MATLAB 2012a (The Mathworks, Inc) is carried out to check if the slopes have an average of 0. If the slope average is zero, the result is regarded as a stable equilibrium; otherwise the distribution is still changing and the kinetics are considered "slow," as it is not finished with the reaction time. Table 3.7 includes the final states and the criteria.

Table 3.7 – Criteria to classify the simulation results.

Cases	Criteria
1. Solution wins	The maximal particle concentration remains 0 within the reaction time.
2. Assembly wins	The final solution and particle phase concentration are both less than 1% of the total concentration.
3. Solution-assembly equilibrium	The final solution and assembly phase are both above 1% of the total concentration. The final particle phase is less than 1% of the total concentration. The assembly concentration does not change for the last 5% of the reaction time.
4. Solution-particle equilibrium	The final solution and particle phase are both above 1% of the total concentration. The final assembly phase is less than 1% of the total concentration. The particle concentration does not change for the last 5% of the reaction time.
5. Three-phase equilibrium	All of the final solution, particle and assembly phase are above 1% of the total concentration. All concentration does not change for the last 5% of the reaction time.
6. Slow kinetics	Either solution phase and/or particle phase are above 1% of the total concentration. Concentrations are still changing within the last 5% of the reaction time.

3.4.3 Aβ(16-22) Solution Preparation

Aβ(16-22) peptides are synthesized using solid phase peptide synthesis, and the synthetic products are purified with high-performance liquid chromatography (HPLC), using a water/acetonitrile gradient with 0.1% trifluoroacetic acid in a C-18 reverse phase column

[21]. After lyophilization, the purified peptide is stored at -20 °C for later use. To disassemble the preexisting assemblies, the peptide powder is dissolved in hexafluoroisopropanol (HFIP) to a concentration of 1 mg/mL and the mixture is sonicated for 2 hours in a bath sonicator. HFIP is removed with a Labconco CentriVap Concentrator-7970010, and the resulting peptide film is placed in a desiccator overnight to remove the remaining HFIP. Immediately prior to conducting an assembly experiment, the peptide film is suspended with HFIP and sonicated for 15 minutes. The sonicated solution is diluted with 40%, 60% or 80% (v:v) acetonitrile in water, to obtain a final HFIP concentration of 1% by volume with the desired peptide concentration. The peptide solution is incubated at 22 °C and aliquots of the solution are removed at predetermined times for measurement of assembly kinetics.

3.4.4 Kinetics of Aβ(16-22) Assembly Measured with CD, TEM and FTIR

Circular dichroism (CD) is used extensively to determine β-sheet content of the peptide solutions. Here a volume of 23 μL of peptide solution is loaded into a demountable window cell with a 0.1 mm path length. The ellipticity is obtained at wavelengths from 185–260 nm under a resolution of 0.2 nm with a bandwidth of 2 nm at a scanning rate of 200 nm/min using a Jasco J-810 spectropolarimeter. After background correction, the spectrum from the average of three scans is saved and the resulting ellipticity at 217 nm is used as the indicator of β-sheet content.

The morphologies of the particles and assemblies are observed with transmission electron microscopy (TEM). To deposit the assemblies and particles onto the grid, the peptide solution is loaded onto the copper grid for one min, and the excess peptide solution is wicked away with filter paper. The resulting grid is negatively stained with 1.5-wt % methylamine tungstate solution for three minutes before the stain solution is wicked away with filter paper. The grid is then stored in a vacuum desiccator overnight to remove the remaining liquid. TEM images are recorded with a Hitachi H-7500 transmission electron microscope. The particle size distribution is obtained by analyzing the TEM images with ImageJ 1.48v (National Institutes of Health, USA).

Once the Aβ(16-22) assemblies are matured and the CD signatures of the solution do not change significantly with time, FTIR is used to probe the final peptide strand arrangement. Peptide solutions with a volume of 8 μL are loaded onto the diamond chip of a JASCO FT/IR-4100. The absorbance spectra are recorded against a background spectrum at a wavenumber range of 1000 to 3600 cm^{-1} at room temperature for 250 scans.

CHAPTER 4.

KINETIC AND STRUCTURAL EVOLUTION OF PH-SENSITIVE PEPTIDE ASSEMBLIES

4.1 Introduction

In the last chapter, Aβ(16-22) is studied as the model for the kinetics of peptide assembly with two-step nucleation. In this chapter, the assembly of Aβ(16-22) is investigated further. The assemblies of Aβ(16-22) are pH-sensitive. At neutral pH, the Aβ(16-22) peptides assemble into solid fibers (Figure 4.1a) with an antiparallel in-register β-sheet structure [21, 155] (Figure 4.1b, c). The glutamic acid (E22) at the C-terminus bears a negative charge and forms a cross-strand salt bridge with the positively-charged lysine residue (K16). However, if the pH value is below the pKa of glutamic acid, it loses its negative charge and thus the salt bridge as well. Aβ(16-22) then assembles into hollow nanotubes (Figure 4.1d) with an antiparallel out-of-register β-sheet structure (Figure 4.1e) [21, 155]. Without the salt bridges, the out-of-register structure is stabilized by the stereo preference of the bulky valine side chain cross-strand paired with the less bulky alanine [155, 163] (Figure 4.1f).

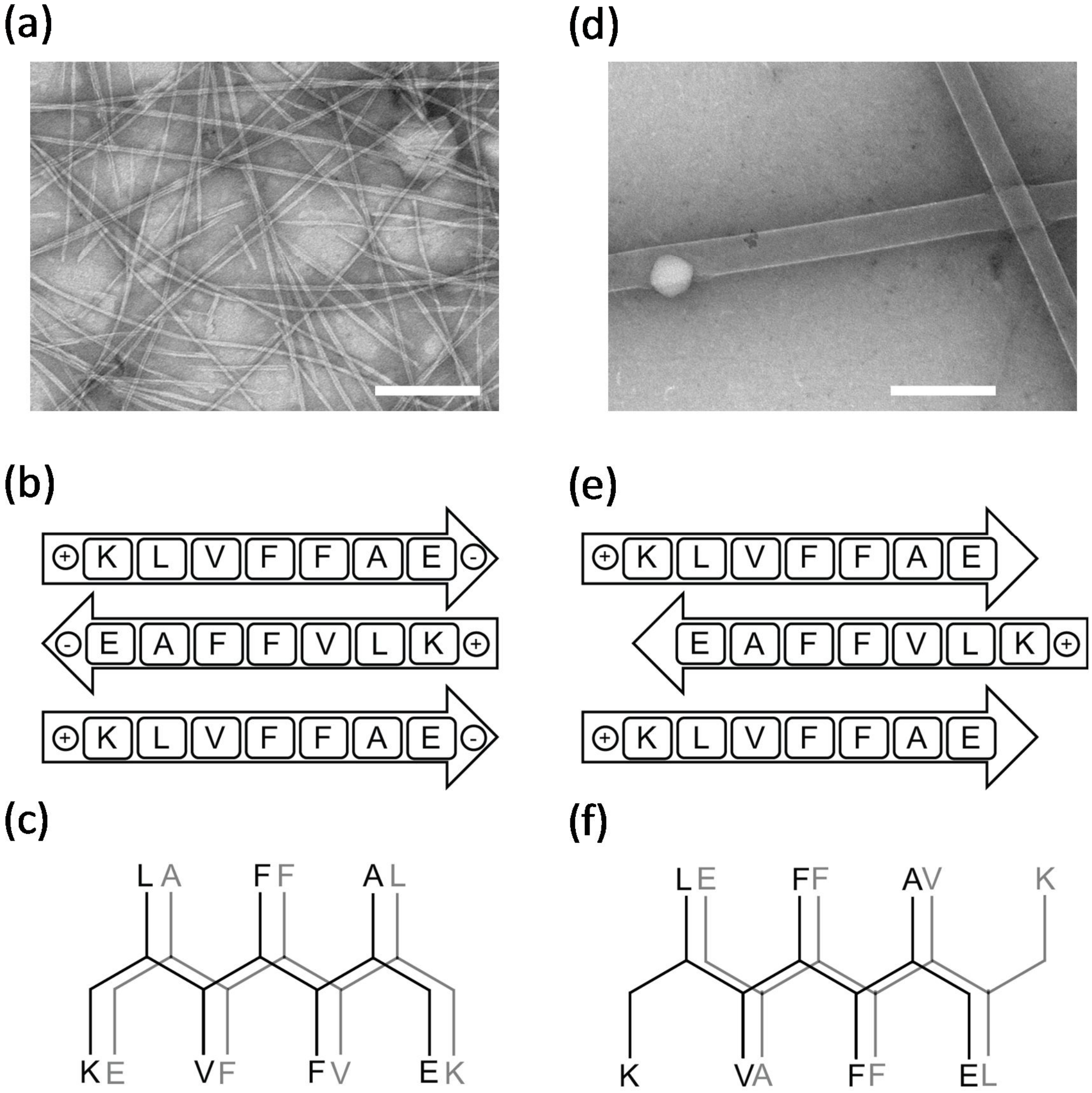

Figure 4.1 – Assembly of Aβ(16-22) at neutral and acidic pH. At neutral pH, Aβ(16-22) assembles into (a) solid fibers with an (b) antiparallel in-register β-sheet with side-chain stacking shown in (c). At acidic pH, Aβ(16-22) assembles into (d) hollow tubes with an (e) antiparallel out-of-register β-sheet with side-chain stacking shown in (f). Scale bar = 200 nm.

Although the structure of Aβ(16-22) assemblies have been widely-studied [164], the assembly kinetics of Aβ(16-22), especially at acidic pH, have not been clearly elucidated [77]. The critical nucleation events of Aβ(16-22) assembly have been computationally simulated [1, 5, 165-167], but these simulations are focused on the early stage of the assembly process. Experimentally, Nilsson *et al.* measured the kinetics of Aβ(16-22) assembly with HPLC, while they did not provide enough microscope images to morphologically explain the possible topology transitions from the possible intermediates to the final products [168]. In Ref. [21] it is shown that the assemblies of Aβ(16-22) emerged from the particle phase, and thus the kinetics follow a two-step nucleation mechanism [64, 169]. Lin and coworkers incubated Aβ(16-22) in phosphate-buffered saline (PBS) to probe its assembly kinetics [77]. Before the Aβ(16-22) fibers arise from metastable particles, a flake-like side product is observed first. As revealed by in situ atomic force microscopy (AFM), the flakes are metastable and dissolve as the fibers mature. However, in this study the coupling registry of the assembled peptides, which influences the stability and the morphology of the assemblies [45, 60], is not followed with instruments like FTIR nor NMR. Also, none of the above experimental works report the assembly process at acidic pH, and kinetics of Aβ(16-22) assembly at acidic pH, which leads to the nanotube structures, have not been demonstrated yet.

Here we report the assembly kinetics of isotope-edited Aβ(16-22), [1-^{13}C]F19 Aβ(16-22), at both neutral and acidic pH to elucidate whether the reaction pathways are pH-dependent. To investigate the possible intermediates and the degree of assembly maturation, circular dichroism (CD) is utilized to probe the β-sheet content, and transmission electron microscopy (TEM) is used to observe the morphologies of the assemblies. In addition, isotope-edited FTIR analyses are performed to follow the reaction pathways and the structural transitions by elucidating the registry populations of the assembled peptides [45]. Combined together, these techniques show that at both neutral and acidic pH, [1-^{13}C]F19 Aβ(16-22) forms the same ribbon intermediates, while the ribbons undergo different reaction pathways in the solution which lead to different final macrostructures.

4.2 Results and Discussion

4.2.1 Development of the IR Basis Sets

The IR basis set necessary for registry analysis is obtained with [1-^{13}C]F19 Aβ(16-22) solutions independent from those for kinetic measurements. Compared to the other residues, the residue F19 of Aβ(16-22) is chosen for ^{13}C-enrichment because it provides the most significant difference in IR bands between different registries [155]; these registries have been previously defined by solid state NMR as antiparallel in-register β-sheet at neutral pH and out-of-register β-sheet at acidic pH [155]. To obtain the IR signatures of the in-register and out-of-register structures, [1-^{13}C]F19 Aβ(16-22) solutions with a concentration of 1.0 mM are incubated at neutral and acidic pH, respectively. Their IR spectra are assigned as the basis set after the spectra remain constant for at least one month, as shown in Figure 4.2.

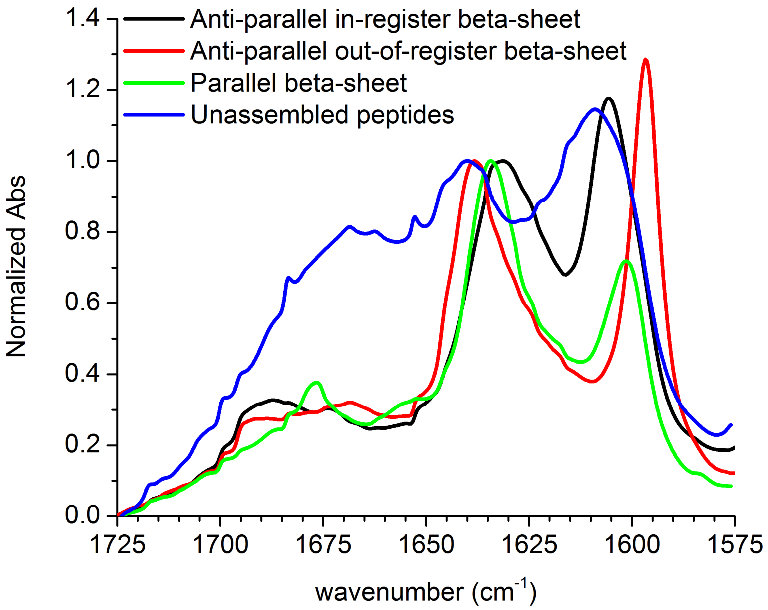

Figure 4.2 – The IR basis set for spectrum deconvolution. For parallel β-sheet signature (black), the spectrum is obtained with [1-^{13}C]F19 Aβ(16-22)E22Q assemblies [60]. The antiparallel in-register β-sheet signature (red) is obtained from [1-^{13}C]F19 Aβ(16-22) assemblies incubated at neutral pH, while the antiparallel out-of-register signature (green) comes from the same peptide incubated at acidic pH. The [1-^{13}C]F19 Aβ(16-22) peptide is dissolved in HFIP to disassemble the possible assemblies, resulting in the unassembled peptide signature (blue) [60].

The maturation of these assembly morphologies is further confirmed by TEM images (Figure 4.3). The images are obtained after the IR spectra of the peptide solution remain constant for more than a month. As shown in Figure 4.3, [1-^{13}C]F19 Aβ(16-22) assembles into fibers at neutral pH and tubes at acidic pH. However, particles are observed along with these assemblies, independent of pH. As the particle phase may contribute to IR absorbance, experiments have been carried out to obtain the IR signature of the pure particle phase for basis set development. To decrease the supersaturation and thus exclude the possible assemblies, fresh [1-^{13}C]F19 Aβ(16-22) solution is incubated with increased temperature, different acetonitrile concentrations, and/or a lower peptide concentration. However, assemblies are always observed together with the particles by

TEM images, and thus currently the IR basis set for the pure particle phase is not available yet. As the particles may not be as ordered as the assemblies, the contribution of the particle phase to the IR spectra is assumed to be similar to the unassembled peptides (Figure 4.2, blue line). The spectrum deconvolution is then used to focus on the registry analysis between the assemblies. Also, to make the basis set robust enough to analyze other possible structures, the parallel β-sheet signature from [1-^{13}C]F19 Aβ(16-22)E22Q (Figure 4.2, black line) and the unassembled peptide signature from [1-^{13}C]F19 Aβ(16-22) (Figure 4.2, blue line) are obtained from Ref. [60]. It should be noticed that, in Ref. [60], [1-^{13}C]F19 Aβ(16-22)E22Q also forms a particle-assembly equilibrium at the end of the incubation, which is similar to the final state of [1-^{13}C]F19 Aβ(16-22) assembly.

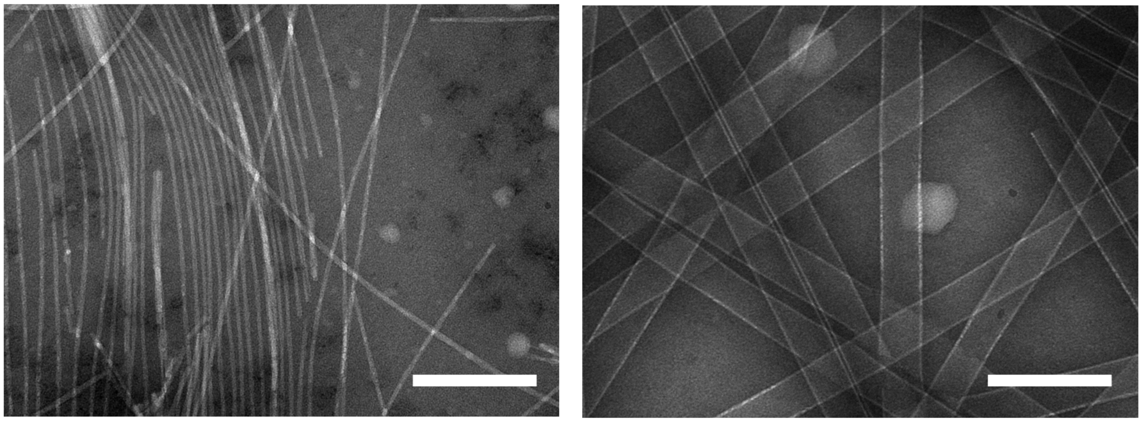

Figure 4.3 – TEM images of 1.0 mM [1-^{13}C]F19 Aβ(16-22) solutions after the IR signature becomes stable at (left) neutral pH and (right) acidic pH. Scale bar = 200 nm.

4.2.2 *Morphological evolution of [1-^{13}C]F19 Aβ(16-22) Assembly*

A series of TEM images reports the morphological changes of the [1-^{13}C]F19 Aβ(16-22) assemblies at acidic pH and neutral pH. As shown in Figure 4.4 and Figure 4.5, the assembly of [1-^{13}C]F19 Aβ(16-22) is pH-responsive, acidic conditions giving hollow nanotubes and neutral conditions giving fibers [164, 170-173]. Despite this morphological difference, the pathway appears remarkably similar. After one hour of incubation, the ribbons are formed together with the particles at both pH (Figure 4.4a and Figure 4.5a), while they are the metastable products which diminish during the incubation. At the acidic pH, the ribbons are the on-pathway intermediates; after they grow longer and wider, they mature into the tube morphology [21, 174] (Figure 4.4b and c). At neutral pH the fibers and the particles grow with the amount of ribbons diminishing (Figure 4.5b and c). After the ribbons disappear, the particles, which are a constant phase throughout the incubation, coexist and equilibrate with either the tubes at acidic pH (Figure 4.4d) and the fiber at the neutral pH (Figure 4.5d). The final morphology difference between fibers and tubes is the result of fibers containing antiparallel in-register β-sheets (Figure 4.1b and c), while the hollow nanotubes contain antiparallel β-sheets that are shifted out-of-register by a single residue (Figure 4.1e and f) [21, 155].

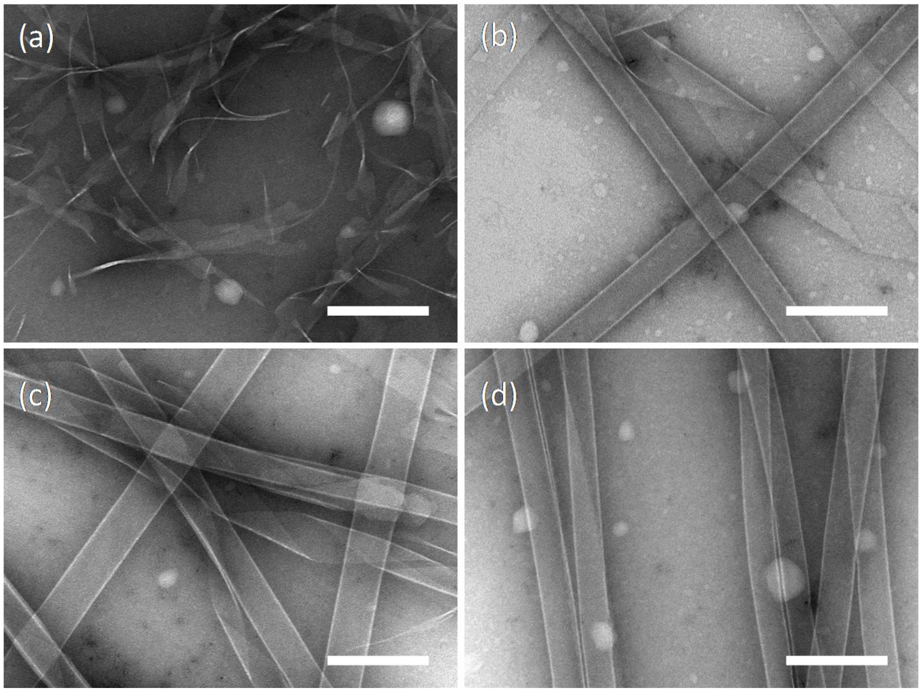

Figure 4.4 – TEM images of 1.0 mM [1-^{13}C]F19 Aβ(16-22) solution at neutral pH after incubation for (a) 1 hr, (b) 7 days, (c) 10 days, and (d) 14 days. Scale bar = 200 nm.

