SELF-ASSEMBLY OF ELASTIN-LIKE PEPTIDES: STUDIES BY
SINGLE MOLECULE IMAGING

By

Guocheng Yang

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Chemical Engineering and Applied Chemistry

University of Toronto

©Copyright by Guocheng Yang (2008)
Understanding the basic mechanisms and dynamics that drive the assembly of molecules into functional structures is critically important in a diverse number of fields, ranging from materials science to drug delivery and biomaterials. In this work, we have focused on examining the self-assembly characteristics, both in solution and at surfaces, of a family of elastin-like peptides (EPs). In addition to directly observing the formation of ordered hexagonally arranged fibrillar EP structures on hydrophobic highly ordered pyrolytic graphite (HOPG), we have studied the dynamics of EP self-assembly process both within physically restricted domains using thermally etched HOPG, and in solution using detergent micelles. We have found that, at surfaces, EP fibril formation occurs via surface stabilization against the hydrophobic surface, while in solution, detergents inhibit EP aggregation at high temperatures and appear to enable the formation of an ordered crystalline structure at low temperatures. These model studies establish a framework for further investigations of peptide self-assembly and the role of hydrophobic interactions in controlling self-assembly.
I want to take this opportunity to thank my supervisor Professor Christopher Yip, as well as reading committee members, Professors Kim Woodhouse and Fred Keeley, for their continuous great help for me to finish this interesting research project. This is a long journey and more difficult than expected in the beginning. I cannot imagine that I am able to finish this work without the guidance from them, and I am glad to learn a lot of useful skills on how to drive a research project.

I also want to thank many other people who have provided me with very useful advice and help, e.g. Dr. Jeff Lee from Professor Lynn Howell’s lab, Dr. Megan Miao from Professor Fred Keeley’s lab and Dr. Jason Hong from Professor Robert Pilliar’s lab. For sure I also want to thank all the current and previous Yip Lab members, including summer students and forth-year thesis students. They brought me many valuable memories besides the research we have worked together.

Finally, I want to thank my parents and my bother’s family. Although they are far away from here and we cannot talk everyday, their continuous support and encouragement helped me overcome many of the difficulties that were in the way of this process.
# TABLE OF CONTENTS

List of figures.................................................................................................................. iv
List of tables..................................................................................................................... xi
Abbreviations................................................................................................................... xii

1 Introduction ......................................................................................................................... 1
  1.1 Molecular Self-Assembly ............................................................................................. 1
  1.2 Biomolecular Self-Assembly ....................................................................................... 4
  1.3 Ordered Self-Assembly of Biomolecules ................................................................... 6
  1.4 Crystallization of Membrane Proteins .................................................................... 14
  1.5 Amphiphilic Molecules ............................................................................................... 18
  1.6 Elastin and Tropoelastin ............................................................................................. 21
  1.7 Atomic Force Microscopy (AFM) .............................................................................. 24
  1.8 Dynamic Light Scattering (DLS) ............................................................................... 33

2 Multi-Layer and Nanoconfined Self-Assembly of EP on HOPG .................................. 39
  2.1 Ordered Self-Assembled Multilayers of EP Fibrils on Graphite Surface ............... 41
  2.2 Trial of Crystallizing EP on Graphite Surface ......................................................... 45
  2.3 The Growth of EP Fibers within Corrals on Graphite Surface ............................... 47
  2.4 Conclusion ..................................................................................................................... 54

3 Self-Assembly of EP-Detergent Complex at Solid-Liquid Interface ............................. 55
  3.1 Ordered Self-Assembly of Detergent on HOPG ....................................................... 56
  3.2 Complex of EP Peptides with Detergents on HOPG at Lower Temperature ....... 63
  3.3 Substrate Facilitated Self-Assembly of EP and Its Complex with Detergent .. 67
  3.4 Particle Size Change Measured by DLS ................................................................. 76
  3.5 Conclusion ..................................................................................................................... 83

4 Molecular Structure Investigation of EP .................................................................... 84
  4.1 FT-IR Spectrum of EP Secondary Structure ........................................................... 84
  4.2 Unfolding of EP by Single Molecule Force Spectroscopy ................................... 91
  4.3 Molecular Dynamics Simulation EP Using CAChe .............................................. 100

5 Conclusion and Discussion .......................................................................................... 105
  5.1 Discussion ..................................................................................................................... 105
  5.2 Future Recommendation ............................................................................................ 111
  5.3 Conclusion ..................................................................................................................... 115

6 Appendices ................................................................................................................. 116
  6.1 Atomic Force Microscopy Imaging .......................................................................... 116
  6.2 FT-IR Reflectance and Absorbance Spectroscopy ............................................... 117
  6.3 Dynamic Light Scattering ......................................................................................... 117
  6.4 Molecular Dynamics Simulation ............................................................................ 117
  6.5 Customized Automatic Temperature Control Circuit ........................................... 118
Figure 1-1. Schematic of proteins self-assembled into symmetric nanostructures. (A) The green semicircle represents a dimeric protein with 2-fold axis of symmetry; the red triangle represents a trimeric protein with 3-fold axis of symmetry. (B) The different conformation from the dimeric and trimeric proteins as they are connected by a rigid linker in different geometries. (C) The model of the fused nanohedral cage formed from the proteins with intersecting symmetric axis. (D) An SEM image of the real particles obtained from the design. (Adapted from reference 11)

Figure 1-2. Surface free energy change of homogenous nucleation. There is an energy barrier at critical nucleus radius $r^*$, after which the volume free energy ($G_B$) overwhelms the surface free energy ($G_S$).