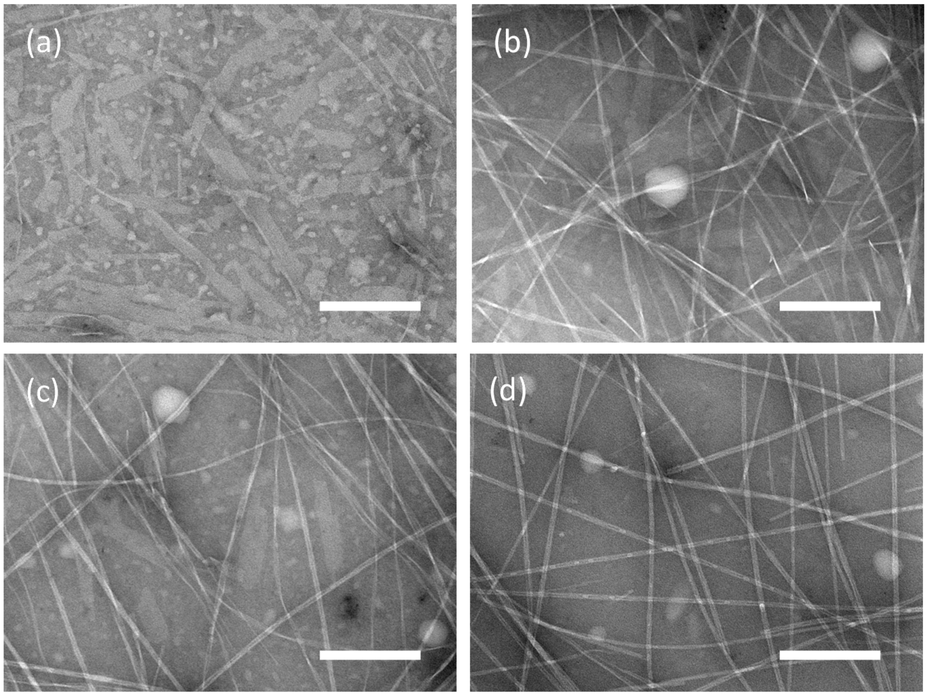

Figure 4.5 – TEM images of 1.0 mM [1-^{13}C]F19 Aβ(16-22) solution at neutral pH after incubation for (a) 1 hr, (b) 1 day, (c) 3 days, and (d) 9 days.

4.2.3 *Maturation of [1-^{13}C]F19 Aβ(16-22) assemblies probed by CD*

CD follows the growth of the β-sheet content. At neutral pH (Figure 4.7, black) the acquired ellipticity is much weaker [175] but the β-sheet signature develops immediately and plateaus after day five, consistent with TEM-determined maturation (Figure 4.5). In contrast, at acidic pH the negative ellipticity at 217 nm appears insensitive to the ribbons (Figure 4.6a and Figure 4.7, blue). To test CD's capability of detecting Aβ(16-22) ribbons, Aβ(16-22) solution is prepared at acidic pH with a lower peptide concentration (0.5 mM), where the ribbons accumulate (Figure 4.8) while they remain insensitive to CD (Figure 4.9). Although CD cannot probe the ribbons at the early time points, the β-sheet content at acidic pH grows as the intermediate ribbons transition into tubes (Figure 4.7, blue). After two days, the negative ellipticity establishes and plateaus after twelve days (Figure 4.7, blue), again consistent with the time dependence obtained with TEM images (Figure 4.4).

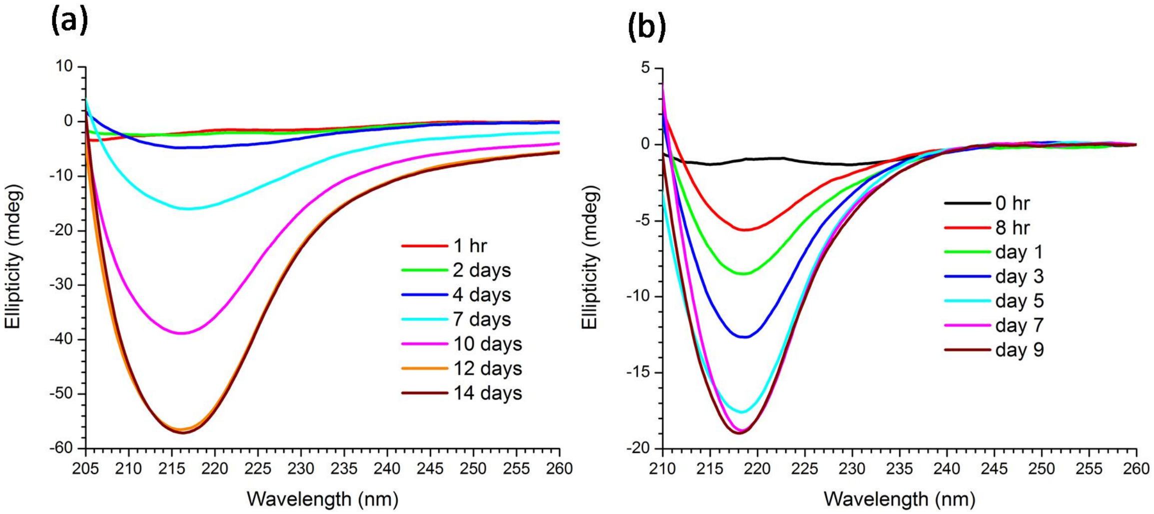

Figure 4.6 – CD spectrum of 1 mM [1-^{13}C]F19 Aβ(16-22) solution in 40% Acetonitrile at (a) acidic pH and (b) neutral pH as a function of time. The negative signature develops as the β-sheet concentration increases. The intensity at 217 nm becomes an indicator of the degree of self-assembly.

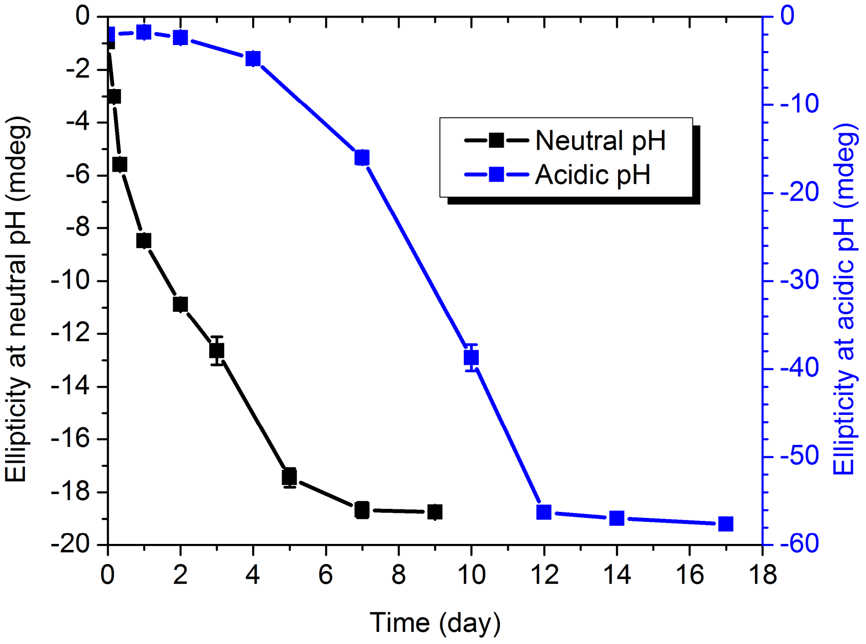

Figure 4.7 – Ellipticity at 217 nm measured by CD for 1 mM [1-^{13}C]F19 Aβ(16-22) solution in 40% acetonitrile under (blue, right y-axis) acidic pH and (black, left y-axis) neutral pH over time.

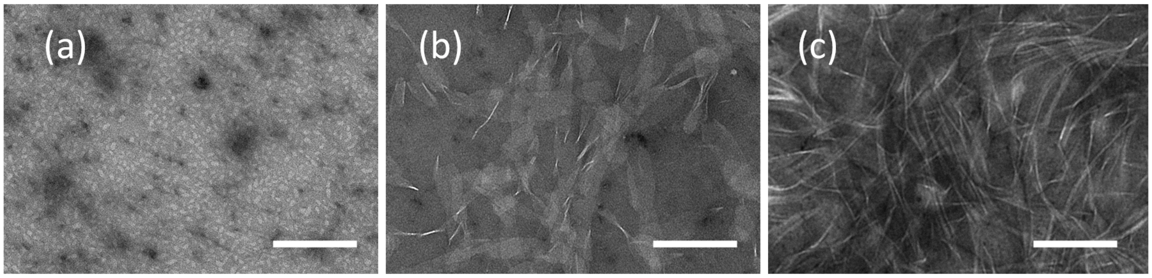

Figure 4.8 – TEM images of 0.5 mM Aβ(16-22) at acidic pH after (a) 0 hr, (b) 1 day and (c) 8 days.

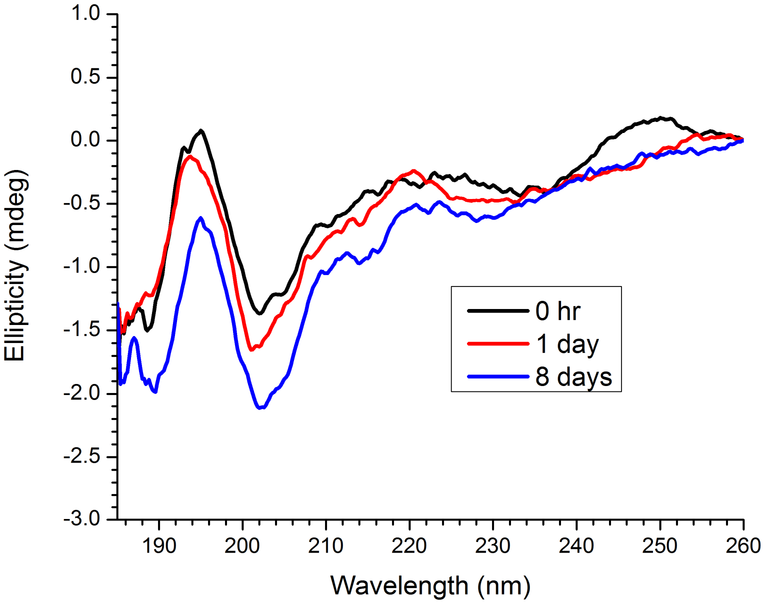

Figure 4.9 – The CD spectra of 0.5 mM Aβ(16-22) at acidic pH.

4.2.4 Isotope-edited IR Analyses for Registry Transition

Isotope-edited FTIR is applied to reveal the coupling between the assembled [1-^{13}C]F19 Aβ(16-22) peptides. With the ^{13}C-enrichment, the [1-^{13}C]F19 Aβ(16-22) assembly shows a ^{12}C peak around 1638 cm^{-1} and a ^{13}C peak around 1597 cm^{-1} (Figure 4.10); for better comparison, the IR spectrum is normalized based on the height of the ^{12}C peak [45, 60]. To detect the possible transition of the inter-strand coupling of the assembled peptides

[45, 60, 155], these IR spectra are scaled with the basis set from the matured assemblies (Figure 4.2), and are represented in terms of the scaling coefficients as shown in Figure 4.11.

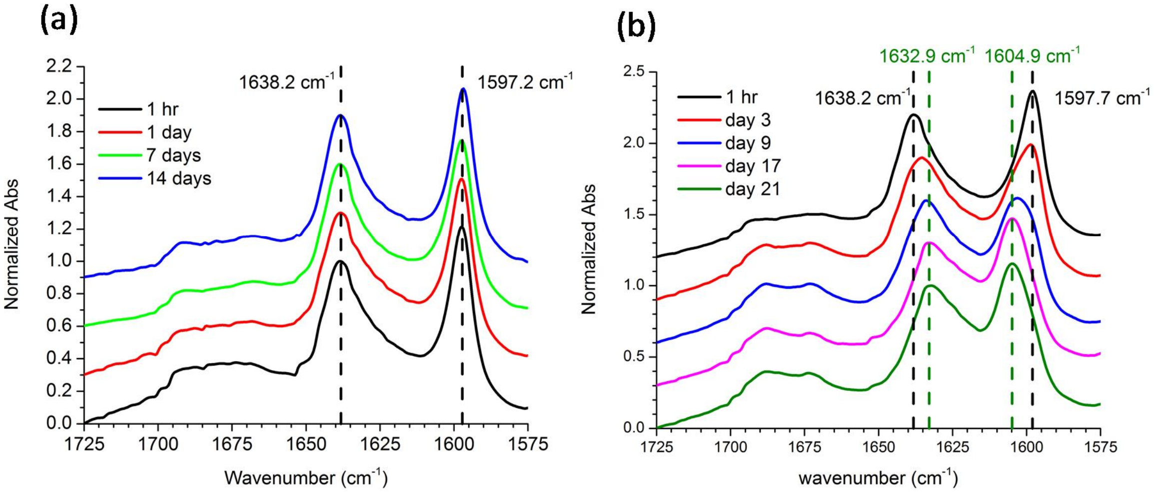

Figure 4.10 – IR spectra of 1 mM [1-^{13}C]F19 Aβ(16-22) solution in 40% Acetonitrile at (a) acidic pH and (b) neutral pH.

The anti-parallel out-of-register β-sheets signature dominates immediately under both acidic and neutral conditions (Figure 4.11), consistent with the ribbons existing as independent of pH (Figure 4.4a and Figure 4.5a). At acidic pH, the ribbons transition to the nanotubes without changing peptide registry; as the morphological change (Figure 4.4) does not affect the peptide registry, the ribbons are the on-pathway intermediates for the tubes. The minor in-register β-sheet component in Figure 4.11a at the early time points imply that the registry may still be flexible and the out-of-register structure is not entirely matured yet (Figure 4.11a). These minor components decreases after the most stable out-of-register structure predominate after day 12, on which all the ribbons disappear by TEM images (Figure 4.4).

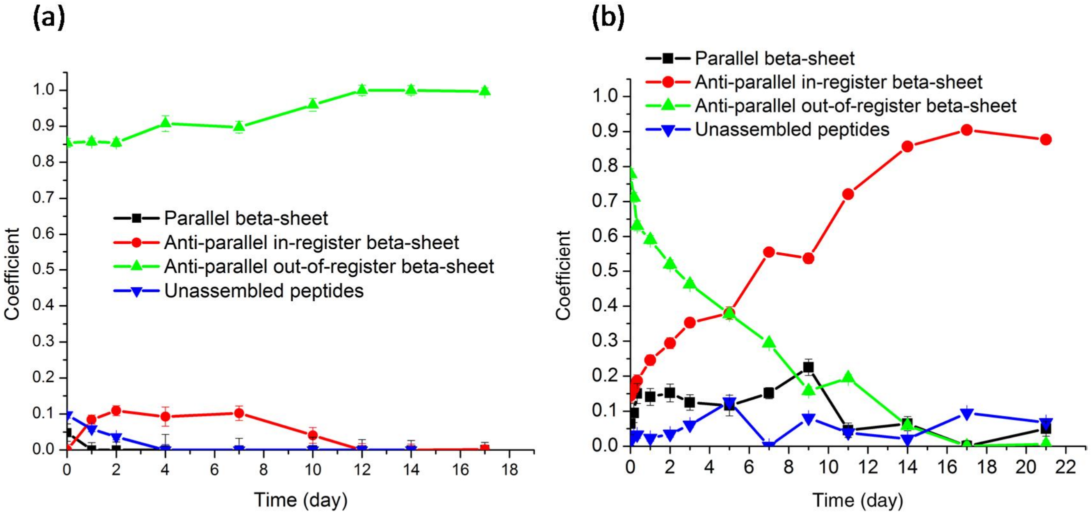

Figure 4.11 – Fits of IR spectra of [1-^{13}C]F19 Aβ(16-22) under (a) acidic pH and (b) neutral pH. The kinetic IR spectra of [1-^{13}C]F19 Aβ(16-22) solution are fitted by the basis set in Figure S3 to reveal any possible registry transition.

Although it is incapable of distinguishing the intermediates and final products at acidic pH, at neutral pH IR probes the transition from the intermediate ribbons to final fibers (Figure 4.10b and Figure 4.11b). The initial antiparallel out-of-register β-sheet structure suggests the same ribbon intermediates are formed pH-independently. However, at neutral pH this initial IR signature does not remain constant, but slowly transitions into the in-register β-sheet signature. The registry transition continues for 14 days, while the CD intensity stops changing on day nine. The inconsistent paces may come from the slow coupling maturation compared to the fiber growth. After the peptides assemble into fibers with the β-sheet structure, the coupling between the peptides may not be stabilized with the in-register structure yet. The flexible couplings then contribute to other structures such as out-of-register β-sheet, before the fibers eventually anneal into the stable form. Thus, the IR transition continues until the assembled peptides mature with the stable in-register structure at neutral pH, as observed in Figure 4.11b.

Combined together, the kinetics and the structural evolution of [1-^{13}C]F19 Aβ(16-22) assembly are revealed. The peptide initially assembles into ribbons at both acidic and neutral pH, which later mature into the final tubes or fibers. Also, IR, CD and TEM are all necessary to provide this comprehensive understanding. The IR probes the peptide coupling patterns, the TEM images reveal the morphological transition, while CD measures the maturation of the assemblies.

4.2.5 Compare the Assembly Pathway of Aβ(16-22) at Neutral and Acidic pH

The kinetics of Aβ(16-22) assembly at both neutral and acidic pH is measured to probe the possible pH-dependent assembly pathway. Through two-step nucleation, the nucleation and growth environments are different yet essential for assembly development. It has been shown that the Aβ(16-22) peptide assemblies nucleate inside the particles, while they mature in the solution [21, 63, 77]. The environmental difference between the particle and the solution phases may impact the transition between different assembly morphologies [21].

Independent of pH, Aβ(16-22) initially forms metastable ribbons which are intermediates with antiparallel out-of-register β-sheet structure. The Dutch congener of Aβ(16-22), E22Q, forms antiparallel out-of-register β-sheet ribbons as well, which transition to fibers with parallel in-register β-sheet structure [60]. The common ribbon morphology with the same out-of-register structure suggest that Aβ(16-22) and E22Q share the similar and pH-independent environment (particles) for ribbon nucleation. As the out-of-register coupling is driven by the stereo preference (Figure 4.1), the formation of Aβ(16-22) ribbons at neutral pH suggests that the particle environment may have affected the pK_a of the peptide residues, and the charges of either K16 and E22 may become weakened. As a result, Aβ(16-22) forms out-of-register ribbons as it does not access the salt bridge even at neutral pH, the same intermediate structure for E22Q and Aβ(16-22) at acidic pH, both of which cannot access inter-strand salt bridges to stabilize the structure. Although the properties of the particles remain to be elucidated, the above results show that the E22Q and Aβ(16-22) particles are probably similar and are not affected by different pH conditions.

Although formed independent of pH, the metastable ribbon intermediates undergo pH-dependent reaction pathways for the final product maturation. At the acidic pH, the ribbons are the on-pathway intermediates as they transform themselves into tubes. Because the out-of-register β-sheet is stable at acidic pH, the tubes keep the registry from the intermediate ribbons. However, at neutral pH, the fibers do not inherit the same registry from ribbons, as their in-register structure is more stable at neutral pH. The ribbons disappear and the fibers mature after one week of incubation, based on the CD (Figure 4.6) and TEM observation (Figure 4.5). The mechanism underlying the registry shift is not defined. One possible mechanism would require two separate nuclei in the particle; the in-register fibers nucleate more slowly, and after they emerge into the solution, the out-of-register ribbons dissolve and the released peptides sequestered into the more stable fibers.

Another possible mechanism, more consistent with the autocatalytic growth of the E22Q fiber in solution, is that the out-of-register structure templates an in-register conformational mutation. The dock-and-lock models for propagation at the ends of the assemblies [176] could then fix an incoming strand so the cross-strand salt bridge directs the locking of the new strand and stabilize it. The multi-stage assembly pathway then diversifies the environmental context as well as the range of structures available to amyloid assembly, and raises fundamental questions about the nature of the particle phase.

4.3 Conclusion

In this study we combined isotope-edited IR, CD, and TEM to analyze the kinetics of Aβ(16-22) assembly at acidic and neutral pH. At both pH conditions, the same ribbon intermediates with out-of-register anti-parallel β-sheet are detected. This implies that the ribbon nucleation environments, the particles, of Aβ(16-22) are similar and not pH-dependent. However, the ribbons undergo different reaction pathways for assembly maturation under neutral and acidic pH conditions. At neutral pH, the out-of-register ribbons are less stable than the in-register fibers as the latter are stabilized with the salt bridges between the lysine and glutamic acid residues. The out-of-register ribbons disappear as the fibers grow and mature. At acidic pH, since the out-of-register coupling is stable, the ribbons are the on-pathway intermediates, which wrap themselves into the final tube structures.

This is the first study reporting the kinetics of Aβ(16-22) assembly at acidic pH. Combined together, the structural transition analyzed with the deconvolution of IR spectra, the morphologies of the assemblies recorded by TEM, and the beta-sheet content measured by CD, make it possible to follow the transition of the peptide assembly process from the molecular level to the macrostructural level. The methodology used and the assembly mechanism revealed in this study may extend disease therapy design and novel material construction.

4.4 Methods

4.4.1 Peptide Solution Preparation

The [1-^{13}C]F19 Aβ(16-22) peptides are synthesized using the F-MOC protected natural abundance and the [1-^{13}C]F19-enriched amino acids with solid state peptide synthesis [21, 45, 60]. The synthetic peptides are purified with high-performance liquid chromatography (HPLC), in a C-18 reverse phase column with a water/acetonitrile gradient with 0.1% trifluoroacetic acid (TFA). The peptide solution from HPLC then undergoes lyophilization and the resulting peptide powder is stored under -20 ^{o}C before use.

To prepare the peptide solution for the kinetic measurements, the peptides are dissolved in hexafluoroisopropanol (HFIP) before sonication for two hours. HFIP is then evaporated with a Labconco CentriVap Concentrator-7970010, and the resulting peptide film is placed in a desiccator overnight to remove residual HFIP. To assemble the peptides at neutral pH, the disaggregated peptides are suspended again with HFIP, and sonicated for 15 min. The peptides in HFIP are then diluted with 40% (v:v) acetonitrile in water to obtain a final HFIP concentration of 1% by volume and the desired peptide concentration. For acidic solution, 40% acetonitrile with 0.1% TFA is used instead of pure 40% acetonitrile to dilute the suspended peptides in HFIP. The peptide solution is then incubated at 22 ^{o}C and aliquots of the solution are removed at predetermined time points for kinetic measurements.

4.4.2 Transmission Electron Microscopy (TEM)

The topologies of the peptide assemblies are reported with transmission electron microscopy (TEM). First, three water drops with an individual volume of 100 μL are loaded onto a piece of parafilm for sample dilution. Then, 5 μL of peptide solution is loaded onto the copper grid for one minute to deposit the assemblies. After one minute of deposition, the grid is put upside down to let the peptide solution contact the three water drops in sequence, which dilutes the peptide solution and washes away access assemblies. The diluted peptide solution on the grid is wicked away with a piece of filter paper and the grid is negatively stained with 1.5-wt % methylamine tungstate solution for 30 seconds. The stain solution is wicked away with filter paper and the grid is placed in a vacuum desiccator to dry overnight. TEM images are recorded with a Hitachi H-7500 transmission electron microscope.

4.4.3 Circular Dichroism (CD)

The β-sheet content of the peptide solution is probed by circular dichroism (CD). CD measurement is performed with 23 μL peptide solution in a demountable window cell with a 0.1 mm path length. The spectrum is recorded using a Jasco J-810 spectropolarimeter with wavelengths from 185–260 nm, a resolution of 0.2 nm, a bandwidth of 2 nm, and a scanning rate of 200 nm per min. After background correction, the average of three consecutive scans is saved and the ellipticity at 217 nm is used to quantify the β-sheet content of the peptide solution.

4.4.4 Attenuated Total Reflectance Fourier Transform Infrared (FTIR)

The coupling of the assembled peptides is measured with isotope-edited FTIR to identify the registry. Eight μL of the peptide solution is loaded onto the diamond chip of a JASCO FT/IR-4100. After the sample dries down, the IR absorbance of the sample is recorded against a background spectrum at a wavenumber range of 1000 to 3600 cm^{-1} at room temperature. The average of 250 consecutive scans is saved and the spectrum is normalized to the peak height of the ^{12}C band.

An IR basis set is established to probe the structural transition of the assemblies with spectrum deconvolution. The registries of matured Aβ(16-22) assemblies have been previously defined by solid state NMR as antiparallel in-registry at neutral pH and out-of-registry at acidic pH [155]. Peptide solutions, independent from those for kinetic measurements, are incubated at either neutral or acidic pH as described in Section 2.1. Their spectra are assigned as the basis set for their corresponding registries after the spectra remain constant for more than a month, when the assembly structure does not change significantly. For spectrum deconvolution with the basis set, the kinetic IR spectra with the range from 1575 cm^{-1} to 1725 cm^{-1} are analyzed as the linear combination of the basis set with the NonlinearModelFit module of *Mathematica 10.0* (Wolfram Research, Inc., 2014).

CHAPTER 5.

CATALYTIC FUNCTION OF PEPTIDE ASSEMBLIES

5.1 Introduction

Protein enzymes catalyze diverse reactions in aqueous environments with high selectivity [24] and yield [23], and their outstanding catalytic ability inspires researchers to construct and design artificial enzymes. However, the complex protein folding mechanism has hindered the development of amino acid-based catalysts, as misfolded and unfolded proteins do not access the correct active sites for reactions.

On the other hand, similar to protein folding for ordered structures, peptide assembly has become an attractive bottom-up method to make catalysts, because solving the protein folding problem is not required [177]. Here we construct and investigate peptide assemblies that are able to catalyze specific enantioselective chemical reactions. The catalytic strength of these assemblies is rationally adjusted by engineering the structure of individual amino acid residues. The size of the binding pocket and enantioselectivity are analyzed with a modified Michaelis-Menten mechanism [178, 179], which also reveals the physical properties of these peptide assemblies. Combined together, the experimental and modeling results help to unveil the catalytic nature of these peptide assemblies.

5.2 Results and Discussion

5.2.1 *The Catalytic Activity from the Assemblies of Ac-KLVFFAE-NH_2 and its Analogs*

The reactive residues on the assembly surface may have catalytic potential. For Ac-KLVFFAE-NH_2 (Aβ(16-22)) and its analogs, the amine groups on the lysine residues (K16) may be catalytic for amine-driven reactions [180]. To test the catalytic ability of these peptide assemblies, here methodol (4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone) is selected as the substrate, which undergoes retro-aldol catalysis [181] and has been applied in several catalytic systems [182-184]. As shown in Figure 5.1 [180], methodol is cleaved into an acetone and a fluorescent aldehyde (6-methoxy-2-naphthaldehyde), the latter of which may be quantified with fluorescence to report the reaction progress.

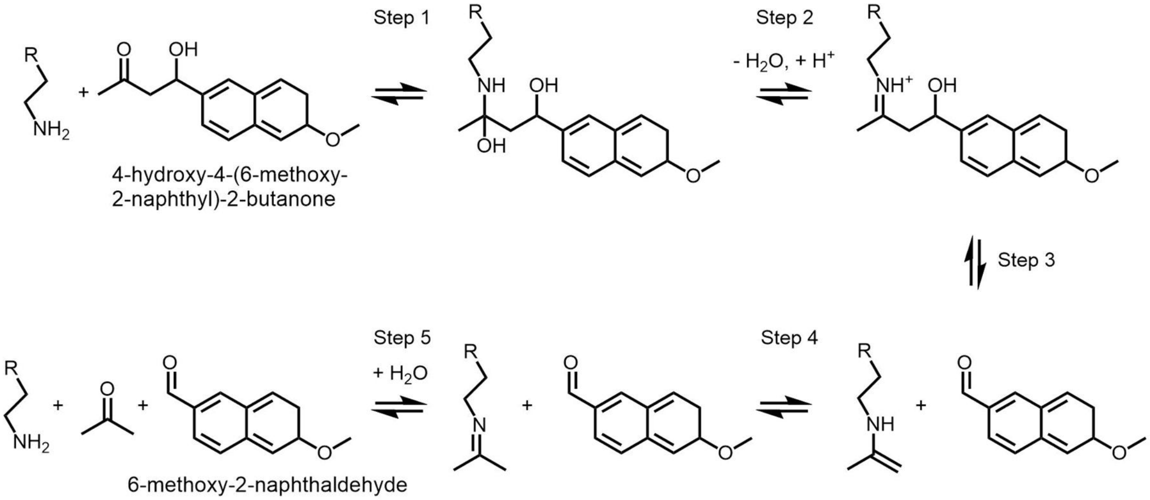

Figure 5.1 – Retro-aldol reaction of methodol (4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone) catalyzed with the amine group on the lysine residue [180].

Several assemblies of the Ac-KLVFFAE-NH_2 analogs are mixed with methodol to test their catalytic capability. Figure 5.2 summarizes the initial rates from these peptide assemblies, including both solid fibers and hollow tubes. The Ac-KLVFFAE-NH_2 fiber is inactive toward the methodol substrate and does not cleave the methodol substrate. As the electrostatic interaction between the positive lysine and negative glutamic acid residues may affect the catalytic ability of the assemblies, this salt bridge is removed by replacing the glutamic acid with the neutral glutamine (Q). This glutamine analog, Ac-KLVFFAQ-NH_2, assembles into fibers as well, but it does not cleave the methodol substrate (Figure 5.2). On the other hand, the E22L analog, Ac-KLVFFAL-NH_2, is able to cleave the methodol substrate (Figure 5.2). Ac-KLVFFAL-NH_2 assembles into tube with anti-parallel β-sheet structure [185], and the flexible lysine residues on the tube surface provide a range of possible docking modes for methodol (Figure 5.3).

Not all peptide tubes are catalytically functional, and the microenvironment in the binding sites is crucial for methodol cleavage. When the valine (V18) in Ac-KLVFFAE-NH_2 is replaced with *tert*-leucine (*ter*-L), the Ac-KL(*ter-L*)FFAE-NH_2 peptide assembles into tubes with the surface similar to that of Ac-KLVFFAL-NH_2 [163]; however the negatively charged glutamic acid in Ac-KL(*ter-L*)FFAE-NH_2 assemblies limits the catalytic efficiency for methodol cleavage, as shown in Figure 5.2. No retro-aldol catalysis is detected for Ac-KL(*ter-L*)FFAQ-NH_2 either, which is the Q analog of Ac-KL(*ter-L*)FFAE-NH_2. Ac-KL(*ter-L*)FFAQ-NH_2, Ac-KL(*ter-L*)FFAE-NH_2, and Ac-KLVFFAL-NH_2 peptides all have the amine groups on the lysine side chains, but they are all catalytically inactive. This suggests that the possible free peptides in these systems cannot cleave the substrate, and the catalytic activity from the Ac-KLVFFAL-NH_2 tubes is associated with the binding pocket only. Otherwise, Ac-KL(*ter-L*)FFAQ-NH_2 and Ac-KL(*ter-L*)FFAE-NH_2 would have shown catalytic ability as well, given the free peptides. Finally, to highlight the important primary amine for the catalysis, the lysine (K16) in Ac-KLVFFAL-NH_2 is replaced with arginine (R), which does not have the primary amine group for retro-aldol reaction. As shown in Figure 5.2, the Ac-RLVFFAL-NH_2 tubes are not able to perform methodol cleavage, indicating the importance of the primary amine group on the lysine side chain.

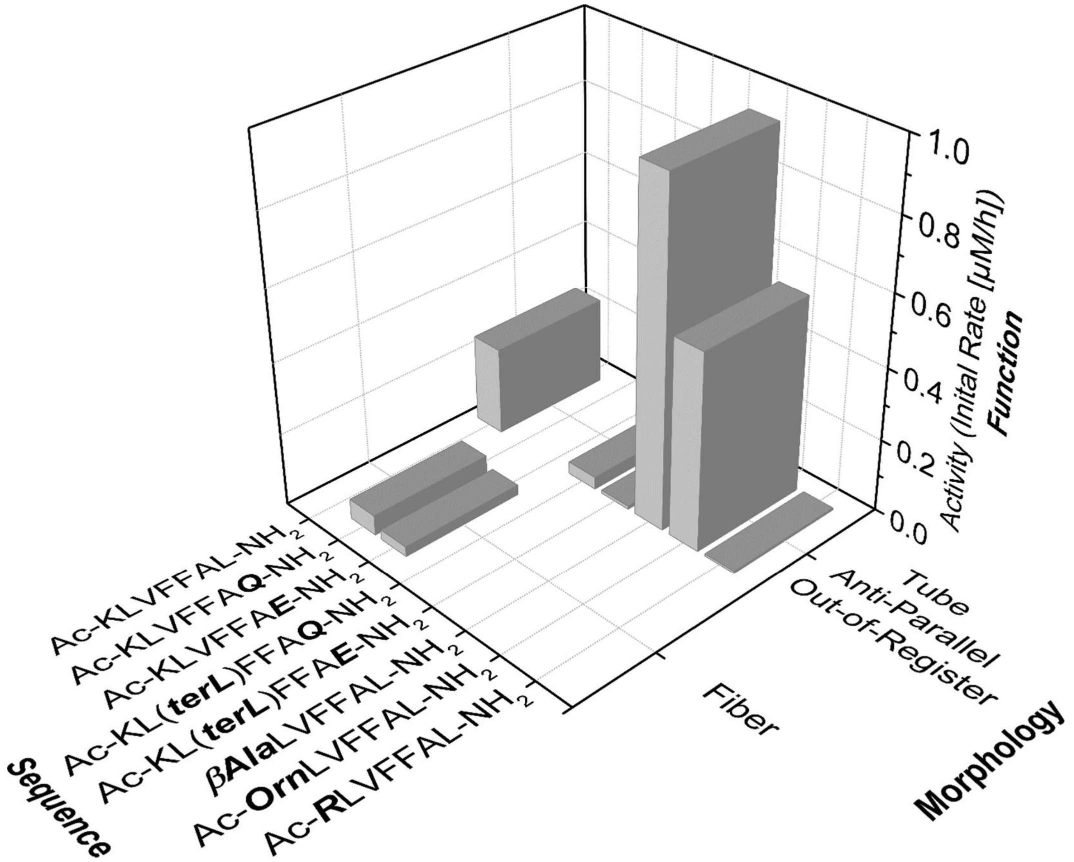

Figure 5.2 – Initial rate of production of 6-methoxy-2-naphthaldehyde from 80 µM (±)-methodol with the assemblies of 500 µM Ac-KLVFFAE-NH_2 and its analogs. Image and data adapted from Ref. [186] with permission.

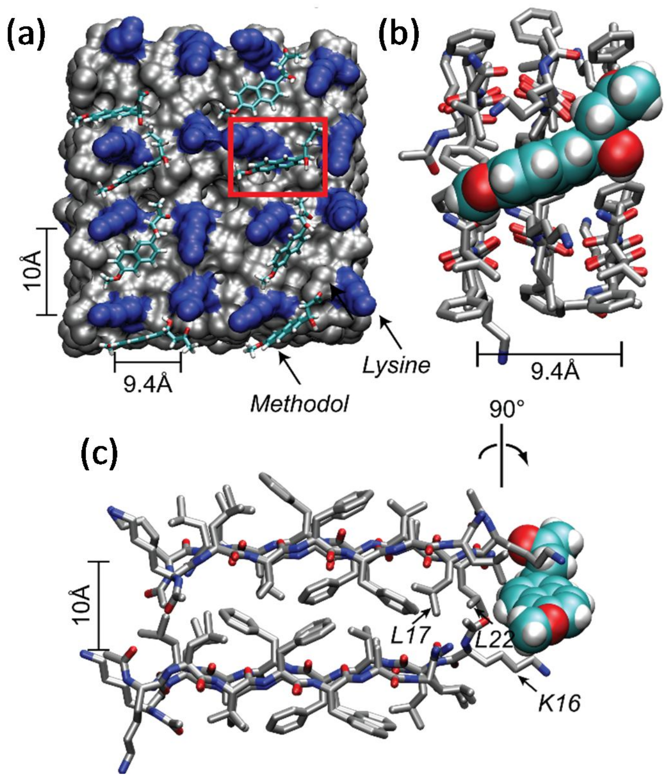

Figure 5.3 – Docking of methodol onto the KLVFFAL tube surface. (a) Molecular dynamics simulation of S-methodol docking onto the surface of KLVFFAL anti-parallel out-of-register amyloid assembly. The lysine residues are colored blue, LVFFAL residues are colored grey, and the methodol substrates are drawn as sticks with carbons colored green, oxygen red and hydrogen white. (b, c) Expansions of methodol (space filling) on tube surface with peptides drawn as sticks. Image and data adapted from Ref. [186] with permission.

The catalytic efficiency of the Ac-KLVFFAL-NH_2 tubes is rationally adjusted by modifying the reactive amine group. The distance of the amine groups from the surface may be crucial to improve the catalytic efficiency; if the amine groups are closer to the tube surface, the substrates may react with the amine groups more efficiently. To test this hypothesis, the amine tether of the lysine residue is shortened by replacing lysine with ornithine and β-alanine (Figure 5.4). As shown in Figure 5.2, the catalytic efficiency is raised when the amine group becomes closer to the tube surface, as Ac-OrnLVFFAL-

NH_2 and βAlaLVFFAL-NH_2 break the methodol more quickly than the Ac-KLVFFAL-NH_2 does.

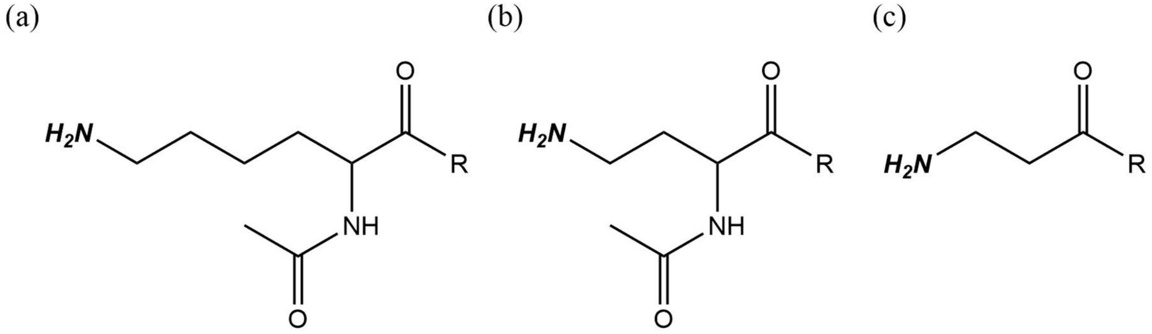

Figure 5.4 – Catalytic amine groups for retro-aldol reaction from (a) Ac-KLVFFAL-NH_2, (b) Ac-OrnLVFFAL-NH_2, and (c) βAlaLVFFAL-NH_2 tubes. The catalytic amine groups are highlighted in bold and italic. The R group represents the LVFFAL-NH_2 group.

Together with the Ac-KLVFFAL-NH_2 tubes, the Ac-OrnLVFFAL-NH_2 and the βAlaLVFFAL-NH_2 tubes have shown catalytic capability, and the catalytic efficiency is enhanced when the primary amine group is closer to the tube surface (Figure 5.2). However, as the βAlaLVFFAL-NH_2 tubes are not stable and transition into ribbons after several days under the reaction condition, the following kinetic analyses are focused on the Ac-KLVFFAL-NH_2 and the Ac-OrnLVFFAL-NH_2 tubes.

5.2.2 Retro-aldol kinetics for the peptide assemblies

To investigate the retro-aldol kinetics of the peptide assemblies in details, 500 μM assembled peptides are mixed with 76.7 μM racemic methodol solution, which contains equal R- and S-methodol; the unreacted R- and S-methodol concentration is determined with chiral HPLC and the concentration of the aldehyde product (6-methoxy-2-

naphthaldehyde) is determined by fluorescence. The kinetic analysis of Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 tubes is shown in Figure 5.5.