Figure 1-3. Schematic illustration of a cluster of molecules forming a nucleus at interface. $\theta$ is the contact angle and $r$ is the radius of projected sphere.

Figure 1-4. Molecular structure of T. thermophilus IleRS (IILE). (Adapted from reference 31)

Figure 1-5. Schematic illustration of the sticky end, which is a short single strand overhang from the double-stranded helical DNA molecule. By using complimentary nucleotide bases, two identical DNA molecules are easily joined together by this type of structure.

Figure 1-6. Schematic illustration of the photoreaction centre from Rhodopseudomonas viridis (1PRC) embedded in a membrane bilayer. Subunits of the protein are shown in different colours: green, cytochrome; blue, M-subunit; brown, L-subunit; purple, H-subunit.

Figure 1-7. Amphiphiles with different shapes. The left-hand amphiphiles (A) tend to form normal phases ($\nu/a_{l_c} < 1$); the middle amphiphiles (B) tend to form lamellar phases ($\nu/a_{l_c} = 1$); the right-hand amphiphiles (C) tend to form reversed phases ($\nu/a_{l_c} > 1$).

Figure 1-8. In situ tapping mode AFM images of lipid bilayer disrupted by wild-type $\alpha$-synuclein. (A) Confluent bilayer of POPC/POPS (1:1) formed on mica before the addition of $\alpha$-synuclein. (B) After adding the wild-type $\alpha$-synuclein at 0.1 mg/ml in PBS, the bilayer gradually undergoes deformation, the aggregates of $\alpha$-synuclein can be found in the zoomed image. (C) After 9 hours, protein aggregates and putative fibrils can be found, as indicated by the arrows. Scale bar is 2 $\mu$m. (Adapted from reference 53.)

Figure 1-9. Illustration of interaction force between two atoms as the function of distance, derived from the Lennard-Jones 6-12 potential model.

Figure 1-10. Schematic of AFM components. The laser beam is projected onto the back of the tip and reflected onto the photodiode detector. The feedback loop controls the
movement of the piezoelectric scanner, on which the sample is positioned. Right-
hand side is the image of Multimode SPM (Veeco, Santa Barbara, CA) equipped
with an "E" scanner, with a lateral scan area of 14.6 × 14.6 µm.------------------------ 27

Figure 1-11. Illustration of the difference in the operation of contact mode AFM (left) and
tapping mode AFM (right) on sample surface. The damage of the sample caused
by the scanning tip is shown by the arrow in contact mode AFM.----------------------- 28

Figure 1-12. Continuous change of endothelial cell morphology captured by in situ AFM. A
series of AFM images clearly show the cell body is retracting after treated by AβP1-
42 (a-c). This effect can be prevented by anti-AβP antibody even after the same
exposure time, there is little change on cell morphology as shown in images d-f.
(Adapted from reference 73)..................................................................................... 29

Figure 1-13. Schematic of AFM force spectroscopy experiment depicting the deflection of
cantilever and the analogy to the extension of a spring under an applied force......... 31

Figure 1-14. Overestimated width profile of sample (solid line) caused by tip shape is
compared with actual size of sample. Wm is measured width at half-height of the
sample and Hmax is the maximum height in the measured profile, R is the radius of
tip and r is the radius of sample particle. (G. Yang, M.A.Sc. thesis, “An In Situ
Atomic Force Microscopy (AFM) Investigation of Tropoelastin-like Peptide
Assembly at Ordered Interface”, 2002)..................................................................... 32

Figure 1-15. Schematic of difference in the incident light and scattered light in the light
scattering measurement. It is shown that the constructive interference and high
scattered intensity found at lower angles. Strong destructive interference and low
scattered intensity happens at higher angle, 90°..................................................... 36

Figure 1-16. A typical dynamic light scattering measurement of 100 nm polystyrene latex
beads. The effective diameter is 93.6 nm and the polydispersity is 0.139, which
indicates the particle size is monodisperse in solution.............................................. 37

Figure 2-1. In situ TMAFM images of the rapid nucleation and growth of EP I fibrils on
HOPG at 50 °C. Note that the elapsed time between images (A) and (B) is ~2 h.
(C) and (D) show the amorphous aggregates of EP I and EP II on mica at room
Copyright © 2002 American Chemical Society.) ..................................................... 40

Figure 2-2. Schematic illustration of EP multi-layers self-assembled on HOPG surface. A
is the first amorphous thin film layer, B is the first ordered layer of fibers in
hexagonal pattern, C is the second ordered layer and D is the third ordered layer..... 41

Figure 2-3. Hexagonal multilayer assembly of EP I on HOPG at 50°C. Image (B) was
captured four hours later than (A), in which two distinctive layers, layer B and layer
C, can be clearly identified. (C) The close-up view from image (A), the arrows
indicate the fiber orientation. (D) 2D Fourier power spectrum of (A), solid arrow
refers to layer B and dotted arrow refers to layer C. (E) and (F) are the section
analysis of the height of fibers seen in image (A) and (B)......................................... 43

Figure 2-4. Images at the same position shows three separated domains grow
independently. Figure 2-4(B) was acquired half an hour after Figure 2-4(A).............. 45
Figure 2-5. Optical microscope image of particles formed on HOPG at 50 °C from 25 µM EP I solution at pH 7.5. ................................................................. 46
Figure 2-6. Optical microscope image of particles formed on HOPG at 40 °C from 25 µM EP II solution at pH 7.5. ................................................................. 46
Figure 2-7. A larger particle formed on HOPG at 40 °C from 125 µM EP II solution at pH 7.5. Inserted is the X-ray diffraction pattern of this particle. The diffraction ring at 3.86 Å is from the ice within the particle. ............................................ 47
Figure 2-8. Corrals formed on HOPG after thermal-etching treatment. (A) The HOPG surface heated at 800 °C for 6 minutes, (B) The HOPG surface heated at 800 °C for 7 minutes. ................................................................. 48
Figure 2-9. Two images of EP I peptide grown within and outside the HOPG corrals at 50°C. Image (B) is captured 28 minutes after image (A). (Height scale: 20 nm) .......... 49
Figure 2-10. A small pit remains visible after surrounding area has been covered by hexagonal fibers. The height and width of the pit can be seen from the section diagram. ........................................................................................................ 50
Figure 2-11. Ordered EP fibers formed within deep pits. The fibers can be clearly identified from the amplitude image (left), while the height image (right) only reveals that it is a very deep pit. The section diagram shows this deep pit has a flat bottom. ........................................................................................................ 51
Figure 2-12. Within thermally-annealed pits having multiple (0001) layers of narrow step edges, EP peptide molecules were unable to form extended fibers as found on the open area. Image (B) was taken 3 hours later than image (A) at the same position. No fibers formed within the lower area, except the individual globular aggregates. (C) The section diagram of the pit showing the curvature. ........................................... 52
Figure 2-13. Illustration of the space requirement for EP adsorbing onto the graphite surface......................................................................................... 54
Figure 3-1. Two types of membrane crystals. (a) Type I: stacks of membrane protein in 2D planes. (b) Type II: membrane proteins crystallized within detergent micelles. ............. 56
Figure 3-2. Self-assembly of EP II with Hexyl-β-D-glucopyranoside (HG) on HOPG surface as a function of temperature. (A) The image captured right after the mixture was deposited on the HOPG at room temperature. Self-assembled fibers in a hexagonal pattern can be clearly identified besides the big particles. Upon heating to 35°C, the fibers start to dissolve, as shown in image (B) and (C). Upon cooling, the fibers reformed on the HOPG surface, image (D). ........................................... 57
Figure 3-3. Illustration of the self assembly of EP and its complex with detergent on HOPG at various temperatures. (A) The adsorption of pure EP molecule on HOPG surface via hydrophobic interaction. White part is the hydrophobic domain; red part is the hydrophilic domain. (B) The assembly of EP molecules with detergents on HOPG surface through the hydrophobic tail of the detergent molecules. The dot line indicates the detergents embedded between the EP and HOPG ........................................................................................................ 58
Figure 3-4. Deflection image (A) and height image (B) of EP II and HG on graphite surface at 10 °C. Besides the large particle with angular edge, the ordered ~ 170 nm long fibers can be clearly seen on the surface of graphite from the deflection image. Image size: 2 × 2 µm.

Figure 3-5. AFM images of ordered structure of HG, OG, DM and Anzergent adsorbed on HOPG at 4 °C. All these detergents form linear semi-cylindrical structures parallel to each other. Inserted: Periodicity revealed by Fast Fourier transform analysis, showing the spacing of the ordered structures. Image size 250 × 250 nm.

Figure 3-6. Self assembly of 1 CMC Zwitterionic® 3-10 on HOPG at 4 °C. In image (A), three separate domains set side by side on graphite surface with different orientation and clear boundary, as indicated by the red line. (B) The close-up image of the distinctive end of each single array of detergents within different domains. The intersection angle between domains is ~ 120°. The alignment of detergent molecules is indicated as perpendicular to the long axis of the rods.

Figure 3-7. Images (1 × 1 µm) of the structures formed from the mixture of 25 µM EP I with 1 CMC detergent on HOPG surface at 4 °C. While EP I - HG complex and EP I - DM complex only formed an amorphous layer on the HOPG surface, the EP I - OG complex and EP I - Zwitterionic complex formed linear arrays which are similar to pure detergent, as shown in the inserted image (250 × 250 nm).

Figure 3-8. Images (1 × 1 µm) of the structures formed from the mixture of 25 µM EP II with 1 CMC detergent on HOPG surface at 4 °C. Besides the EP II - HG fibers seen previously, only EP II - Anzergent® 3-10 complex formed fiber-like structure. EP II - OG complex formed linear arrays similar to pure detergent (inserted image: 250 × 250 nm) and EP II - DM complex formed an amorphous layer.

Figure 3-9. The morphology of 25 µM EP I on a pre-formed OG monolayer at 4 °C after a 2-hour incubation. The linear alignment of OG can be clearly identified from image (A), as indicated by the red arrow. (B) EP I form globular structure on the pre-layer instead of those fibers in the hexagonal pattern. (Image size: 1 × 1 µm)

Figure 3-10. 25 µM EP I on OG pre-layer at elevated temperature. Image (A) is captured right after the EP has been put on top of OG at 4 °C. There are only small particles adsorbed on the OG layer. (B) Image captured after the EP-OG complex has been heated to 46 °C. No specific structural feature can be found besides the small particles.