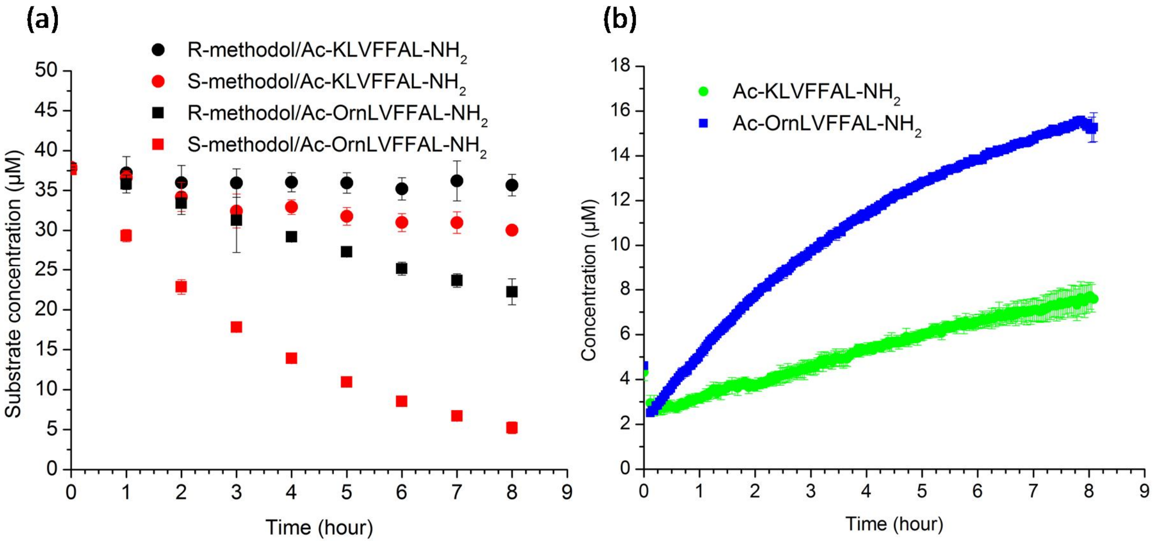

Figure 5.5 – The progress of retro-aldol reaction for methodol with Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 tubes. (a) Chiral HPLC of 500 µM peptide and 76.7 µM racemic methodol. (b) Naphthaldehyde fluorescence of 500 µM peptides with 76.7 µM racemic methodol. Data adapted from Ref. [187]. Error bars are included with size similar to the data points.

The chiral HPLC measures the racemic substrate consumption, which provides the catalytic efficiency and the selectivity of the peptide assemblies. Figure 5.5a shows that both the Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 tubes enantioselectively cleave the racemic methodol: both tubes have a higher conversion rate for S- over R-methodol (Figure 5.5a). Given a shorter tether length for the amine group, Ac-OrnLVFFAL-NH_2 breaks the methodol more quickly than the Ac-KLVFFAL-NH_2 does, consistent with its higher initial rate indicated in Figure 5.2.

The aldehyde product rebinds to both of the Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 tubes. After eight hours of reaction, around 10 μM of methodol is reacted with Ac-KLVFFAL-NH_2 tubes (Figure 5.5a), while only 6 μM of the aldehyde product is detected (Figure 5.5b). For Ac-OrnLVFFAL-NH_2 tubes, they cleave around 35 μM methodol but only 15 μM aldehyde product is detected, a difference of 20 μM aldehyde product. As the aldehyde product shows fluorescence when it is dissolved in the solution but not in the binding pocket, the inconsistency between the substrate and product concentration implies that the aldehyde may rebind onto the tube surface. The product-rebinding behavior is detected in Figure 5.5b as well: before the methodol solution is mixed with the Ac-KLVFFAL-NH_2 assemblies, it contains 2.5 μM of initial fluorescent aldehyde, the unreacted reagent after methodol synthesis. However, the aldehyde concentration drops to ~2 μM immediately after the mixture, which indicates the instant binding between the tubes and the aldehyde upon mixing.

To clarify the catalytic kinetics, the binding site concentration is required. Unlike protein enzymes, whose concentration may be directly derived from the protein concentration, here the binding site concentration of the tubes remains to be determined. Not all primary amine groups are capable of methodol cleavage; half of the amine groups are buried in the bi-layer structure of the peptide tubes [185], and the concentration of the active groups on the tube surface remains to be determined. To experimentally determine the number of peptides per binding site, 300 μM assembled Ac-KLVFFAL-NH_2 peptides are titrated with 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione. 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione has a similar size as the methodol substrate, binds onto the tube surface to form an enaminone, but does not undergo further cleavage (Figure 5.6).

The titration analysis shows that at least four peptides are needed to form a binding site (Figure 5.7) for 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione.

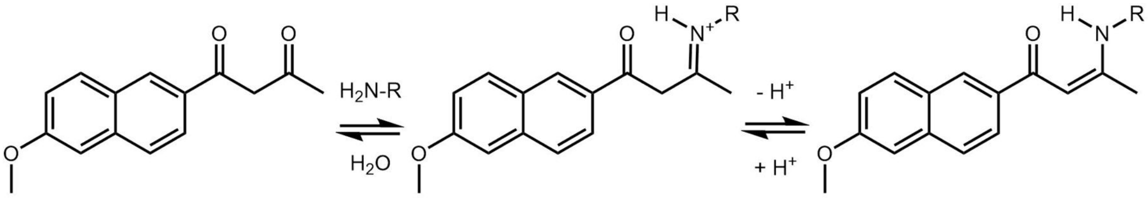

Figure 5.6 – The binding step between the amine group on the lysine residue and 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione [188].

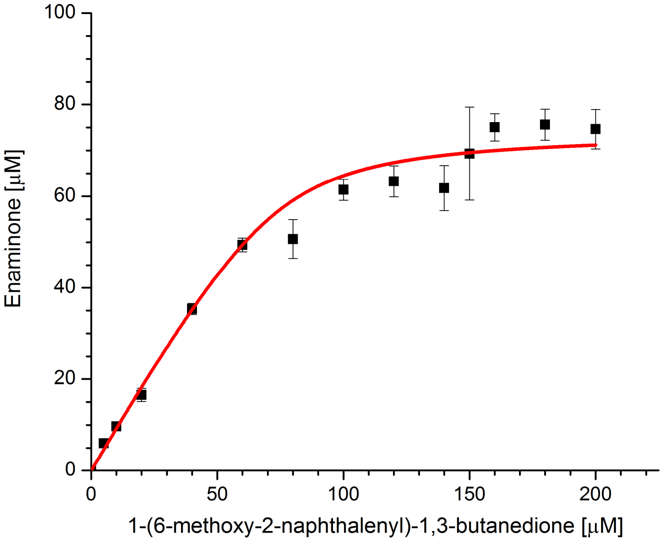

Figure 5.7 – Titrating 300 µM Ac-KLVFFAL-NH_2 with 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione to determine the number of peptides per binding site. Enaminone formation upon addition of 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione to 300µM Ac-KLVFFAL-NH_2 monitored by UV absorbance at 350 nm. When the amine groups on the tube surface is saturated with 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione and form enaminone as indicated in Figure 5.6, the enaminone concentration is about 76 µM. This suggests that the tube requires around 4 peptides (300 µM /76 µM) to form a binding site. Image adapted from Ref. [186] with permission.

To resolve the size of an active binding site, to verify the product-rebinding mechanism, and to investigate the enantioselectivity, the Michaelis-Menten mechanism is modified for a bi-substrate system with product-rebinding to fit the experimental data. Retro-aldol kinetics is fit to this minimal model that contains the following steps (Figure 5.8):

(1) Equilibria between the methodol (S- and R-substrates) on the tube surface (*E*).

(2) Carbon-carbon bond cleavage.

(3) Product (P_1 for aldehyde and P_2 for acetone) release.

(4) Equilibrium between the product, 6-amino-2-naphthaldehyde (P_1), and tube surface.

$$E+S \underset{k_{forS}*K_{dS}}{\overset{k_{forS}}{\rightleftharpoons}} ES \xrightarrow{k_{catS}} E+P_1+P_2$$

$$E+R \underset{k_{forR}*K_{dR}}{\overset{k_{forR}}{\rightleftharpoons}} ER \xrightarrow{k_{catR}} E+P_1+P_2$$

$$EP_1 \underset{k_a}{\overset{k_aK_d}{\rightleftharpoons}} E+P_1$$

Figure 5.8 – The modified Michaelis-Menten mechanism for retro-aldol reaction. *E* is binding site, *S* is S-methodol, *ES* is the bound complex, P_1 is 6-methoxy-2-naphthaldehyde, P_2 is acetone, *R* is R-methodol, *ER* is the bound complex, and EP_1 is the enzyme-6-methoxy-2-naphthaldehyde complex.

The methodol consumption and aldehyde production are used to evaluate the parameters of the model in Figure 5.8. The parameter optimization is carried out by minimizing the sum of square error (SSE) between the calculated and experimental values, and the SSE is normalized for the number of data points and the scale of the measurements [135]. Latin Hypercube sampling [136] is used to sample 500 various initial guesses for the kinetic parameters. The optimization procedure was carried out with the **`fmincon`** function in MATLAB 2012a (Mathwork, Inc).

The modified Michaelis-Menten mechanism is able to describe the reaction progress well (Figure 5.9), with eight peptides per binding sites for Ac-OrnLVFFAL-NH_2 and nine for Ac-KLVFFAL-NH_2 tubes (Figure 5.10). Also, adjusting the amine position does not alter the tubes' reactive mechanism, as Ac-OrnKLVFFAL-NH_2 tubes follow the same reactions as the Ac-KLVFFAL-NH_2 tubes, but with different rate constants. The optimal parameter sets (Table 5.1) for both tubes suggest that the binding affinity for the R-methodol is stronger than that for the S-methodol ($K_{dR} < K_{dS}$). However, the S-methodol is consumed more quickly because it has a stronger chemical activity ($k_{catS} > k_{catR}$). When the number of binding site is abundant, the R-methodol cannot take the advantage of its better binding affinity, and thus S-methodol is preferred by the tubes for kinetic reasons.

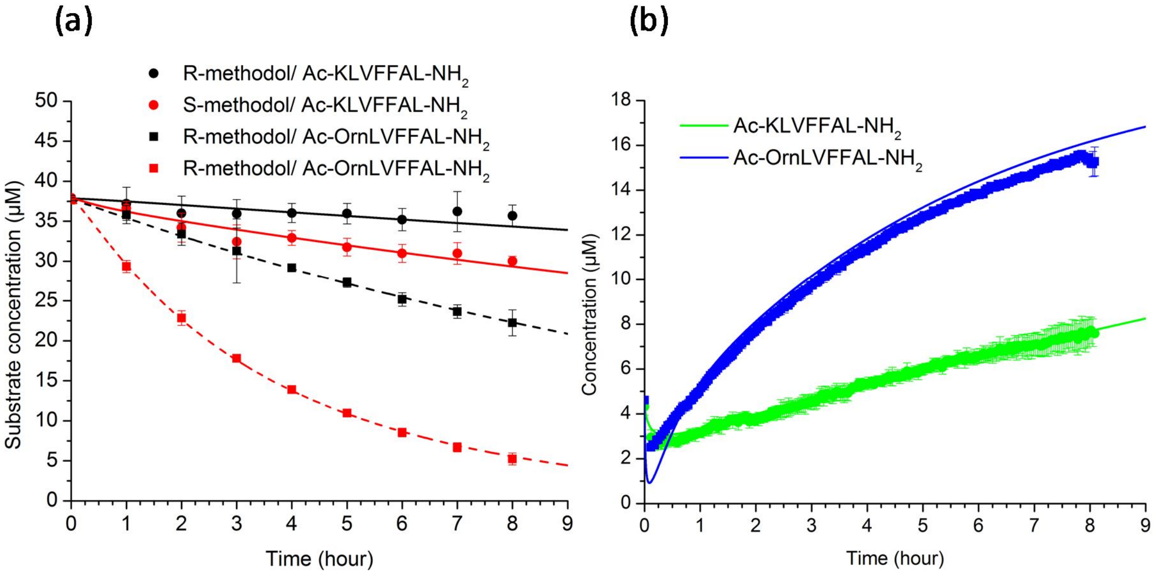

Figure 5.9 – The fits to the progress of retro-aldol reaction for methodol with Ac-KLVFFAL-NH$_2$ and Ac-OrnLVFFAL-NH$_2$ tubes. (a) Chiral HPLC of 500 µM peptide and 76.7 µM racemic methodol. (b) Naphthaldehyde fluorescence of 500 µM peptides with 76.7 µM racemic methodol.

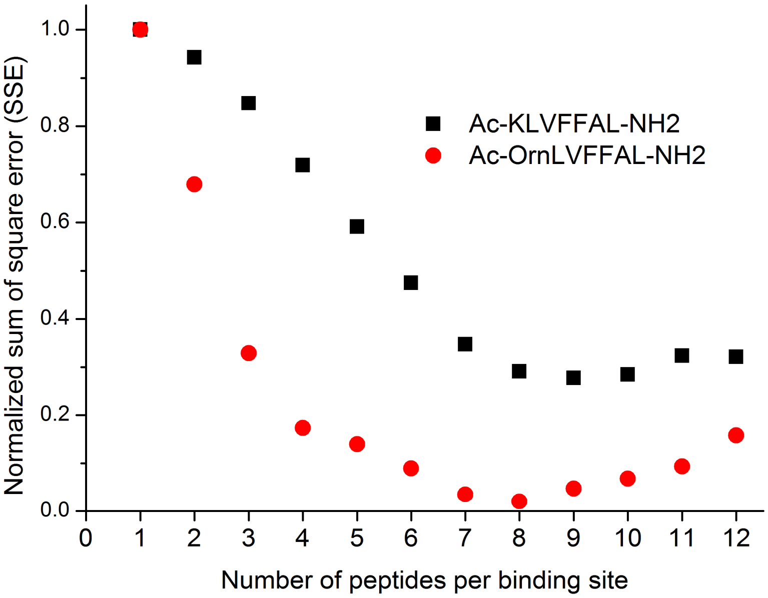

Figure 5.10 – Normalized sum of square error (SSE) for R- and S-methodol consumption on Ac-KLVFFAL-NH$_2$ (red) and Ac-(Orn)LVFFAL-NH$_2$ (black) assemblies. Methodol consumption and production 6-amino-2-naphthaldehyde in Figure 5.5 are fit to the model in Figure 5.8 as a function of the number of peptides per binding site.

Table 5.1 – The parameters of the best fits of Ac-KLVFFAL-NH$_2$ and Ac-OrnLVFFAL-NH$_2$ tubes. The confidence intervals of those parameters are at the *95% confidence level*.

Parameter	Ac-KLVFFAL-NH$_2$	Ac-OrnLVFFAL-NH$_2$
k_{forR} (s^{-1} μM^{-1})	$4.63 \pm 4.73 \times 10^{-5}$	$2.09 \pm 0.85 \times 10^{-4}$
k_{forS} (s^{-1} μM^{-1})	$6.19 \pm 15.7 \times 10^{-6}$	$4.92 \pm 0.82 \times 10^{-5}$
k_{catR} (s^{-1})	$3.46 \pm 1.17 \times 10^{-6}$	$1.87 \pm 0.06 \times 10^{-5}$
k_{catS} (s^{-1})	$4.31 \pm 1.57 \times 10^{-5}$	$1.48 \pm 0.08 \times 10^{-4}$
k_a (s^{-1} μM^{-1})	$2.67 \pm 1.97 \times 10^{-5}$	$2.66 \pm 0.68 \times 10^{-4}$
K_{dR} (μM)	$2.44 \pm 317 \times 10^{-3}$	$4.54 \pm 2.95 \times 10^{-2}$
K_{dS} (μM)	$3.18 \pm 2.34 \times 10^{2}$	$1.55 \pm 1.56 \times 10^{0}$
K_d (μM)	$5.83 \pm 3.83 \times 10^{-1}$	$9.93 \pm 1.47 \times 10^{-1}$

The R- and S- methodol has different binding and chemical propensities, which suggests the invertible enantioselectivity with different methodol/peptide ratio. The peptide tubes break S-methodol more quickly when the number of binding sites is not limited, due to its stronger chemical reactivity. When binding sites are scarce, the R-methodol would be enantioselectively consumed as it has a stronger binding affinity and would occupy most of the binding sites; this hinders the binding of the S-methodol and the following chemical steps. Simulation shows that both tubes undergo such invertible enantioselectivity when the catalyst load is limited (Figure 5.11), which is defined as the ratio of the binding site concentration over the substrate concentration:

$$\text{Catalyst load} = \frac{[\text{binding site}]}{[\text{substrate}]} \tag{5.1}$$

Similar to the enantiomeric excess [189, 190], the enantioselectivity in Figure 5.11 is defined as:

$$\text{Enantioselectivity} = \frac{R_c - S_c}{R_c + S_c} \times 100\% \tag{5.2}$$

where R_c and S_c are the concentration of consumed R- and S-methodol.

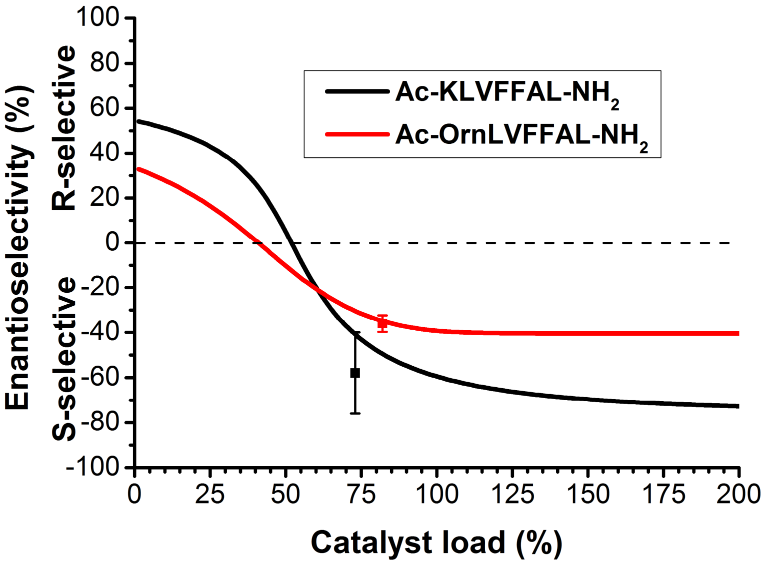

Figure 5.11 – The invertible enantioselectivity of the Ac-KLVFFAL-NH$_2$ and the Ac-OrnLVFFAL-NH$_2$ tubes. Reaction time is eight hours and the parameter sets in Table 5.1 are used. The two data points represent the experimental results from Figure 5.5a.

Different combinations of methodol/peptide concentrations are applied experimentally to verify the hypothesis about the invertible enantioselectivity. The catalyst load is decreased for all conditions to test whether the tubes enantioselectively cleave the R-methodol; however, the S-methodol is enantioselectively preferred under all conditions tested here. Although the new data sets do not support the hypothesis, the joint fits of these data sets provide more insight about the properties of the catalytic peptide assemblies.

Joint fits are carried out with the modified Michaelis-Menten mechanism (Figure 5.8) for the extended data sets as a function of numbers of peptides per binding site. With the new methodol consumption and aldehyde generation data, the mechanistic model in Figure 5.8 suggests the number of peptides per binding sites equal to 6 (Figure 5.12). The normalized SSE shows similar trends for both the Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 assemblies, with both having the lowest at six. This number is slightly greater than the experimentally determined value, which is four (Figure 5.7). The size difference in the calculated results and the experimental value indicates that not all binding sites are active for retrol-aldol reaction; the methodol may need to bind into the binding pocket with some specific relative positions with the amine group to trigger retro-aldol reaction.

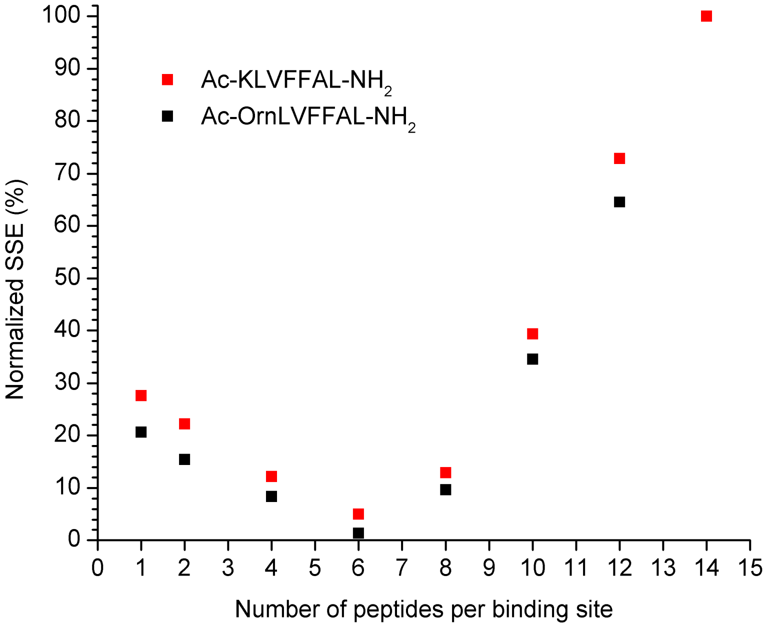

Figure 5.12 – Normalized sum of square error (SSE) of the joint fits for R- and S-methodol consumption on Ac-KLVFFAL-NH_2 (red) and Ac-OrnLVFFAL-NH_2 (black) assemblies.