Figure 3-11. An amorphous particle formed from the mixture of 25 µM EP I and 1 CMC OG on a pre-layer of OG on top of the HOPG.

Figure 3-12. Illustration of the effect of surface property on the adsorption of EP peptides. EPs are able to form fibers in an ordered pattern on HOPG due to hydrophobic interaction. When the surface is covered by the detergents, similar structures cannot form due to the loss of proper interaction.
Figure 3-13. Structure change of EP I fibers formed on HOPG before and after the injection of OG. (A) 25 µM EP I self-assembled into normal fibers in a hexagonal pattern on HOPG at 45 °C, alone with some big aggregates. After the injection of 1 CMC OG, image (B) shows the big aggregates and part of the fibers were washed away. (C) When the solution was cooled down to room temperature, the remaining fibers stayed on the surface but no specific structure was formed. .......................... 71

Figure 3-14. 5 µM EP I formed fibers on HOPG surface after injected into 1 CMC OG at 40 °C (A). The section line shows the difference in height of two types of fibers, the thick fibers (as indicted by the arrow) are ~ 1 nm taller than the thinner fibers. (B) After ~ 1 hour, when the solution was cooled down to room temperature, more and more thick fibers appeared on the surface. .......................................................... 73

Figure 3-15. At 4 °C, a two-dimensional ordered structure was found from the mixture of 25 µM EP I and 1 CMC OG (A) and it continued to expand during a 1-hour period (B). Image (C) is the close-up view, showing the typical crystalline structure. The section diagram of the designated area in image (C) shows the step height is ~ 4.6 nm. .................................................................................................................. 75

Figure 3-16. Crystal size of EP I with OG increase during a 1-hour period. ......................... 76

Figure 3-17. Aggregate size of EP I and EP II with 1 CMC OG as a function of temperature. The orange curve indicates the increased size of EP I and OG complex upon heating from 25 °C to 50 °C. The blue curve represents the size change of the same composition upon cooling to 5 °C. Another group of size change curves were shown for the combination of EP II with OG, the red line is the size heating curve and the deep blue line is the cooling curve. .......................................................... 78

Figure 3-18. Particle size measurement of 25 µM EP I and 1 CMC OG at 50 °C obtained by dynamic light scattering. ........................................................................................................ 79

Figure 3-19. Aggregate size of EP I and EP II with 1 CMC DM as a function of temperature. The red line and orange line are the heating curves of EP II and EP I with DM, 25 °C to 50 °C. The deep blue line and the blue lines are the cooling curves, from 50 °C to 5 °C. .................................................................................................................. 80

Figure 3-20. Aggregate size of EP I and EP II with 1 CMC Anzergent®3-10 as a function of temperature. The orange and red lines represent the heating curves of EP I and EP II with Anzergent®3-10, from 25 °C to 50 °C. The blue and deep blue lines represent the cooling process, from 50 °C to 5 °C. .................................................................................................................. 81

Figure 3-21. Aggregate size of EP I and EP II with 1 CMC HG as a function of temperature. The orange line is from the mixture of EP II with HG, from 10 °C to 50 °C. The red and blue lines are the heating and cooling curves of EP I with HG. ... 82

Figure 4-1. General diagram of interferometer used for FT-IR. .............................................. 85

Figure 4-2. Infrared transmission-absorbance spectra of EP in KBr pellet in air collected at room temperature. The absorption peaks at ~ 1630 cm⁻¹ and 1550 cm⁻¹ belong to amide I (*) and amide II band (**). .................................................................................................................. 88

Figure 4-3. Image shows the location on the EP I thin film where the IR spectra were collected. .................................................................................................................. 89
Figure 4-4. Infrared spectra of EP thin films, from 2000 to 1000 cm\(^{-1}\), adsorbed on HOPG, collected on Nicolet™ Continuum™ infrared microscope, using reflectance mode. Amide I (*) and amide II (**) bands are indicated on graph. .......... 90

Figure 4-5. Well-defined recombinant titin segments produces force extension curves from the unfolding of individual Ig domains. (A) is from the stretching of Ig8, (C) is from the unfolding of Ig4, the superimposed force curves in (B) and (D) reveals the 25 nm pattern, (F) demonstrates the unfolding of individual Ig domains. (Adapted from reference 109) .............................................................................................................................. 93

Figure 4-6. Schematic diagram for different patterns in breaking hydrogen bond between β-strands. (A) The “zipper” model of H-bond breakage, in which each bond breaks sequentially. (B) The “shear pattern” model of H-bond breakage, in which the orientation of H-bond is perpendicular to the applied force. ......................... 94

Figure 4-7. Single molecule force spectroscopy on a single fiber formed from EP I. (A) the entire fiber before the force spectroscopy. (B) The dissected fiber after the force spectroscopy. Insert: two force curves recorded between these images, in which the sawtooth pattern can be clearly identified. ............................................................. 95

Figure 4-8. Histogram distribution of peaks found in the single molecule force spectroscopy of EP I pulled at 250 nm/s, arrows indicate the characteristic unfolding length as identified from the local maximum. ......................................................... 96