With six peptides per binding site, the modified Michaelis-Menten mechanism fits the kinetics well for both Ac-KLVFFAL-NH_2 (Figure 5.13) and Ac-OrnLVFFAL-NH_2 (Figure 5.14) catalyzed substrate consumption and product accumulation. For the first 8 hours given 500 μM peptides and 80 μM substrate, the fits show that the binding step, the chemical steps and the product-rebinding step are essential. The same parameters sets (Table 5.2) also fit the other conditions with different peptide and substrate concentration as a joint fit. As shown in Figure 5.13c, d and Figure 5.14c, d, the joint model fits follow the decrease of R- and S- methodol concentrations with doubled substrate concentration (160 μM) for up to 128 hours with two peptide concentrations (500 and 300 μM). Even though the catalyst load is decreased, the enantioselectivity of both tubes does not change as predicted in Figure 5.11, and the tubes still break S-methodol more quickly. Also with the joint fits, both the Ac-KLVFFAL-NH_2 and the Ac-OrnLVFFAL-NH_2 tubes bind R- and S-methodol without preference, given the close dissociation constants (K_{dR} and K_{dS}) in Table 5.2. Meanwhile, both type of tubes have a stronger chemical activity for the S-methodol over the R-methodol ($k_{catS} > k_{catR}$); as both the two types of tubes have close binding affinity for R- and S-methodol, but break S-methodol more quickly, the enantioselectivity is not invertible by varying the catalyst load .

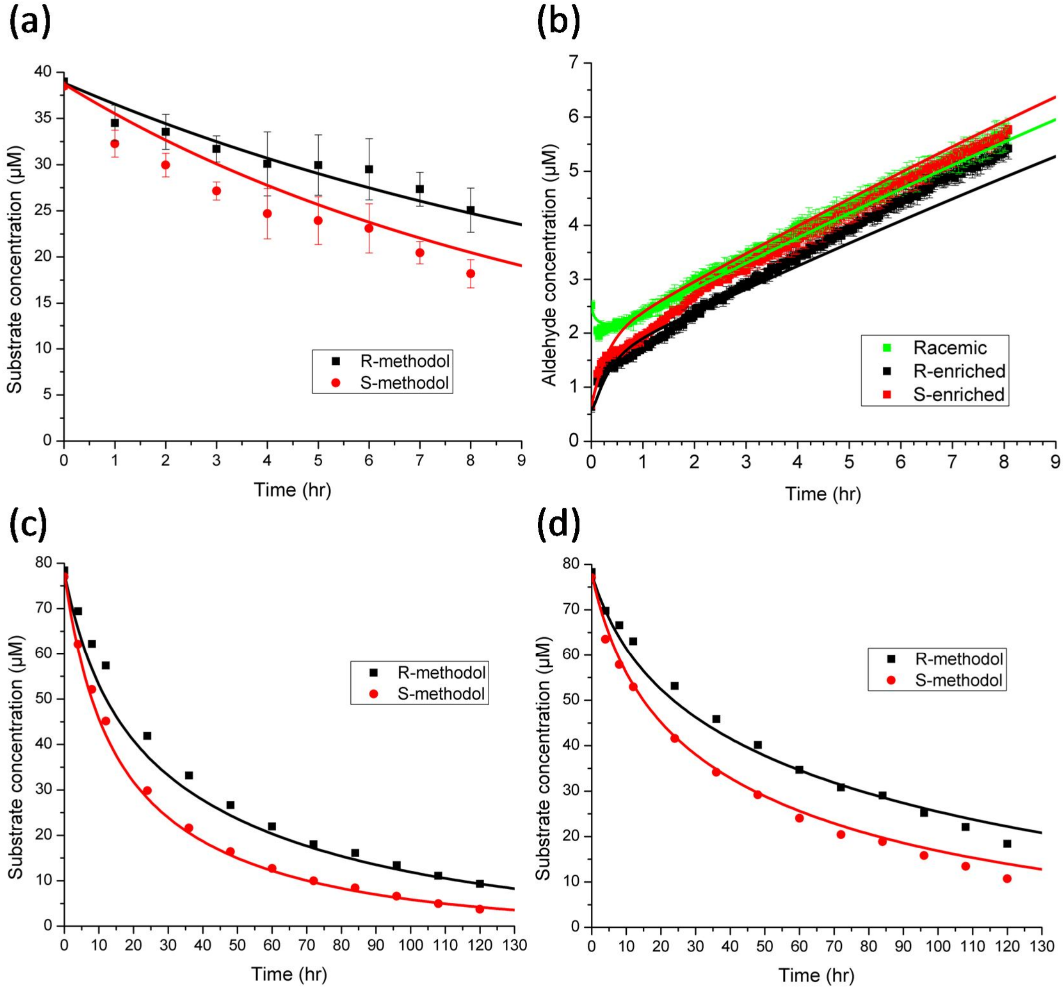

Figure 5.13 – Best fits of Ac-KLVFFAL-NH_2 retro-aldol catalysis with 6 peptides per binding site. (a) Chiral HPLC of 500 μM peptide and 76.7 μM racemic methodol. (b) Naphthaldehyde fluorescence of 500 μM peptides with 76.7 μM racemic methodol (green), 79.2 μM R-enriched methodol (black) and 79.1 μM S-enriched methodol (red). (c) Chiral HPLC of 500 μM peptide and 155.3 μM racemic methodol. (d) Chiral HPLC of 300 μM peptide and 155.4 μM racemic methodol. Solid lines are best fits with Figure 5.8. (a) Initial concentrations are: $[E]$ = 83.3 μM, $[S]$ = 38.3 μM, $[R]$ = 38.5 μM, $[P_1]$ = 3.1 μM. (b) The initial concentrations of the racemic solution are: $[E]$ = 83.3 μM, $[S]$ = 38.3 μM, $[R]$ = 38.5 μM, $[P_1]$ = 3.1 μM. The initial concentrations of the R-enriched substrate are: $[E]$ = 83.3 μM, $[S]$ = 11.9 μM, $[R]$ = 67.3 μM, $[P_1]$ = 0.77 μM. The initial concentrations of the S-enriched substrate are: $[E]$ = 83.3 μM, $[S]$ = 67.2 μM, $[R]$ = 11.9 μM, $[P_1]$ = 0.95 μM. (c) The initial concentrations are: $[E]$ = 83.3 μM, $[S]$ = 77.0 μM, $[R]$ = 78.4 μM, $[P_1]$ = 4.6 μM. (d) The initial concentrations are: $[E]$ = 50 μM, $[S]$ = 77.0 μM, $[R]$ = 78.4 μM, $[P_1]$ = 4.6 μM.

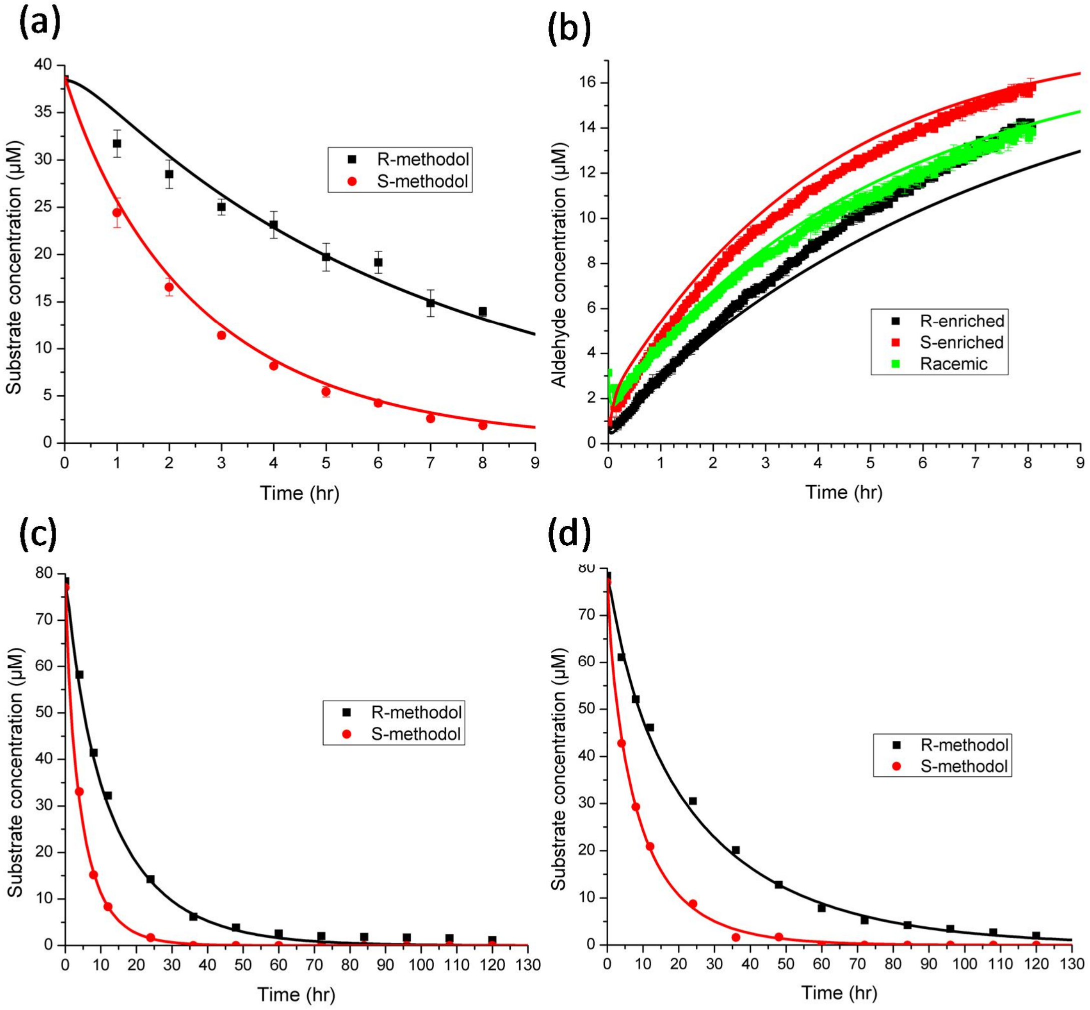

Figure 5.14 – Best fits of Ac-OrnLVFFAL-NH_2 retro-aldol catalysis with 6 peptides per binding site. (a) Chiral HPLC of 500 μM peptide and 76.7 μM racemic methodol. (b) Naphthaldehyde fluorescence of 500 μM peptides with 76.7 μM racemic methodol (green), 79.2 μM R-enriched methodol (black) and 79.1 μM S-enriched methodol (red). (c) Chiral HPLC of 500 μM peptide and 155.3 μM racemic methodol. (d) Chiral HPLC of 300 μM peptide and 155.4 μM racemic methodol. Solid lines are best fits from Figure 5.8 with (a) initial concentrations [*E*] = 83.3 μM, [*S*] = 38.3 μM, [*R*] = 38.5 μM, [P_1] = 3.1 μM. (b) Initial concentrations of the racemic solution (green) [*E*] = 83.3 μM, [*S*] = 38.3 μM, [*R*] = 38.5 μM, [P_1] = 3.1 μM, for R-enriched substrate (black) [*E*] = 83.3 μM, [*S*] = 11.9 μM, [*R*] = 67.3 μM, [P_1] = 0.77 μM and initial concentrations of the S-enriched substrate (red) [*E*] = 83.3 μM, [*S*] = 67.2 μM, [*R*] = 11.9 μM, [P_1] = 0.95 μM. (c) The initial concentrations are: [*E*] = 83.3 μM, [*S*] = 77.0 μM, [*R*] = 78.4 μM, [P_1] = 4.6 μM. (d) The initial concentrations are: [*E*] = 50 μM, [*S*] = 77.0 μM, [*R*] = 78.4 μM, [P_1] = 4.6 μM.

Table 5.2 – The optimal parameter sets for the joint fits of Ac-KLVFFAL-NH$_2$ and Ac-OrnLVFFAL-NH$_2$ tubes. The confidence intervals of those parameters are at the *95% confidence level*. K_d for both tubes are acquired independently by titrating the tubes with the 6-amino-2-naphthaldehyde.

Parameter	Ac-KLVFFAL-NH$_2$	Ac-OrnLVFFAL-NH$_2$
k_{forR} (s^{-1} μM^{-1})	$3.98 \pm 5.60 \times 10^{-5}$	$9.94 \pm 1.62 \times 10^{-6}$
k_{forS} (s^{-1} μM^{-1})	$9.18 \pm 45.0 \times 10^{-4}$	$1.17 \pm 0.25 \times 10^{-4}$
k_{catR} (s^{-1})	$4.08 \pm 1.26 \times 10^{-5}$	$6.00 \pm 0.69 \times 10^{-5}$
k_{catS} (s^{-1})	$6.20 \pm 2.39 \times 10^{-5}$	$1.90 \pm 0.35 \times 10^{-4}$
k_a (s^{-1} μM^{-1})	$1.60 \pm 0.07 \times 10^{-5}$	$7.47 \pm 0.20 \times 10^{-5}$
K_{dR} (μM)	$7.02 \pm 3.05 \times 10^{1}$	$1.21 \pm 0.35 \times 10^{1}$
K_{dS} (μM)	$7.96 \pm 4.33 \times 10^{1}$	$2.51 \pm 0.86 \times 10^{1}$
K_d (μM)	5.91×10^{1}	6.49×10^{1}

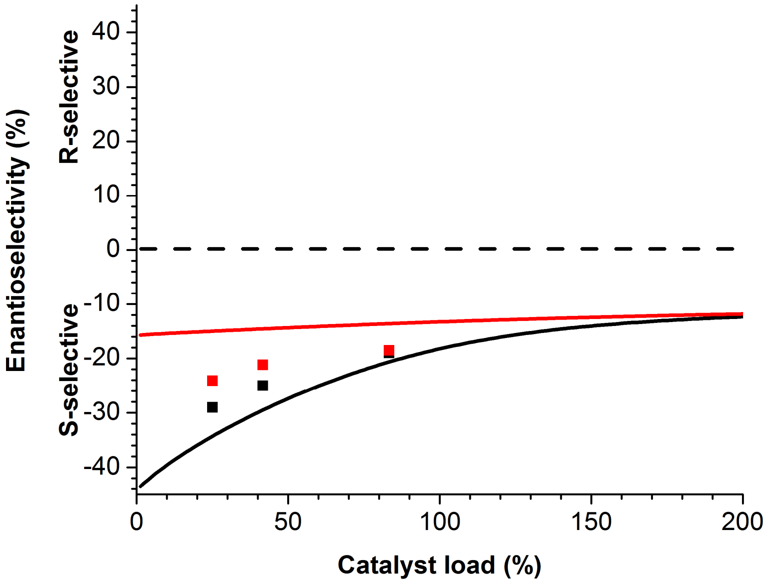

Figure 5.15 – The prediction of enantioselectivity of the Ac-KLVFFAL-NH$_2$ and the Ac-OrnLVFFAL-NH$_2$ tubes. Reaction time is eight hours and the parameter sets in Table 5.2 are used. The data points represent the experimental results from Figure 5.13 and Figure 5.14.

The joint fits also reveal the physical properties of the peptide tubes. The fitting results follow the long-term data well with the same mechanism (Figure 5.8) and the same parameter set (Table 5.2). This suggests that the tubes are stable, the number of binding sites remains constant, and the catalytic nature of these binding pockets does not change during the catalytic reaction. The joint fits also suggest that the model is capable describing the data from different substrate and peptide concentration, and the mechanism remain the same even though the substrate and the peptide assemblies are mixed in different ratio.

To compare the catalytic efficiency between different catalysts, the constants k_{cat}/K_M in the Michaelis-Menten kinetics (equation 5.3) are often used [191]:

$$\frac{dP}{dt} = \frac{k_{cat}[E]_{tot}[S]_l}{K_m + [S]_l} \tag{5.3}$$

where P is the product, $[E]_{tot}$ is the total enzyme concentration, and $[S]_l$ is the substrate concentration. However, our peptide-methodol system does not satisfy the assumption to obtain equation 5.3. Although the rapid equilibration predicted by our simulations satisfies the basic quasi-steady-state assumption of Michaelis-Menten kinetics, our model does not follow k_{cat}/K_M calculations as [ER] and [ES] do not remain constant due to the limited substrate solubility. In addition to limitations of substrate solubility, product binding also makes the kinetics deviate from Michaelis-Menten kinetics early in the initial rates. Thus, equation 5.3 cannot be directly applied to the peptide tubes and k_{cat}/K_M calculated from this mechanism for our system will not accurately reflect its catalyst efficiency.

The enantioselectivity of the peptide assemblies is clarified with modeling simulation and discrimination. To test either the binding or the chemical step has a stronger impact to the enantioselectivity, parameter estimation for the modified Michaelis-Menten model (Figure 5.8) is carried out with restriction of equal chemical step constants ($k_{catR} = k_{catS}$) or equal substrate dissociation constants ($K_{dR} = K_{dS}$) for both Ac-KLVFFAL-NH_2 (Table 5.3) and Ac-OrnLVFFAL-NH_2 (Table 5.4) tubes. The goodness of fits for these restricted models is then compared to the original model by the Akaike information criterion (AIC) [192, 193] and the Akaike weight [129], as mentioned in Chapter 2. Here, AIC (equation 5.4) is used instead of AICc (equation 2.1) as the data set is large enough ($\frac{N_d}{N_p} \geq 40$, where N_d is the number of data points and N_p is the number of parameters.) [194]. AIC measures the consistency of the models upon the data by rewarding the model with the less SSE or the less number of parameters, and is calculated based on:

$$\text{AIC} = N_d \ln\left(\frac{SSE}{N_d}\right) + 2N_p \tag{5.4}$$

where SSE is the sum of squares error. The Akaike weight, on the other hand, indicates the probability of a model to be more consistent with the data:

$$\text{Akaike weight} = \frac{\exp\left(-\frac{1}{2}\Delta_i\left(\text{AIC}\right)\right)}{\sum_{j=1}^{N_M}\exp\left(-\frac{1}{2}\Delta_j\left(\text{AIC}\right)\right)} \tag{5.5}$$

where N_M is the total number of models for comparison and $\Delta_i(\text{AIC})$ is the difference between the AIC of model *i* and the minimal AIC from the model set.

As shown in Table 5.3 and Table 5.4, the original model has the lowest AIC value and the greatest Akaike weight, most consistent with unequal dissociation constants and chemical steps. This suggests that the original model is not overfitting the data. However, the restriction on the chemical step constants ($k_{catR} = k_{catS}$) has a greater *SSE* than the dissociation constants for both tubes; this suggests that the chemical step has a stronger impact on the enantioselective consumption for both peptide assemblies. Therefore the nanotubes show a stronger binding affinity toward R-methodol, but still cleave S-methodol selectively.

Table 5.3 – AIC analysis of the best fits for the modified Michaelis-Menten model for Ac-KLVFFAL-NH_2 tubes. N_d is the number of data points, and N_p is the number of parameters.

Model	N_d	*SSE*	N_p	AIC	Akaike weight (%)
Original	793	7.94×10^1	7	-1810.80	92.83
$k_{catR} = k_{catS}$	793	2.00×10^2	6	-1079.94	0.00
$K_{dR} = K_{dS}$	793	8.01×10^1	6	-1805.68	7.17

Table 5.4 – AIC analysis of the best fits for the modified Michaelis-Menten model for Ac-OrnLVFFAL-NH_2 tubes. N_d is the number of data points, and N_p is the number of parameters.

Model	N_d	SSE	N_p	AIC	Akaike weight (%)
No restriction	793	2.83×10^1	7	-2629.40	100.00
$k_{catR} = k_{catS}$	793	1.07×10^2	6	-1573.16	0.00
$K_{dR} = K_{dS}$	793	4.07×10^1	6	-2343.13	0.00

5.3 Conclusion

In this chapter, the catalytic ability of the peptide assemblies is demonstrated and analyzed, which are able to cleave the methodol substrate. Among all the peptide assemblies tested, both Ac-KLVFFAL-NH_2 and Ac-OrnLVFFAL-NH_2 peptides show capability to catalyze the enantioselective chemical reaction with significant stability.