Figure 4-9. Fast Fourier transform of the frequency of the unfolding length of EP I reveals the well-defined spacing of 2.59 nm, which is believed from the repeat -PGVGVA- segment. .................................................................................................................. 97

Figure 4-10. Histogram distribution of peaks found in the single molecule force spectroscopy of EP II pulled at 250 nm/s, arrows indicate the characteristic unfolding length as identified from local maximum. ......................................................................................... 98

Figure 4-11. Fast Fourier transform of the frequency of the unfolding length of the EP II reveals the well-defined spacing of 2.46 nm, which is in match with the extended length of -PGVGVA- segment. ......................................................................................... 99

Figure 4-12. Molecular conformation of EP I simulated by CAChe at 25 °C and 50 °C. (A) is the molecular surface of EP I at 25 °C and 50 °C drawn by SwissPDB viewer. (B) and (C) are EP I and its surrounding water molecules with 2 Å range at 25 °C and 50 °C. .................................................................................................................. 101

Figure 4-13. Diagram of radius of gyration change of EP I in solution at 25 °C (A) and at 50 °C (B). Below are the solvent accessible surface area of EP I at 25 °C (C) and at 50 °C (D), the black line is for the hydrophobic surface. ................................................................. 103

Figure 5-1. Molecular structure of a series of peptides capable of self-assembling in aqueous solution, RAD16-I, RAD16-II, EAK16-I and EAK16-II. (A) They contain both hydrophobic and hydrophilic surfaces, which function as pegs and holes to facilitate the further extension of these “peptide Lego”s. (B) AFM image of nanofibers formed from RADA16-I, which has been used as scaffold for cell culture and regenerative medicine. (Adapted from reference 120) ............................................. 106
Figure 5-2. Illustration of the zigzag and armchair edge of graphite edge, highlighted by grey lines connecting circles which represent the carbon atoms. This structure may explain the incompatibility with large biomolecules...

Figure 5-3. Illustration of the schematic showing the temperature-dependent ordered self-assembly of EP and its complex with detergent in solution and on HOPG surface.

Figure 6-1. Circuit diagram of the automatic switch introduced between the Igloo Koolmate and the Omega temperature controller. By using this unit, the Koolmate is able to work as either a cooler or heater.
LIST OF TABLES

Table 1-1. Sequence details of cDNA templates, exon 20, 21, 23, 24; and EP peptides, EP I, EP II and EP IV. The repeat hydrophobic segment is underlined in exon 24. ........ 24
Table 3-1. Chemical structure and properties of detergents used for screening the crystallization of EP. .................................................................................................................. 60
Table 4-1. Typical peak frequency in FT-IR spectrum of amide I and II band from the secondary structure of proteins. .................................................................................................................. 87
Table 4-2. Wavenumbers of amide I and amide II Bands in FT-IR Spectrum of EP I, II and IV in KBr pellet and on HOPG surface. ...................................................................................................... 91
ABBREVIATIONS

8CB  4'-Octyl-4-biphenylcarbonitrile
α-Syn  α-Synuclein
AβP  Amyloid β protein
AD  Alzheimer's disease
AFM  Atomic force microscopy
CD  Circular dichroism
CMC  Critical micelle concentration
DLS  Dynamic light scattering
DDM  n-Dodecyl-β-D-maltoside
DM  n-Decyl-β-D-maltopyranoside
DNA  Deoxyribonucleic acid
DTT  Dithiothreitol
EBP  Elastin-binding protein
EP  Elastin-like peptide
(ET)$_3$I$_3$  Bis(ethylenedithiolo)-tetrathiafulvalene triiodide
FFT  Fast Fourier transform
FT-IR  Fourier transformed infrared spectroscopy
HOPG  Highly ordered pyrolytic graphite
HOMO  Highest occupied molecular orbital
HG  n-Hexyl-β-D-glucopyranoside
Ig  Immunoglobulin
LUMO  Lowest unoccupied molecular orbital
MAS  Magic angle spinning
MD  Molecular dynamics
MHC  Major histocompatibility complex
NMR  Nuclear magnetic resonance
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPC</td>
<td>Neural progenitor cell</td>
</tr>
<tr>
<td>OG</td>
<td>n-Octyl-β-D-glucopyranoside</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PDC</td>
<td>Protein detergent complex</td>
</tr>
<tr>
<td>POPC</td>
<td>1-Palmitoyl-2-oleoyl phosphatidylcholine</td>
</tr>
<tr>
<td>POPS</td>
<td>1-Palmitoyl-2-oleoyl phosphatidylyserine</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SEM</td>
<td>Scanning electron microscope</td>
</tr>
<tr>
<td>STM</td>
<td>Scanning tunnelling microscopy</td>
</tr>
<tr>
<td>SPM</td>
<td>Scanning probe microscopy</td>
</tr>
<tr>
<td>SLS</td>
<td>Static light scattering</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer ribonucleic acid</td>
</tr>
<tr>
<td>TMAFM</td>
<td>Tapping mode atomic force microscopy</td>
</tr>
<tr>
<td>ssNMR</td>
<td>Solid-state nuclear magnetic resonance</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
Chapter 1

1 Introduction

1.1 Molecular Self-Assembly

The self-assembly of molecules is referred to as the spontaneous aggregation and ordered formation of constituents when they are mixed in correct proportions under appropriate conditions. For example, crystallization, which can be considered a form of molecular self-assembly into the crystalline solid-state, has been used extensively as a means of purification or separation. Thomas Osborne asserted that protein crystallization is “not only interesting in itself, but is important as presumably furnishing a means for making preparations of an undoubted purity, which will afford a sure basis for further studies of their properties.” However, the actual mechanisms of molecular self-assembly and the principles controlling the process of crystallization remain elusive.