As shown with the modeling results, these catalytic assemblies cleave the substrate by following the Michaelis-Menten mechanism with a product-rebinding step. The modeling and simulation clarify the number of peptides required for an active binding site, the critical step for enantioselectivity, and the stability of the tubes. Both type of the peptide tubes follow the same mechanism but with different rates, indicating that the catalytic efficiency may be systematically and rationally adjusted without altering the overall catalytic nature.

5.4 Methods

5.4.1 Synthesis of Methodol

Racemic 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone is synthesized based on Ref. [181, 195] via condensation of acetone with 6-methoxy-2-napthaldehyde. Acetone (196 μL, 2.50eq) in tetrahydrofuran (THF) is gently added into a freshly prepared solution of lithium diisopropylamide (LDA) under −78°C. The mixture is stirred at this temperature for 30 min, and then 6-methoxy-2-naphthaldehyde (100 mg, 1.07 mmol) in 5 ml THF is gently added into the mixture within one minute. After it is stirred for another 30 minutes under −78°C, the reaction is quenched by saturated NH_4Cl under room temperature. The product is extracted with ethyl acetate and dried with $MgSO_4$, and then flash purified with SiO_2, ethyl acetate and hexane.

For R-methodol preparation, acetone (100 ml) is mixed with 6-methoxy-2-naphthaldehyde (887 mg, 4.76 mmol) and L-proline (248 mg, 0.40eq), and stirred at room temperature for 5 hours. The mixture is then dried down under vacuum, re-suspended in 40 ml vinyl acetate and treated with *P.cepacia* (1.32 g, 245 mg/mmol substrate), a 4Å molecular sieve was added. The mixture is placed at room temperature under N_2 protection overnight, and the reaction crude is filtered and applied to flash column.

For S-methodol, 6-methoxy-2-naphthaldehyde (500 mg, 2.68 mmol) and D-proline (124 mg, 0.40eq) are mixed with 60 ml acetone at room temperature for three hours. After the solids are removed by filtration, the filtrate is dried down under vacuum and applied to flash column directly.

5.4.2 Synthesis of 1-(6-methoxy-2-naphthalenyl)-1,3-butanedione

244 mg of 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone is dissolved in 10 mL chloroform, and the solution is incubated with 1.5 equivalents of Dess-Martin periodinane in chloroform at room temperature overnight. The product is purified by silica gel chromatography with hexanes and ethyl acetate.

5.4.3 Preparation of Peptide Assembly Solution

The peptides are synthesized using solid phase peptide synthesis, and the synthetic products are purified with high-performance liquid chromatography (HPLC), using a water/acetonitrile gradient with 0.1% trifluoroacetic acid in a C-18 reverse phase column. After lyophilization, the purified peptide is stored at -20 ^{o}C for later use [21].

To assemble the peptides, the peptides (2.5 mM) are dissolved in 40% acetonitrile (ACN) aqueous solvent with 0.1% trifluoroacetic acid (TFA). Dissolution is assisted by two minutes of continuous vortexing, followed by 15 minutes of sonication until solution becomes clear. The solution is placed under 4 °C for two weeks to allow the peptides to assemble.

5.4.4 Transmission Electron Microscopy

The morphologies of the assembled peptides are observed with transmission electron microscopy (TEM). The peptide solution is diluted for ten times, and then the diluted portion is loaded onto the copper grid for 1 min. The excess solution is wicked away with filter paper. The resulting grid is negatively stained with 2% uranyl acetate in water for 1-2 minutes for fibers and 4-5 minutes for nanotubes before the stain solution is removed with filter paper. The grid is then stored in a vacuum desiccator overnight to remove the remaining liquid. TEM images are recorded with a Hitachi H-7500 transmission electron microscope.

5.4.5 Molecular Dynamic Simulation

Methodol docking onto the KLVFFAL tube surface is performed manually with energy minimization in Macromodel (Schrodinger, Inc. Portland, OR) [196]. To define the lysine accessibility across the tube surface, a one ns molecular dynamics (MD) run (Desmond) [197] at 300 K with SPC explicit water model [198] and the OPLS2005 all atom force field [199] is carried out starting with eight peptides assembled with hydrogen bonding per laminate, with totally five laminates of KLVFFAL peptides in SPC water. Initial structures were allowed to relax for 100 ps before the run. Boundary conditions were set to create infinite H-bonded β-sheets and infinite β-sheet stacking.

5.4.6 Retro-aldol Analysis

The kinetics of retro-aldol reaction are followed by fluorescence for product concentration and HPLC for substrate concentration. Right before the kinetic measurements, the peptide assemblies are centrifuged under 13,800 × g for 60 minutes and then are re-suspended with distilled water. The peptide solution is buffered to pH 7.5 with 50 mM phosphate and 300 mM NaCl, with 5% DMSO to enhance the solubility of methodol. After the methodol solution is mixed with the peptide solution, fluorescence measurements for the reaction mixture are conducted on a Synergy HT Multi-detection microplate reader. Triplicate samples of 200 μL are measured with excitation at 360 nm and fluorescence emission at 460 nm, and data are collected every 45 seconds, with the plates shaken for 20 revolutions per second.

The enantioselectivity of the peptide assemblies is determined by chiral HPLC analysis using a Chiralpak AD-RH (Daicel) column. After the peptide solution is mixed with the substrate methodol, the remaining R- and S-methodol are separated and quantified with the chiral HPLC kinetically. The elution time points for R- and S-methodol are 18.6 and 21.3 minutes, respectively, when the reaction mixture is eluted from HPLC isocratically with 40% ACN at a flow rate of 0.5 mL/min.

CHAPTER 6. CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

The main goal of this thesis is to provide a comprehensive study about building the peptide-based biomaterials. The crucial building blocks capable of assembly, the assembly pathway, and the catalytic function of the assemblies are analyzed experimentally and mathematically. The synthesis and selection of the building blocks are demonstrated via the dynamic combinatorial networks, which yields sequence-defined peptide oligomers. Next, the assembly mechanism is simulated with a two-step nucleation model, and the model is experimentally tested with a pH-sensitive peptide, Aβ(16-22); extension of this model for a polydisperse system provides insights about the physical and chemical phases. Given the structures studied, the catalytic function of the peptide nanotube assemblies is analyzed to reveal the underlying mechanism which selectively breaks the substrate.

The building blocks determine the properties of the assemblies. It is demonstrated in Chapter 2 that the peptide oligomers as the building blocks may be synthesized from a dynamic combinatorial network (DCN). Two DCNs are constructed with chemically modified peptides, *NF*-CHO and *NFF*-CHO. The two DCNs behave similarly and they undergo the same physical transitions. The metastable particle phase nucleate as the average oligomer length increases, and later the fibers nucleate inside these particles. The experiments reveal that the emergent physical phases cause the species redistribution, and the modeling simulation supports this interplay between the chemical and the physical phases. As both networks yield trimer with high fidelity, construction of the DCNs may become a new way to synthesize sequence-defined oligomers, the building blocks for the mesoscale materials.

Next, to understand the multi-stage assembly pathway observed in the DCN, a peptide assembly model with two-step nucleation is constructed for a monodisperse system. It addresses that the final thermodynamic state of the solution is determined by the solubilities of the particles and the assemblies. The phase with a higher solubility will diminish while the other survives. However, this process may be kinetically hindered with slow particle growth and dissolution. The utility of the monodisperse model is tested experimentally with the assembly kinetics of the Aβ(16-22) peptide. Aβ(16-22) assembles more quickly as the acetonitrile concentration increases, and this trend is rationally explained by the model that the kinetics is limited when the peptide solubility for the particles decreases. Next, the addition of the peptide oligomerization makes the model capable of simulating the polydisperse systems with reversible oligomerization. For the polydisperse model, the equilibrium of the chemical distribution (oligomer concentrations) and the physical distribution (overall oligomer mass in different physical phases) may be achieved independently. This suggests that once the desired chemical or physical distribution is achieved, the system may be harvested before the other phase become stationary.

The reaction pathway and the morphological selection are further investigated with the pH sensitive Aβ(16-22). The pH-dependent Aβ(16-22) assembly reveals the importance of the environments for assembly nucleation and growth. As Aβ(16-22) assembles with two-step nucleation, the assemblies nucleate in the particle while grow in the aqueous solution. At both acidic and neutral pH, Aβ(16-22) assembles into the metastable ribbons in the particles. The same intermediate products from both pH conditions suggest that the particles, where the ribbons nucleate, are similar and pH-independent. The metastable ribbons diminish after the other thermodynamically stable morphologies arise in the solvent, the propagation environment.

Finally, the catalytic function of the peptide assemblies is analyzed. The Ac-KLVFFAL-NH_2 and the Ac-OrnLVFFAL-NH_2 nanotubes catalyze a retro-aldol reaction on the methodol substrate, and the catalytic reaction progress is fitted by a modified Michaelis-Menten mechanism. The result suggests similar size of the binding pocket for both peptides. The consistent parameter set suggests that the peptide nanotubes are stable throughout the entire reaction time, and their chemical activity does not decay during the reaction. The enantioselectivity for the methodol substrate may come from the chemical selectivity of the peptide nanotubes, but not from the different binding activity, based on the analysis.

To fully understand and engineer the materials, their building block, construction pathway, and functions are all essential and should not be studied separately from each other. The above work provides a comprehensive board view about studying the construction of functional biomaterial by analyzing the building block, the assembly mechanism, and the catalytic function of the peptide assemblies. This achievement brings a broader view for studies to follow, which may make the peptide-based materials more viable and accessible.

6.2 Recommendations

6.2.1 Construction of DCNs with α-helical Structure

In the human body, collagen is the most abundant protein, which is constituted mainly by proline (P) and glycine (G). Each chain in collagen is about 1400 amino acids long, and three chains would coil together into a tight triple helix. Glycine-proline-X is a common repeating sequence in collagen, which results in the helical structures with intramolecular hydrogen bonds.

By mimicking the collagen structure, the *NGP*-CHO monomer may be used to achieve the α-helical structure to template the growth of longer oligomers via DCN. As the *NGP*-CHO monomer undergoes oligomerization, the oligomers may have propensity to assemble into the triple-helix structure, as the $[NGP]_n$ sequence is similar to the glycine-proline-X sequence observed in collagen. The end of the helical structure then may propagate by the incorporating more *NGP*-CHO molecules, and thus longer oligomers within the DCN may be synthesized.

6.2.2 NF-CHO Tetramer Formation from Templating

In the DCNs investigated in the thesis, both the *NF*-CHO and *NFF*-CHO trimers assemble into nanofibers with the β-sheet structure. As a result, the longest stable product detected in these two networks is trimer, while the other longer oligomers are not identified. The trimer assemblies may be kinetic products, and they predominate before the assemblies of other longer oligomers nucleate.

To test if the *NF*-CHO molecules may oligomerize into tetramer (4-mer), the *NF*-CHO network may be seeded with the assemblies of peptide NFNFNFNF, which might template the formation of tetramer and thus make the tetramer assemblies predominate, instead of the trimer assemblies. Given the tetramer assemblies built from seeding, the trimer assemblies are shown to be kinetic products only, while the tetramers are more thermodynamically stable.

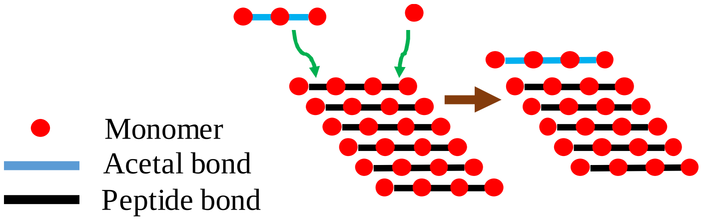

Figure 6.1– The templated formation of the *NF*-CHO tetramer via the (NF)$_4$ assemblies.

6.2.3 Peptide Assembly via Two-step Nucleation with Different Morphologies

In this thesis, it is simulated that the peptides assemble into linear fibers from spherical particles. The model may be modified to fit other systems with different morphologies. For example, β-lactoglobulin has an amorphous metastable phase before its crystal arise [64, 90]; the bis-FF dipeptide fibers are nucleated inside the particles, but those fibers further transform into tubes [71]. It requires different rate equations to express these transition from the metastable phase to the final phase, and the nucleation mechanism may also be different from what is discussed in this thesis. Addition of these different morphological combinations may make the model more robust to simulate the multi-step nucleation process.

6.2.4 Emergent Catalytic Assemblies from the Dynamic Combinatorial Networks

The ordered structures of the Ac-KLVFFAL-NH_2 and the Ac-OrnLVFFAL-NH_2 provide catalytic binding sites for a retro-aldol reaction. With the primary amine group on the tube surface, the peptide tubes are able to catalyse the retro-aldol reaction and cleave the methodol substrate.

As the primary amine group on the assembly surface is crucial for catalysis, the DCN building blocks may be designed as Ac-*KF*-CHO and *NFL*-NH_2 to access catalytic function. Ac-*KLVF*-CHO and *NAL*-NH_2 may dimerize into Ac-KLVFNAL-NH_2 with the reversible acetal linkage. If the dimers assemble into ordered structures similar to the Ac-KLVFFAL-NH_2 tubes, they may be able to catalyse the retro-aldol reaction, and become a DCN product from the DCN with emergent catalytic function.

6.2.5 Transition from Out-of-register Ribbons to In-register Fibers

When Aβ(16-22) is incubated at neutral pH, it assemble into metastable out-of-register ribbons, which then diminish and transition into the stable fiber with in-register β-sheet structure. However, the nucleation mechanism of the in-register structure remains to be clarified. The hypothesis for the place of fiber nucleation includes: 1) inside the particle, and 2) on the ribbon surface. If the fiber nucleates inside the particles, NMR may be used to probe the in-register structure in the particle. If the fiber nucleate on the ribbon surface with a secondary nucleation process, the out-of-register ribbon from Ac-KLVFFAL-NH_2 may be seeded into the Aβ(16-22), which should make the fibers nucleate more quickly given more nucleation sites. The results of these experiments may help to clarify the transition from the metastable product to the final stable fibers.

REFERENCES

1. Barz, B., Wales, D.J., and Strodel, B., *A Kinetic Approach to the Sequence-Aggregation Relationship in Disease-Related Protein Assembly.* Journal of Physical Chemistry B, 2014. **118**(4): p. 1003-1011.

2. Hardy, J. and Selkoe, D.J., *Medicine - the Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics.* Science, 2002. **297**(5580): p. 353-356.

3. Laganowsky, A., Liu, C., Sawaya, M.R., Whitelegge, J.P., Park, J., Zhao, M., Pensalfini, A., Soriaga, A.B., Landau, M., Teng, P.K., Cascio, D., Glabe, C., and Eisenberg, D., *Atomic View of a Toxic Amyloid Small Oligomer.* Science, 2012. **335**(6073): p. 1228-1231.

4. Li, X.L., James, S., and Lei, P., *Interactions between Alpha-Synuclein and Tau Protein: Implications to Neurodegenerative Disorders.* Journal of Molecular Neuroscience, 2016. **60**(3): p. 298-304.

5. Meinke, J.H. and Hansmann, U.H.E., *Aggregation of Beta-Amyloid Fragments.* The Journal of Chemical Physics, 2007. **126**(1): p. 014706-5.

6. Kodali, R. and Wetzel, R., *Polymorphism in the Intermediates and Products of Amyloid Assembly.* Current Opinion in Structural Biology, 2007. **17**(1): p. 48-57.

7. Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G., *Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis.* Science, 2003. **300**(5618): p. 486-489.

8. Caughey, B. and Lansbury, P.T., *Protofibrils, Pores, Fibrils, and Neurodegeneration: Separating the Responsible Protein Aggregates from the Innocent Bystanders.* Annual Review of Neuroscience, 2003. **26**: p. 267-298.

9. Hwang, W., Zhang, S.G., Kamm, R.D., and Karplus, M., *Kinetic Control of Dimer Structure Formation in Amyloid Fibrillogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(35): p. 12916-12921.

10. Mehta, A.K., Rosen, R.F., Childers, W.S., Gehman, J.D., Walker, L.C., and Lynn, D.G., *Context Dependence of Protein Misfolding and Structural Strains in Neurodegenerative Diseases.* Biopolymers, 2013. **100**(6): p. 722-730.

11. Lee, J., Culyba, E.K., Powers, E.T., and Kelly, J.W., *Amyloid-Beta Forms Fibrils by Nucleated Conformational Conversion of Oligomers.* Nature Chemical Biology, 2011. **7**(9): p. 602-609.

12. Yan, X.H., He, Q., Wang, K.W., Duan, L., Cui, Y., and Li, J.B., *Transition of Cationic Dipeptide Nanotubes into Vesicles and Oligonucleotide Delivery.* Angewandte Chemie International Edition, 2007. **46**(14): p. 2431-2434.

13. Nagai, Y., Unsworth, L.D., Koutsopoulos, S., and Zhang, S.G., *Slow Release of Molecules in Self-Assembling Peptide Nanofiber Scaffold.* Journal of Controlled Release, 2006. **115**(1): p. 18-25.

14. Habibi, N., Kamaly, N., Memic, A., and Shafiee, H., *Self-Assembled Peptide-Based Nanostructures: Smart Nanomaterials toward Targeted Drug Delivery.* Nano Today, 2016. **11**(1): p. 41-60.

15. Pawar, R., Ben-Ari, A., and Domb, A.J., *Protein and Peptide Parental Controlled Delivery.* Expert Opinion on Biological Therapy, 2004. **4**(8): p. 1203-1212.

16. Reches, M. and Gazit, E., *Casting Metal Nanowires within Discrete Self-Assembled Peptide Nanotubes.* Science, 2003. **300**(5619): p. 625-627.

17. Huang, J., Lin, L., Sun, D., Chen, H., Yang, D., and Li, Q., *Bio-Inspired Synthesis of Metal Nanomaterials and Applications.* Chemical Society Reviews, 2015. **44**(17): p. 6330-6374.

18. Singh, N., Conte, M.P., Ulijn, R.V., Miravet, J.F., and Escuder, B., *Insight into the Esterase Like Activity Demonstrated by an Imidazole Appended Self-Assembling Hydrogelator.* Chemical Communications, 2015. **51**(67): p. 13213-13216.

19. Cohen, F.E. and Kelly, J.W., *Therapeutic Approaches to Protein-Misfolding Diseases.* Nature, 2003. **426**(6968): p. 905-909.

20. Spillantini, M.G., Schmidt, M.L., Lee, V.M.Y., Trojanowski, J.Q., Jakes, R., and Goedert, M., *Alpha-Synuclein in Lewy Bodies.* Nature, 1997. **388**(6645): p. 839-840.

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

22. Dadon, Z., Samiappan, M., Wagner, N., and Ashkenasy, G., *Chemical and Light Triggering of Peptide Networks under Partial Thermodynamic Control.* Chemical Communications, 2012. **48**(10): p. 1419-1421.

23. Lad, C., Williams, N.H., and Wolfenden, R., *The Rate of Hydrolysis of Phosphomonoester Dianions and the Exceptional Catalytic Proficiencies of*

Protein and Inositol Phosphatases. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(10): p. 5607-5610.

24. Ringe, D. and Petsko, G.A., *How Enzymes Work.* Science, 2008. **320**(5882): p. 1428-1429.

25. Breslow, R. and Dong, S.D., *Biomimetic Reactions Catalyzed by Cyclodextrins and Their Derivatives.* Chemical Reviews, 1998. **98**(5): p. 1997-2011.

26. Qi, D.F., Tann, C.M., Haring, D., and Distefano, M.D., *Generation of New Enzymes Via Covalent Modification of Existing Proteins.* Chemical Reviews, 2001. **101**(10): p. 3081-3111.

27. Guler, M.O. and Stupp, S.I., *A Self-Assembled Nanofiber Catalyst for Ester Hydrolysis.* Journal of the American Chemical Society, 2007. **129**(40): p. 12082-12083.

28. Rodriguez-Llansola, F., Escuder, B., and Miravet, J.F., *Remarkable Increase in Basicity Associated with Supramolecular Gelation.* Organic & Biomolecular Chemistry, 2009. **7**(15): p. 3091-3094.

29. Huang, Z., Guan, S., Wang, Y., Shi, G., Cao, L., Gao, Y., Dong, Z., Xu, J., Luo, Q., and Liu, J., *Self-Assembly of Amphiphilic Peptides into Bio-Functionalized Nanotubes: A Novel Hydrolase Model.* Journal of Materials Chemistry B, 2013. **1**(17): p. 2297-2304.