As an example of a native self-assembly process, consider insulin, a small protein hormone that regulates carbohydrate metabolism. Consisting of a 21 amino acid A-chain connected via two disulphide bonds to a 30 amino acid B-chain, insulin is synthesized within the β-cell of the pancreas via the precursor compounds: preproinsulin and proinsulin. After cleavage of the signal sequence from preproinsulin, the proinsulin is transported into Zn\(^{2+}\)- and Ca\(^{2+}\)-rich vesicles in the Golgi and assembled into hexameric \((\text{Zn}^{2+})_2(\text{Ca}^{2+})(\text{Proin})_6\) forms. After removal of the C-peptide, \((\text{Zn}^{2+})_2(\text{Ca}^{2+})(\text{Proin})_6\) is converted into \((\text{Zn}^{2+})_2(\text{Ca}^{2+})(\text{In})_6\), which is much less soluble than the precursor. As a result, insulin is stored as a hexamer in well-defined two-dimensional arrays, resembling the actual insulin protein crystal, in the β-cell. In vitro studies have shown that the insulin hexamer is much more resistant to chemical and/or physical degradation than the insulin monomer. Fusion of the vesicle with the plasma membrane results in expulsion of the insulin crystals into the intercellular fluid. The change in pH and the loss of ligands due to dilution in the intercellular environment causes the insulin hexamer to dissociate into its constituent monomers. This dissociation process is critical to the

action of insulin as it is the bioactive monomer that associates with the transmembrane insulin receptor. Insulin is also known to form fibrillar structures, an irreversible process that is thought to be related to type II diabetes. Identifying the basic mechanisms that control insulin self-assembly from the monomer into either the stable hexamer, or the fibril form, and also how insulin associates with its receptor, is thus critically important for understanding the biological function of this hormone.

Other reasons for the recent interest in molecular self-assembly arise from the emerging needs of nanotechnology. While there is much debate over the definition of nanotechnology, it is generally accepted that this term refers to the exploitation of phenomena that are present only on molecular length scales and where the properties of a nanostructured material are uniquely different from those that are present in conventional bulk forms. Nanotechnology is recognized as an interdisciplinary technology emerging from physics, chemistry, biology and material science, to develop structures at a scale between individual molecules and 100 nm, in which the building components are single molecules. The external form of the materials reflects the shape of molecules (building blocks) and the geometric arrangement of these building blocks. Using this approach, researchers are able to build materials with very specific physical and chemical properties. For example, nanostructured gold exhibits unique fluorescence properties that are directly related to the number of gold atoms within the structure: \(\text{Au}_5\) emits at 385 nm (UV), \(\text{Au}_8\) emits at 455 nm (blue), \(\text{Au}_{13}\) emits at 510 nm (green) and \(\text{Au}_{23}\) emits at 760 nm (red). This size-dependent emission is generated from a discrete, quantum-confined electron transition that only occurs within a critical size scale.

If the self-assembled material structure is assumed to be a network, the building blocks are considered nodes and the intermolecular interaction or coordination bonds represent node connections. The desired one, two or three-dimensional structure can be achieved by choosing suitable nodes and connection sites. For example, oligomeric proteins can be used as nodes provided they exhibit an appropriate structural symmetry. Two-dimensional protein layers can
be used as biosensors or molecular sieves while tetrahedral protein cages can be used for delivering drugs or genes, or stabilizing and sequestering other molecules. For example, the dimeric M1 matrix protein of influenza virus, is designed to connect with a trimeric protein, bromoperoxidase, through a “rigid” connection. This complex is capable of forming highly symmetrical nanohedral structures including 12 copies of engineered fusion proteins. (Figure 1-1)

Figure 1-1. Schematic of proteins self-assembled into symmetric nanostructures. (A) The green semicircle represents a dimeric protein with 2-fold axis of symmetry; the red triangle represents a trimeric protein with 3-fold axis of symmetry. (B) The different conformation from the dimeric and trimeric proteins as they are connected by a rigid linker in different geometries. (C) The model of the fused nanohedral cage formed from the proteins with intersecting symmetric axis. (D) An SEM image of the real particles obtained from the design. (Adapted from reference 11)
Rational exploitation of molecular self-assembly requires the use of all the resources of molecular chemistry combined with specifically engineered noncovalent intermolecular interactions. For example, there is keen interest in the design of porous materials for hydrogen storage. The Yaghi group demonstrated that it is possible to build a highly porous crystalline structure in which the maximum open space reaches 91.1% of the total volume. This metal-organic framework is built from octahedral Zn-O-C clusters with various molecular struts made of biphenyl, tetrahydropyrene, pyrene and terphenyl. It should be noted that this material is thermally stable despite its hollow structure. It will not decompose until heated to 400 °C. Another very interesting application of these materials is that, due to their precisely controlled internal structure, they can be used as membrane filters to separate molecules that differ by only 0.1 nm in diameter, such as \( o \)-xylene and \( p \)-xylene. In contrast to commercially available zeolite membranes, these materials rely on pure structural recognition rather than physical adsorption to effect separation.