30. Jin, Q., Zhang, L., Cao, H., Wang, T., Zhu, X., Jiang, J., and Liu, M., *Self-Assembly of Copper(II) Ion-Mediated Nanotube and Its Supramolecular Chiral Catalytic Behavior.* Langmuir, 2011. **27**(22): p. 13847-13853.

31. Djalali, R., Chen, Y., and Matsui, H., *Au Nanowire Fabrication from Sequenced Histidine-Rich Peptide.* Journal of the American Chemical Society, 2002. **124**(46): p. 13660-13661.

32. Matsui, H., Pan, S., Gologan, B., and Jonas, S.H., *Bolaamphiphile Nanotube-Templated Metallized Wires.* Journal of Physical Chemistry B, 2000. **104**(41): p. 9576-9579.

33. Guler, M.O., Claussen, R.C., and Stupp, S.I., *Encapsulation of Pyrene within Self-Assembled Peptide Amphiphile Nanofibers.* Journal of Materials Chemistry, 2005. **15**(42): p. 4507-4512.

34. Keller, S., Sauer, I., Strauss, H., Gast, K., Dathe, M., and Bienert, M., *Membrane-Mimetic Nanocarriers Formed by a Dipalmitoylated Cell-Penetrating Peptide.* Angewandte Chemie-International Edition, 2005. **44**(33): p. 5252-5255.

35. van Hell, A.J., Costa, C., Flesch, F.M., Sutter, M., Jiskoot, W., Crommelin, D.J.A., Hennink, W.E., and Mastrobattista, E., *Self-Assembly of Recombinant*

Amphiphilic Oligopeptides into Vesicles. Biomacromolecules, 2007. **8**(9): p. 2753-2761.

36. Yishay-Safranchik, E., Golan, M., and David, A., *Controlled Release of Doxorubicin and Smac-Derived Pro-Apoptotic Peptide from Self-Assembled KLD-Based Peptide Hydrogels.* Polymers for Advanced Technologies, 2014. **25**(5): p. 539-544.

37. Silva, R.F., Araujo, D.R., Silva, E.R., Ando, R.A., and Alves, W.A., *L-Diphenylalanine Microtubes as a Potential Drug-Delivery System: Characterization, Release Kinetics, and Cytotoxicity.* Langmuir, 2013. **29**(32): p. 10205-10212.

38. Kokkoli, E., Mardilovich, A., Wedekind, A., Rexeisen, E.L., Garg, A., and Craig, J.A., *Self-Assembly and Applications of Biomimetic and Bioactive Peptide-Amphiphiles.* Soft Matter, 2006. **2**(12): p. 1015-1024.

39. Shi, J.J., Xiao, Z.Y., Kamaly, N., and Farokhzad, O.C., *Self-Assembled Targeted Nanoparticles: Evolution of Technologies and Bench to Bedside Translation.* Accounts of Chemical Research, 2011. **44**(10): p. 1123-1134.

40. Vauthey, S., Santoso, S., Gong, H.Y., Watson, N., and Zhang, S.G., *Molecular Self-Assembly of Surfactant-Like Peptides to Form Nanotubes and Nanovesicles.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(8): p. 5355-5360.

41. Xu, H., Wang, J., Han, S., Wang, J., Yu, D., Zhang, H., Xia, D., Zhao, X., Waigh, T.A., and Lu, J.R., *Hydrophobic-Region-Induced Transitions in Self-Assembled Peptide Nanostructures.* Langmuir, 2009. **25**(7): p. 4115-4123.

42. Yoon, Y.-R., Lim, Y.-b., Lee, E., and Lee, M., *Self-Assembly of a Peptide Rod-Coil: A Polyproline Rod and a Cell-Penetrating Peptide Tat Coil.* Chemical Communications, 2008(16): p. 1892-1894.

43. Lee, N.R., Bowerman, C.J., and Nilsson, B.L., *Effects of Varied Sequence Pattern on the Self-Assembly of Amphipathic Peptides.* Biomacromolecules, 2013. **14**(9): p. 3267-3277.

44. Israelachvili, J.N., *Intermolecular and Surface Forces, 3rd Edition*. Intermolecular and Surface Forces, 3rd Edition. 2011. 1-674.

45. Smith, J.E., Liang, C., Tseng, M., Li, N., Li, S., Mowles, A.K., Mehta, A.K., and Lynn, D.G., *Defining the Dynamic Conformational Networks of Cross- Peptide Assembly.* Israel Journal of Chemistry, 2015. **55**(6-7): p. 763-769.

46. Yu, T., Lee, O.S., and Schatz, G.C., *Steered Molecular Dynamics Studies of the Potential of Mean Force for Peptide Amphiphile Self-Assembly into Cylindrical Nanofibers.* Journal of Physical Chemistry A, 2013. **117**(32): p. 7453-7460.

47. Wallace, J.A. and Shen, J.K., *Probing the Strand Orientation and Registry Alignment in the Propagation of Amyloid Fibrils.* Biochemistry, 2010. **49**(25): p. 5290-5298.

48. Dong, H., Paramonov, S.E., Aulisa, L., Bakota, E.L., and Hartgerink, J.D., *Self-Assembly of Multidomain Peptides: Balancing Molecular Frustration Controls Conformation and Nanostructure.* Journal of the American Chemical Society, 2007. **129**(41): p. 12468-12472.

49. Fan, Z. and Govorov, A.O., *Plasmonic Circular Dichroism of Chiral Metal Nanoparticle Assemblies.* Nano Letters, 2010. **10**(7): p. 2580-2587.

50. Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Refregiers, M., and Kardos, J., *Accurate Secondary Structure Prediction and Fold Recognition for Circular Dichroism Spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(24): p. E3095-E3103.

51. Bartolini, M., Bertucci, C., Cavrini, V., and Andrisano, V., *Beta-Amyloid Aggregation Induced by Human Acetylcholinesterase: Inhibition Studies.* Biochemical Pharmacology, 2003. **65**(3): p. 407-416.

52. Moran, S.D. and Zanni, M.T., *How to Get Insight into Amyloid Structure and Formation from Infrared Spectroscopy.* Journal of Physical Chemistry Letters, 2014. **5**(11): p. 1984-1993.

53. Krimm, S. and Abe, Y., *Intermolecular Interaction Effects in Amide I Vibrations of Beta Polypeptides.* Proceedings of the National Academy of Sciences of the United States of America, 1972. **69**(10): p. 2788-2792.

54. Moore, W.H. and Krimm, S., *Transition Dipole Coupling in Amide I Modes of Beta Polypeptides.* Proceedings of the National Academy of Sciences of the United States of America, 1975. **72**(12): p. 4933-4935.

55. Chirgadze, Y.N. and Nevskaya, N.A., *Infrared-Spectra and Resonance Interaction of Amide-One Vibration of Parallel-Chain Pleated Sheet.* Biopolymers, 1976. **15**(4): p. 627-636.

56. Toniolo, C. and Palumbo, M., *Solid-State Infrared-Absorption Spectra and Chain Arrangement in Some Synthetic Homooligopeptides in Intermolecularly Hydrogen-Bonded Pleated-Sheet Beta-Conformation.* Biopolymers, 1977. **16**(1): p. 219-224.

57. Goormaghtigh, E., Cabiaux, V., and Ruysschaert, J.M., *Determination of Soluble and Membrane Protein Structure by Fourier Transform Infrared Spectroscopy. III. Secondary Structures.* Sub-cellular biochemistry, 1994. **23**: p. 405-50.

58. Paul, C. and Axelsen, P.H., *Beta-Sheet Structure in Amyloid Beta Fibrils and Vibrational Dipolar Coupling.* Journal of the American Chemical Society, 2005. **127**(16): p. 5754-5755.

59. Sarroukh, R., Goormaghtigh, E., Ruysschaert, J.M., and Raussens, V., *ATR-FTIR: A "Rejuvenated" Tool to Investigate Amyloid Proteins.* Biochimica Et Biophysica Acta-Biomembranes, 2013. **1828**(10): p. 2328-2338.

60. Liang, C., Ni, R., Smith, J.E., Childers, W.S., Mehta, A.K., and Lynn, D.G., *Kinetic Intermediates in Amyloid Assembly.* Journal of the American Chemical Society, 2014. **136**(43): p. 15146-15149.

61. Moran, S.D., Woys, A.M., Buchanan, L.E., Bixby, E., Decatur, S.M., and Zanni, M.T., *Two-Dimensional IR Spectroscopy and Segmental C-13 Labeling Reveals the Domain Structure of Human Gamma D-Crystallin Amyloid Fibrils.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(9): p. 3329-3334.

62. Mora, A.K., Singh, P.K., Patro, B.S., and Nath, S., *Picogreen: A Better Amyloid Probe Than Thioflavin-T.* Chemical Communications, 2016. **52**(82): p. 12163-12166.

63. Liang, Y., Lynn, D.G., and Berland, K.M., *Direct Observation of Nucleation and Growth in Amyloid Self-Assembly.* J Am Chem Soc, 2010. **132**(18): p. 6306-6308.

64. Sauter, A., Roosen-Runge, F., Zhang, F.J., Lotze, G., Jacobs, R.M.J., and Schreiber, F., *Real-Time Observation of Nonclassical Protein Crystallization Kinetics.* Journal of the American Chemical Society, 2015. **137**(4): p. 1485-1491.

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

66. Morris, A.M., Watzky, M.A., Agar, J.N., and Finke, R.G., *Fitting Neurological Protein Aggregation Kinetic Data Via a 2-Step, Minimal "Ockham's Razor" Model: The Finke-Watzky Mechanism of Nucleation Followed by Autocatalytic Surface Growth.* Biochemistry, 2008. **47**(8): p. 2413-2427.

67. Arosio, P., Beeg, M., Nicoud, L., and Morbidelli, M., *Time Evolution of Amyloid Fibril Length Distribution Described by a Population Balance Model.* Chemical Engineering Science, 2012. **78**: p. 21-32.

68. Cohen, S.I.A., Vendruscolo, M., Welland, M.E., Dobson, C.M., Terentjev, E.M., and Knowles, T.P.J., *Nucleated Polymerization with Secondary Pathways. I. Time Evolution of the Principal Moments.* Journal of Chemical Physics, 2011. **135**(6): p. 065105.

69. Vekilov, P.G., *The Two-Step Mechanism of Nucleation of Crystals in Solution.* Nanoscale, 2010. **2**(11): p. 2346-2357.

70. ten Wolde, P.R. and Frenkel, D., *Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations.* Science, 1997. **277**(5334): p. 1975-1978.

71. Levin, A., Mason, T.O., Adler-Abramovich, L., Buell, A.K., Meisl, G., Galvagnion, C., Bram, Y., Stratford, S.A., Dobson, C.M., Knowles, T.P.J., and Gazit, E., *Ostwald's Rule of Stages Governs Structural Transitions and Morphology of Dipeptide Supramolecular Polymers.* Nature Communications, 2014. **5**: p. 5219.

72. Auer, S., Ricchiuto, P., and Kashchiev, D., *Two-Step Nucleation of Amyloid Fibrils: Omnipresent or Not?* Journal of Molecular Biology, 2012. **422**(5): p. 723-730.

73. Lomakin, A., Chung, D.S., Benedek, G.B., Kirschner, D.A., and Teplow, D.B., *On the Nucleation and Growth of Amyloid Beta-Protein Fibrils: Detection of Nuclei and Quantitation of Rate Constants.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(3): p. 1125-1129.

74. Krishnan, R. and Lindquist, S.L., *Structural Insights into a Yeast Prion Illuminate Nucleation and Strain Diversity.* Nature, 2005. **435**(7043): p. 765-772.

75. Luiken, J.A. and Bolhuis, P.G., *Primary Nucleation Kinetics of Short Fibril-Forming Amyloidogenic Peptides.* Journal of Physical Chemistry B, 2015. **119**(39): p. 12568-12579.

76. Liu, Y., Wang, X., and Ching, C.B., *Toward Further Understanding of Lysozyme Crystallization: Phase Diagram, Protein-Protein Interaction, Nucleation Kinetics, and Growth Kinetics.* Crystal Growth & Design, 2010. **10**(2): p. 548-558.

77. Lin, D.D., Luo, Y., Wu, S., Ma, Q.Q., Wei, G.H., and Yang, X.J., *Investigation of the Aggregation Process of Amyloid-Beta-(16-22) Peptides and the Dissolution of Intermediate Aggregates.* Langmuir, 2014. **30**(11): p. 3170-3175.

78. Lomakin, A., Asherie, N., and Benedek, G.B., *Liquid-Solid Transition in Nuclei of Protein Crystals.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(18): p. 10254-10257.

79. Vivares, D., Kaler, E.W., and Lenhoff, A.M., *Quantitative Imaging by Confocal Scanning Fluorescence Microscopy of Protein Crystallization Via Liquid-Liquid Phase Separation.* Acta Crystallographica Section D: Biological Crystallography, 2005. **61**: p. 819-825.

80. Gliko, O., Neumaier, N., Pan, W., Haase, I., Fischer, M., Bacher, A., Weinkauf, S., and Vekilov, P.G., *A Metastable Prerequisite for the Growth of Lumazine*

Synthase Crystals. Journal of the American Chemical Society, 2005. **127**(10): p. 3433-3438.

81. Pan, W., Galkin, O., Filobelo, L., Nagel, R.L., and Vekilov, P.G., *Metastable Mesoscopic Clusters in Solutions of Sickle-Cell Hemoglobin.* Biophysical Journal, 2007. **92**(1): p. 267-277.

82. Kuznetsov, Y.G., Malkin, A.J., and McPherson, A., *The Liquid Protein Phase in Crystallization: A Case Study - Intact Immunoglobulins.* Journal of Crystal Growth, 2001. **232**(1-4): p. 30-39.

83. Vorontsova, M.A., Maes, D., and Vekilov, P.G., *Recent Advances in the Understanding of Two-Step Nucleation of Protein Crystals.* Faraday Discussions, 2015. **179**: p. 27-40.

84. Schope, H.J., Bryant, G., and van Megen, W., *Two-Step Crystallization Kinetics in Colloidal Hard-Sphere Systems.* Physical Review Letters, 2006. **96**(17): p. 175701.

85. Anderson, V.J. and Lekkerkerker, H.N.W., *Insights into Phase Transition Kinetics from Colloid Science.* Nature, 2002. **416**(6883): p. 811-815.

86. Savage, J.R. and Dinsmore, A.D., *Experimental Evidence for Two-Step Nucleation in Colloidal Crystallization.* Physical Review Letters, 2009. **102**(19): p. 198302.

87. Rodriguez-Navarro, C., Kudlacz, K., Cizer, O., and Ruiz-Agudo, E., *Formation of Amorphous Calcium Carbonate and Its Transformation into Mesostructured Calcite.* Crystengcomm, 2015. **17**(1): p. 58-72.

88. Nielsen, M.H., Aloni, S., and De Yoreo, J.J., *In Situ Tem Imaging of* $CaCO_3$ *Nucleation Reveals Coexistence of Direct and Indirect Pathways.* Science, 2014. **345**(6201): p. 1158-1162.

89. Wang, J., Mueller, M., and Wang, Z.-G., *Nucleation in A/B/AB Blends: Interplay between Microphase Assembly and Macrophase Separation.* Journal of Chemical Physics, 2009. **130**(15): p. 154902.

90. Sauter, A., Roosen-Runge, F., Zhang, F.J., Lotze, G., Feoktystov, A., Jacobs, R.M.J., and Schreiber, F., *On the Question of Two-Step Nucleation in Protein Crystallization.* Faraday Discussions, 2015. **179**: p. 41-58.

91. Debeljuh, N., Barrow, C.J., and Byrne, N., *The Impact of Ionic Liquids on Amyloid Fibrilization of Aβ16-22: Tuning the Rate of Fibrilization Using a Reverse Hofmeister Strategy.* Physical Chemistry Chemical Physics, 2011. **13**(37): p. 16534-16536.

92. Tomiyama, T., Matsuyama, S., Iso, H., Umeda, T., Takuma, H., Ohnishi, K., Ishibashi, K., Teraoka, R., Sakama, N., Yamashita, T., Nishitsuji, K., Ito, K., Shimada, H., Lambert, M.P., Klein, W.L., and Mori, H., *A Mouse Model of Amyloid Beta Oligomers: Their Contribution to Synaptic Alteration, Abnormal Tau Phosphorylation, Glial Activation, and Neuronal Loss in Vivo.* Journal of Neuroscience, 2010. **30**(14): p. 4845-4856.

93. Xue, W.F., Homans, S.W., and Radford, S.E., *Systematic Analysis of Nucleation-Dependent Polymerization Reveals New Insights into the Mechanism of Amyloid Self-Assembly.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(26): p. 8926-8931.

94. Andrews, J.M. and Roberts, C.J., *A Lumry-Eyring Nucleated Polymerization Model of Protein Aggregation Kinetics: 1. Aggregation with Pre-Equilibrated Unfolding.* Journal of Physical Chemistry B, 2007. **111**(27): p. 7897-7913.

95. Pan, W.C., Kolomeisky, A.B., and Vekilov, P.G., *Nucleation of Ordered Solid Phases of Proteins Via a Disordered High-Density State: Phenomenological Approach.* Journal of Chemical Physics, 2005. **122**(17): p. 174905.

96. Aich, A., Pan, W.C., and Vekilov, P.G., *Thermodynamic Mechanism of Free Heme Action on Sickle Cell Hemoglobin Polymerization.* Aiche Journal, 2015. **61**(9): p. 2861-2870.

97. Kashchiev, D., Vekilov, P.G., and Kolomeisky, A.B., *Kinetics of Two-Step Nucleation of Crystals.* Journal of Chemical Physics, 2005. **122**(24): p. 244706.

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

99. Otto, S., *Dynamic Combinatorial Chemistry: A New Method for Selection and Preparation of Synthetic Receptors.* Current Opinion in Drug Discovery & Development, 2003. **6**(4): p. 509-520.

100. Otto, S., *Dynamic Molecular Networks: From Synthetic Receptors to Self-Replicators.* Accounts of Chemical Research, 2012. **45**(12): p. 2200-2210.

101. Rowan, S.J., Cantrill, S.J., Cousins, G.R.L., Sanders, J.K.M., and Stoddart, J.F., *Dynamic Covalent Chemistry.* Angewandte Chemie-International Edition, 2002. **41**(6): p. 898-952.

102. Goodwin, J.T. and Lynn, D.G., *Template-Directed Synthesis - Use of a Reversible-Reaction.* Journal of the American Chemical Society, 1992. **114**(23): p. 9197-9198.

103. Star, A., Goldberg, I., and Fuchs, B., *Dioxadiazadecalin/Salen Tautomeric Macrocycles and Complexes: Prototypal Dynamic Combinatorial Virtual*

Libraries. Angewandte Chemie-International Edition, 2000. **39**(15): p. 2685-2689.

104. Cacciapaglia, R., Di Stefano, S., and Mandolini, L., *Metathesis Reaction of Formaldehyde Acetals: An Easy Entry into the Dynamic Covalent Chemistry of Cyclophane Formation.* Journal of the American Chemical Society, 2005. **127**(39): p. 13666-13671.

105. Godoy-Alcantar, C., Yatsimirsky, A.K., and Lehn, J.M., *Structure-Stability Correlations for Imine Formation in Aqueous Solution.* Journal of Physical Organic Chemistry, 2005. **18**(10): p. 979-985.

106. Giuseppone, N. and Lehn, J.M., *Protonic and Temperature Modulation of Constituent Expression by Component Selection in a Dynamic Combinatorial Library of Imines.* Chemistry-a European Journal, 2006. **12**(6): p. 1715-1722.

107. Zameo, S., Vauzeilles, B., and Beau, J.M., *Direct Composition Analysis of a Dynamic Library of Imines in an Aqueous Medium.* European Journal of Organic Chemistry, 2006(24): p. 5441-5444.

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

109. Giuseppone, N. and Lehn, J.M., *Constitutional Dynamic Self-Sensing in a Zinc(II)/Polyiminofluorenes System.* Journal of the American Chemical Society, 2004. **126**(37): p. 11448-11449.

110. Eliel, E.L. and Wilen, S.H., *Stereochemistry of Organic Compounds*. 1994, New York: Wiley.

111. Star, A., Goldberg, I., and Fuchs, B., *Diazadioxadecalin and Salen Podands and Macrocycles within Dynamic Combinatorial Virtual Libraries: Structure, Prototropy, Complexation and Enantioselective Catalysis.* Journal of Organometallic Chemistry, 2001. **630**(1): p. 67-77.