1.2 Biomolecular Self-Assembly

Biomolecules are macromolecules that naturally exist in living organisms, working together to exert their proper function. They can be classified as proteins, polysaccharides, nucleic acids or lipids. Proteins are linear polymers made of 20 basic common amino acids, which are described as polar, nonpolar and charged, according to their functional side group. Polysaccharides are made of simple sugars (monosaccharides, carbohydrates) with glycosidic bonds. They are the major source of energy, a component of cell walls and interact with proteins in performing their biological function. Nucleic acids are made of nucleosides that carrying genetic information and direct protein synthesis. Naturally occurring membranes are made up of amphiphilic lipid molecules and contain embedded membrane proteins and cholesterol. They separate cells or compartments within cells from their environments by acting as barriers.

In order to ensure their proper function or to protect themselves, biomolecules are able to spontaneously assemble into ordered structures. For example, under extreme
environmental stress, DNA could protect itself through a non-enzymatic physical structural transition from an open conformation (which is vulnerable to damage) into ordered structures.\textsuperscript{17} It can also form a highly ordered, tightly packed DNA-Dps (DNA-binding protein) co-crystal designed to protect DNA against oxidizing agents and nucleases.\textsuperscript{18}

Protein self-assembly is determined by subtle factors that control the molecular structure at various levels. The sequence of amino acids within a protein is referred to as its primary structure. The local structure within the protein is referred to as its secondary structure, e.g. $\alpha$-helix (right-handed spiral), $\beta$-sheet (parallel aligned short segments) and random coils. These secondary structures are the building blocks for the tertiary structure of a protein, which helps to define its biological function. Some proteins have a further level of organization termed quaternary structure, in which multiple subunits of protein molecules assemble into a larger structure. An excellent example of this structural hierarchy can be seen in the previously described insulin example. The insulin monomer comprises an $\alpha$-helical region and a $\beta$-sheet rich domain. It self-assembles via hydrogen-bonding to form a $\beta$-sheet dimer. These dimers then self-associate around Zn\textsuperscript{2+} ions to form the stable hexamer storage form, which is referred as the quaternary structure of insulin.

The folding and stability of a protein's structure are maintained by several noncovalent forces, including electrostatic, van der Waals, hydrogen bonding and hydrophobic interactions. The energies of these non-covalent bonds are relatively weak, e.g. $\sim 2\text{-}7.5$ kJ/mol for hydrogen bonds, $\sim 1$ kJ/mol for Van der Waals forces and $\sim 3$ kJ/mol for hydrophobic interaction between $-\text{CH}_2$, so that the protein structure is flexible within a small structural range.\textsuperscript{19} This flexibility allows the protein to perform its proper function; however, this flexibility is also a disadvantage as it can facilitate conformational changes into mis-folded or non-native structures.

Proteins can aggregate and lose their biological function or activity due to changes in local environment. The misfolding of proteins has been linked to certain diseases, e.g.
Alzheimer’s disease, Creutzfeld-Jakob disease and type-II diabetes. People have put great efforts into studying the misfolding of proteins, not only to cure diseases, but also to discover the fundamental mechanisms that control protein structure and function.

1.3 Ordered Self-Assembly of Biomolecules

A very simple example of biomolecular self-assembly is the formation and maintenance of cell membranes and vesicles from their constituent lipids. Under biological conditions, the bilayer is fluid and the component lipids are able to diffuse within the plane of the leaflet. The fluidity of the bilayer is useful because it allows for spontaneous repair of damage, and also allows for diffusion and interaction of embedded molecules. The shape of the individual lipid molecules determines the form of assembly. Wedge-shaped lipids tend to form spherical micelles. Cylindrical molecules tend to form extended bilayers. Instead of the original vision of a dilute solution of proteins in a liquid solvent, the biological membrane is a liquid-like structure viewed as a lipid-protein composite with protein concentrations in the membrane of up to 30,000 per μm². Lipid bilayers and embedded membrane proteins are dynamic structures. The component lipids associated with proteins can exchange with the surrounding lipid bilayers. However, little is known about the dynamics of this structure formation, e.g. how one single cell divides into two cells with similar size and identical composition, and how fusion happens between two contacting membranes and the magnitude of the associated free energy change. Although Huang et al. have provided evidence that the shape and density of the fusion intermediates are stalk-like structures, the free energy barrier for fusion is still unknown, and membrane protein-induced fusion is even more complicated.