112. Sutton, L.R., Donaubauer, W.A., Hampel, F., and Hirsch, A., *Tris(Thioacetals) from Benzene Hexathiol: Towards Covalent Self-Assembly.* Chemical Communications, 2004(15): p. 1758-1759.

113. Wipf, P., Mahler, S.G., and Okumura, K., *Metathesis Reactions of Pyrazolotriazinones Generate Dynamic Combinatorial Libraries.* Organic Letters, 2005. **7**(20): p. 4483-4486.

114. Takahashi, Y. and Mihara, H., *Construction of a Chemically and Conformationally Self-Replicating System of Amyloid-Like Fibrils.* Bioorganic & Medicinal Chemistry, 2004. **12**(4): p. 693-699.

115. Bourbo, V., Matmor, M., Shtelman, E., Rubinov, B., Ashkenasy, N., and Ashkenasy, G., *Self-Assembly and Self-Replication of Short Amphiphilic Beta-Sheet Peptides.* Origins of Life and Evolution of Biospheres, 2011. **41**(6): p. 563-567.

116. Shi, B.L., Stevenson, R., Campopiano, D.J., and Greaney, M.F., *Discovery of Glutathione S-Transferase Inhibitors Using Dynamic Combinatorial Chemistry.* Journal of the American Chemical Society, 2006. **128**(26): p. 8459-8467.

117. Rodriguez-Docampo, Z., Eugenieva-Ilieva, E., Reyheller, C., Belenguer, A.M., Kubik, S., and Otto, S., *Dynamic Combinatorial Development of a Neutral Synthetic Receptor That Binds Sulfate with Nanomolar Affinity in Aqueous Solution.* Chemical Communications, 2011. **47**(35): p. 9798-9800.

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

119. Brady, P.A. and Sanders, J.K.M., *Selection Approaches to Catalytic Systems.* Chemical Society Reviews, 1997. **26**(5): p. 327-336.

120. Sanders, J.K.M., *Supramolecular Catalysis in Transition.* Chemistry-a European Journal, 1998. **4**(8): p. 1378-1383.

121. Brisig, B., Sanders, J.K.M., and Otto, S., *Selection and Amplification of a Catalyst from a Dynamic Combinatorial Library.* Angewandte Chemie-International Edition, 2003. **42**(11): p. 1270-1273.

122. Vial, L., Sanders, J.K.M., and Otto, S., *A Catalyst for an Acetal Hydrolysis Reaction from a Dynamic Combinatorial Library.* New Journal of Chemistry, 2005. **29**(8): p. 1001-1003.

123. Lutz, J.-F., *Sequence-Controlled Polymerizations: The Next Holy Grail in Polymer Science?* Polymer Chemistry, 2010. **1**(1): p. 55-62.

124. Miao, X., Cao, W., Zheng, W., Wang, J., Zhang, X., Gao, J., Yang, C., Kong, D., Xu, H., Wang, L., and Yang, Z., *Switchable Catalytic Activity: Selenium-Containing Peptides with Redox-Controllable Self-Assembly Properties.* Angewandte Chemie International Edition, 2013. **52**(30): p. 7781-7785.

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

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

127. Fife, T.H. and Hutchins, J.E.C., *General-Acid-Catalyzed Ring Opening of Oxazolidines. Hydrolysis of 2-[4-(Dimethylamino)Styryl]-N-Phenyl-1,3-Oxazolidine.* J. Org. Chem, 1980. **45**: p. 2099-2104.

128. Hurvich, C.M. and Tsai, C.L., *Regression and Time-Series Model Selection in Small Samples.* Biometrika, 1989. **76**(2): p. 297-307.

129. Wagenmakers, E.J. and Farrell, S., *AIC Model Selection Using Akaike Weights.* Psychonomic Bulletin & Review, 2004. **11**(1): p. 192-196.

130. Iggland, M. and Mazzotti, M., *Population Balance Modeling with Size-Dependent Solubility: Ostwald Ripening.* Crystal Growth & Design, 2012. **12**(3): p. 1489-1500.

131. Li, Y., Lubchenko, V., Vorontsova, M.A., Filobelo, L., and Vekilov, P.G., *Ostwald-Like Ripening of the Anomalous Mesoscopic Clusters in Protein Solutions.* Journal of Physical Chemistry B, 2012. **116**(35): p. 10657-10664.

132. Chen, C., Tan, J., Hsieh, M.-C., Pan, T., Goodwin, J.T., Mehta, A.K., Grover, M.A., and Lynn, D.G., *Design of Multi-Phase Dynamic Chemical Networks.* Nat Chem, 2017. **In press**: p. 10.1038/nchem.2737.

133. Chen, Q.Y., Vekilov, P.G., Nagel, R.L., and Hirsch, R.E., *Liquid-Liquid Phase Separation in Hemoglobins: Distinct Aggregation Mechanisms of the Beta 6 Mutants.* Biophysical Journal, 2004. **86**(3): p. 1702-1712.

134. Galkin, O., Chen, K., Nagel, R.L., Hirsch, R.E., and Vekilov, P.G., *Liquid-Liquid Separation in Solutions of Normal and Sickle Cell Hemoglobin.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(13): p. 8479-8483.

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

136. McKay, M.D., Beckman, R.J., and Conover, W.J., *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code.* Technometrics, 1979. **21**(2): p. 239-245.

137. Kashchiev, D. and van Rosmalen, G.M., *Review: Nucleation in Solutions Revisited.* Crystal Research and Technology, 2003. **38**(7-8): p. 555-574.

138. Tauer, K. and Kuhn, I., *Modeling Particle Formation in Emulsion Polymerization - an Approach by Means of the Classical Nucleation Theory.* Macromolecules, 1995. **28**(7): p. 2236-2239.

139. Flory, P.I., *Thermodynamics of High Polymer Solutions.* Journal of Chemical Physics, 1942. **10**(1): p. 51-61.

140. Flory, P.I., *Principles of Polymer Chemistry.* Cornell University Press, Ithaca, N.Y., 1953.

141. Roelands, C.P.M., ter Horst, J.H., Kramer, H.J.M., and Jansens, P.J., *Analysis of Nucleation Rate Measurements in Precipitation Processes.* Crystal Growth & Design, 2006. **6**(6): p. 1380-1392.

142. Madras, G. and McCoy, B.J., *Reversible Crystal Growth-Dissolution and Aggregation-Breakage: Numerical and Moment Solutions for Population Balance Equations.* Powder Technology, 2004. **143**: p. 297-307.

143. Vetter, T., Iggland, M., Ochsenbein, D.R., Hanseler, F.S., and Mazzotti, M., *Modeling Nucleation, Growth, and Ostwald Ripening in Crystallization Processes: A Comparison between Population Balance and Kinetic Rate Equation.* Crystal Growth & Design, 2013. **13**(11): p. 4890-4905.

144. Lindenberg, C. and Mazzotti, M., *Effect of Temperature on the Nucleation Kinetics of Alpha L-Glutamic Acid.* Journal of Crystal Growth, 2009. **311**(4): p. 1178-1184.

145. Matthews, H.B., Miller, S.M., and Rawlings, J.B., *Model Identification for Crystallization: Theory and Experimental Verification.* Powder Technology, 1996. **88**(3): p. 227-235.

146. Gherras, N. and Fevotte, G., *On the Use of Process Analytical Technologies and Population Balance Equations for the Estimation of Crystallization Kinetics. A Case Study.* Aiche Journal, 2012. **58**(9): p. 2650-2664.

147. Kashchiev, D. and Sato, K., *Kinetics of Crystallization Preceded by Metastable-Phase Formation.* Journal of Chemical Physics, 1998. **109**(19): p. 8530-8540.

148. Anthony, N.R., Mehta, A.K., Lynn, D.G., and Berland, K.M., *Mapping Amyloid-Beta(16-22) Nucleation Pathways Using Fluorescence Lifetime Imaging Microscopy.* Soft Matter, 2014. **10**(23): p. 4162-4172.

149. Dong, J.J., Shokes, J.E., Scott, R.A., and Lynn, D.G., *Modulating Amyloid Self-Assembly and Fibril Morphology with Zn(II).* Journal of the American Chemical Society, 2006. **128**(11): p. 3540-3542.

150. Tkachenko, A.V. and Maslov, S., *Spontaneous Emergence of Autocatalytic Information-Coding Polymers.* Journal of Chemical Physics, 2015. **143**(4): p. 045102.

151. Kamihira, M., Naito, A., Tuzi, S., Nosaka, A.Y., and Saito, H., *Conformational Transitions and Fibrillation Mechanism of Human Calcitonin as Studied by High-Resolution Solid-State C-13 NMR.* Protein Science, 2000. **9**(5): p. 867-877.

152. Lomakin, A., Teplow, D.B., Kirschner, D.A., and Benedek, G.B., *Kinetic Theory of Fibrillogenesis of Amyloid Beta-Protein.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(15): p. 7942-7947.

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

154. Tao, K., Wang, J., Zhou, P., Wang, C., Xu, H., Zhao, X., and Lu, J.R., *Self-Assembly of Short Abeta(16-22) Peptides: Effect of Terminal Capping and the Role of Electrostatic Interaction.* Langmuir, 2011. **27**(6): p. 2723-2730.

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

156. Mamajanov, I., MacDonald, P.J., Ying, J.Y., Duncanson, D.M., Dowdy, G.R., Walker, C.A., Engelhart, A.E., Fernandez, F.M., Grover, M.A., Hud, N.V., and Schork, F.J., *Ester Formation and Hydrolysis During Wet-Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction.* Macromolecules, 2014. **47**(4): p. 1334-1343.

157. Abed, S., Boileau, S., and Bouteiller, L., *Supramolecular Association of Acid-Terminated Poly(Dimethylsiloxane). 2. Molecular Weight Distributions.* Macromolecules, 2000. **33**(22): p. 8479-8487.

158. Ahmed, S.F. and Poehlein, G.W., *Kinetics of Dispersion Polymerization of Styrene in Ethanol .1. Model Development.* Industrial & Engineering Chemistry Research, 1997. **36**(7): p. 2597-2604.

159. Yasuda, M., Yokoyama, H., Seki, H., Ogino, H., Ishimi, K., and Ishikawa, H., *Simulation of Particle Growth in the Dispersion Polymerization of Styrene: The Termination Rate Constant in Particles.* Macromolecular Theory and Simulations, 2001. **10**(1): p. 54-62.

160. Freire, F., Almeida, A.M., Fisk, J.D., Steinkruger, J.D., and Gellman, S.H., *Impact of Strand Length on the Stability of Parallel-Beta-Sheet Secondary Structure.* Angewandte Chemie-International Edition, 2011. **50**(37): p. 8735-8738.

161. Stanger, H.E., Syud, F.A., Espinosa, J.F., Giriatt, I., Muir, T., and Gellman, S.H., *Length-Dependent Stability and Strand Length Limits in Antiparallel Beta-Sheet Secondary Structure.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(21): p. 12015-12020.

162. Wang, L. and Broadbelt, L.J., *Tracking Explicit Chain Sequence in Kinetic Monte Carlo Simulations.* Macromolecular Theory and Simulations, 2011. **20**(1): p. 54-64.

163. Liang, Y., Pingali, S.V., Jogalekar, A.S., Snyder, J.P., Thiyagarajan, P., and Lynn, D.G., *Cross-Strand Pairing and Amyloid Assembly.* Biochemistry, 2008. **47**(38): p. 10018-10026.

164. Balbach, J.J., Ishii, Y., Antzutkin, O.N., Leapman, R.D., Rizzo, N.W., Dyda, F., Reed, J., and Tycko, R., *Amyloid Fibril Formation by A Beta(16-22), a Seven-Residue Fragment of the Alzheimer's Beta-Amyloid Peptide, and Structural Characterization by Solid State NMR.* Biochemistry, 2000. **39**(45): p. 13748-13759.

165. Wang, J.Q., Tao, K., Zhou, P., Pambou, E., Li, Z.Y., Xu, H., Rogers, S., King, S., and Lu, J.R., *Tuning Self-Assembled Morphology of the A Beta(16-22) Peptide by Substitution of Phenylalanine Residues.* Colloids and Surfaces B-Biointerfaces, 2016. **147**: p. 116-123.

166. Chiricotto, M., Melchionna, S., Derreumaux, P., and Sterpone, F., *Hydrodynamic Effects on Beta-Amyloid (16-22) Peptide Aggregation.* Journal of Chemical Physics, 2016. **145**(3).

167. Favrin, G., Irback, A., and Mohanty, S., *Oligomerization of Amyloid a Beta(16-22) Peptides Using Hydrogen Bonds and Hydrophobicity Forces.* Biophysical Journal, 2004. **87**(6): p. 3657-3664.

168. Senguen, F.T., Lee, N.R., Gu, X., Ryan, D.M., Doran, T.M., Anderson, E.A., and Nilsson, B.L., *Probing Aromatic, Hydrophobic, and Steric Effects on the Self-Assembly of an Amyloid-B Fragment Peptide.* Mol Biosyst, 2011. **7**(2): p. 486-96.

169. Bera, M.K. and Antonio, M.R., *Crystallization of Keggin Heteropolyanions Via a Two-Step Process in Aqueous Solutions.* Journal of the American Chemical Society, 2016. **138**(23): p. 7282-7288.

170. Burkoth, T.S., Benzinger, T.L.S., Urban, V., Morgan, D.M., Gregory, D.M., Thiyagarajan, P., Botto, R.E., Meredith, S.C., and Lynn, D.G., *Structure of the Beta-Amyloid((10-35)) Fibril.* Journal of the American Chemical Society, 2000. **122**(33): p. 7883-7889.

171. Petkova, A.T., Buntkowsky, G., Dyda, F., Leapman, R.D., Yau, W.M., and Tycko, R., *Solid State NMR Reveals a pH-Dependent Antiparallel Beta-Sheet Registry in Fibrils Formed by a Beta-Amyloid Peptide.* Journal of Molecular Biology, 2004. **335**(1): p. 247-260.

172. Lansbury, P.T., Costa, P.R., Griffiths, J.M., Simon, E.J., Auger, M., Halverson, K.J., Kocisko, D.A., Hendsch, Z.S., Ashburn, T.T., Spencer, R.G.S., Tidor, B., and Griffin, R.G., *Structural Model for the Beta-Amyloid Fibril Based on*

Interstrand Alignment of an Antiparallel-Sheet Comprising a C-Terminal Peptide. Nature Structural Biology, 1995. **2**(11): p. 990-998.

173. Gordon, D.J., Balbach, J.J., Tycko, R., and Meredith, S.C., *Increasing the Amphiphilicity of an Amyloidogenic Peptide Changes the Beta-Sheet Structure in the Fibrils from Antiparallel to Parallel.* Biophysical Journal, 2004. **86**(1): p. 428-434.

174. Selinger, J.V., MacKintosh, F.C., and Schnur, J.M., *Theory of Cylindrical Tubules and Helical Ribbons of Chiral Lipid Membranes.* Physical Review E, 1996. **53**(4): p. 3804-3818.

175. Ghafouri, R. and Bruinsma, R., *Helicoid to Spiral Ribbon Transition.* Physical Review Letters, 2005. **94**(13).

176. Nguyen, P.H., Li, M.S., Stock, G., Straub, J.E., and Thirumalai, D., *Monomer Adds to Preformed Structured Oligomers of Aβ-Peptides by a Two-Stage Dock-Lock Mechanism.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(1): p. 111-6.

177. Rufo, C.M., Moroz, Y.S., Moroz, O.V., Stoehr, J., Smith, T.A., Hu, X., DeGrado, W.F., and Korendovych, I.V., *Short Peptides Self-Assemble to Produce Catalytic Amyloids.* Nature Chemistry, 2014. **6**(4): p. 303-309.

178. Michaelis, L. and Menten, M.L., *Die Kinetik Der Invertinwirkung.* Biochemische Zeitschrift, 1913. **49**: p. 333–369.

179. Chen, W.W., Niepel, M., and Sorger, P.K., *Classic and Contemporary Approaches to Modeling Biochemical Reactions.* Genes & Development, 2010. **24**(17): p. 1861-1875.

180. Jiang, L., Althoff, E.A., Clemente, F.R., Doyle, L., Rothlisberger, D., Zanghellini, A., Gallaher, J.L., Betker, J.L., Tanaka, F., Barbas, C.F., III, Hilvert, D., Houk, K.N., Stoddard, B.L., and Baker, D., *De Novo Computational Design of Retro-Aldol Enzymes.* Science, 2008. **319**(5868): p. 1387-1391.

181. List, B., Barbas, C.F., and Lerner, R.A., *Aldol Sensors for the Rapid Generation of Tunable Fluorescence by Antibody Catalysis.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(26): p. 15351-15355.

182. Schmidt, J., Ehasz, C., Epperson, M., Klas, K., Wyatt, J., Hennig, M., and Forconi, M., *The Effect of the Hydrophobic Environment on the Retro-Aldol Reaction: Comparison to a Computationally-Designed Enzyme.* Organic & Biomolecular Chemistry, 2013. **11**(48): p. 8419-8425.

183. Muller, M.M., Windsor, M.A., Pomerantz, W.C., Gellman, S.H., and Hilvert, D., *A Rationally Designed Aldolase Foldamer.* Angew Chem Int Ed Engl, 2009. **48**(5): p. 922-5.

184. Tanaka, F., Fuller, R., and Barbas, C.F., 3rd, *Development of Small Designer Aldolase Enzymes: Catalytic Activity, Folding, and Substrate Specificity.* Biochemistry, 2005. **44**(20): p. 7583-92.

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

186. Omosun, T.O., Hsieh, M.-C., Childers, W.S., Das, D., Mehta, A.K., Anthony, N.R., Pan, T., Grover, M.A., Berland, K.M., and Lynn, D.G., *Catalytic Diversity in Self-Propagating Peptide Assemblies.* Nat Chem, 2017. **In press**: p. 10.1038/nchem.2738.

187. Omosun, T.O., *Cross-Beta Folds of Amyloid as a Versatile Self-Propagating Catalyst.* Emory University Doctoral Thesis, 2016.

188. Giger, L., Caner, S., Obexer, R., Kast, P., Baker, D., Ban, N., and Hilvert, D., *Evolution of a Designed Retro-Aldolase Leads to Complete Active Site Remodeling.* Nat Chem Biol, 2013. **9**(8): p. 494-8.

189. Pu, L. and Yu, H.B., *Catalytic Asymmetric Organozinc Additions to Carbonyl Compounds.* Chemical Reviews, 2001. **101**(3): p. 757-824.

190. Wang, Z.X., Tu, Y., Frohn, M., Zhang, J.R., and Shi, Y., *An Efficient Catalytic Asymmetric Epoxidation Method.* Journal of the American Chemical Society, 1997. **119**(46): p. 11224-11235.

191. Stroppolo, M.E., Falconi, M., Caccuri, A.M., and Desideri, A., *Superefficient Enzymes.* Cellular and Molecular Life Sciences, 2001. **58**(10): p. 1451-1460.

192. Xue, W.-F., Homans, S.W., and Radford, S.E., *Systematic Analysis of Nucleation-Dependent Polymerization Reveals New Insights into the Mechanism of Amyloid Self-Assembly.* Proceedings of the National Academy of Sciences, 2008. **105**(26): p. 8926-8931.

193. Hurvich, C.M. and Tsai, C.L., *Regression and Time Series Model Selection in Small Samples.* Biometrika, 1989. **76**(2): p. 297-307.

194. Burnham, K. and Anderson, D., *Model Selection and Multimodel Inference* NY: Springer, 2002.

195. Lassila, J.K., Baker, D., and Herschlag, D., *Origins of Catalysis by Computationally Designed Retroaldolase Enzymes.* Proceedings of the National

Academy of Sciences of the United States of America, 2010. **107**(11): p. 4937-4942.

196. Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W.C., *Macromodel - an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics.* Journal of Computational Chemistry, 1990. **11**(4): p. 440-467.

197. Bowers, K.J., Chow, E., Xu, H., Dror, R.O., Eastwood, M.P., Gregersen, B.A., Klepeis, J.L., Kolossvary, I., Moraes, M.A., and Sacerdoti, F.D. *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters*. in *SC 2006 Conference, Proceedings of the ACM/IEEE*. 2006. IEEE.

198. Berendsen, H., Postma, J., Van Gunsteren, W., and Hermans, J., *Interaction Models for Water in Relation to Protein Hydration.* Intermolecular forces, 1981. **11**(1): p. 331-342.

199. Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., and Jorgensen, W.L., *Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins Via Comparison with Accurate Quantum Chemical Calculations on Peptides.* The Journal of Physical Chemistry B, 2001. **105**(28): p. 6474-6487.