The Stupp group has successfully used the self-assembly of amphiphilic molecules to make materials for different biological applications. The general principle behind these approaches is very similar. One single molecule contains a long alkyl chain as the hydrophobic tail and several amino acids as the hydrophilic head. These molecules will self-assemble into cylindrical micelles in solution. Besides the basic amphiphilic composition of the molecules, there are four consecutive cysteines placed between the head group and tail, which are used to
trigger the self-assembly process. After treated with dithiothreitol (DTT) to reduce cysteine residues to free thiols, the peptides have a negative charge at pH 8, so that they are soluble due to the electrostatic repulsion between molecules. When the pH is adjusted below 4 to neutralize the negative charge group, the molecules become insoluble as a result of the hydrophobic interaction between alkyl tails and form cylindrical micelles of similar diameter. This is a reversible process which allows for self-correction of improperly aligned molecules. If the thiol group were oxidized by iodine, the fibers are permanently bound even when the pH is brought back to 8. Phosphorylated serine residues were introduced in order to enhance the interaction of these fibres with calcium ions and help direct mineralization of hydroxyapatite, while the RGD sequence was also introduced to improve cell adhesion. It was clearly shown that hydroxyapatite form plate-shaped polycrystalline aggregates along the fiber axis. These self-assembled fibrils have also been used to enhance the directional differentiation of neurons.

With the neurite-promoting laminin epitope IKVAV segment incorporated into the fibrils, 35% of incubated neural progenitor cells (NPC) differentiate rapidly into neurons instead of astrocytes when grown on these fibers. This is the highest level achieved compared with other approaches, and has been attributed to the high surface density of the growth factor on the fiber surface and the highly hydrated structure, which provides a network for diffusion of nutrients, bioactive factors and oxygen.

The self-assembly of larger biomolecules is far more complicated than lipids. The crystallization of proteins is a well-known example. Although some proteins e.g. haemoglobin, can be crystallized without much effort, it is very difficult to crystallize many other proteins and solve their three-dimensional structure via X-ray diffraction. The driving force for protein crystallization is from the difference of the chemical potential ($\Delta\mu$) of protein molecules in the solution and the crystal.

$$\Delta\mu = -kT \ln(c/s)$$
where \( c \) is the protein concentration in solution (mol/L), \( s \) is the solubility of protein (mol/L), \( k \) is the Boltzmann’s constant \( (1.38 \times 10^{-23} \text{ J/K}) \) and \( T \) is the absolute temperature (K). The solubility of protein is defined as the concentration of protein for which the solution is in dynamic equilibrium with the crystal. The ratio \( c/s \) is the supersaturation of solution.

Nucleation is the initiating process for crystal or stable aggregate formation. Such a process formed purely by the molecules of interest is called homogeneous nucleation. In homogeneous nucleation, the Gibbs free energy change for the solution arises from the bond formation between molecules within the ordered structure and the unsatisfied bonds present on the surface of the growing crystal. For the association of \( j \) monomers, the free energy change \( \Delta G_j^0 \) can be expressed as:

\[
\Delta G_j^0 = \nu j G_B + \beta j^\gamma G_S
\]

where \( G_B \) (favourable) is the bulk free energy per unit volume and \( G_S \) (unfavourable) is the free energy per unit area of the surface, \( \nu \) is the volume of growth unit. The coefficients \( \beta \) and \( \gamma \) depend on the shape of the nuclei.

At the beginning of nucleation, the increase in the positive/unfavourable surface free energy overwhelms the negative/favourable volume free energy change so that there is a free energy barrier for molecules to form a critical nucleus. When the nucleus grows to a certain size, the favourable volume free energy \( G_B \) dominates the free energy change and stable crystal growth begins. (Figure 1-2)
Figure 1-2. Surface free energy change of homogenous nucleation. There is an energy barrier at critical nucleus radius $r^*$, after which the volume free energy ($G_B$) overwhelms the surface free energy ($G_S$).

Compared with homogeneous nucleation, heterogeneous nucleation is more likely to occur in the presence of interfaces, such as the surface of the container or impurities, which are believed to reduce the surface free energy barrier for molecules during the self-assembly process.\textsuperscript{28,29}
When a cluster of molecules form on a surface as shown in Figure 1-3, the contact angle $\theta$ is determined by Young's relation,

$$\cos \theta = \frac{\sigma_{sv} - \sigma_{sl}}{\sigma}$$ \hspace{1cm} \text{(1-3)}$$

where $\sigma$ is the liquid-vapour interface tension, $\sigma_{sv}$ is the solid-vapour interface tension and $\sigma_{sl}$ is the solid-liquid interface tension. If the contact angle $\theta$ is 0°, the surface is wetting; if the contact angle $\theta$ is 90°, the surface is non-wetting. If the contact angle of water on surface is less than 30°, the surface is hydrophilic. The volume of a cluster of molecules on surface is given by

$$V = \frac{4}{3} \pi r^3 g(\cos \theta)$$ \hspace{1cm} \text{(1-4)}$$

where

Figure 1-3. Schematic illustration of a cluster of molecules forming a nucleus at interface. $\theta$ is the contact angle and $r$ is the radius of projected sphere.
\[ g(m) = \frac{1}{4}(2 + m)(1 - m)^2 \]  

where \( r \) is the radius of the cluster. The liquid-vapour interface area and the solid-liquid interface area are given by

\[ S_{lv} = 4\pi r^2 \left( \frac{1 - \cos \theta}{2} \right) \]  \hspace{1cm} 1-6

and

\[ S_{sl} = \pi (r \sin \theta)^2 \]  \hspace{1cm} 1-7

The free energy change of a cluster of molecules formed on surface is given by

\[ G = V\Delta G_v + S_{lv}\sigma + S_{sl}\sigma_{sl} \]  \hspace{1cm} 1-8

where \( V \) is the volume of the cluster, \( \Delta G_v \) is the free energy change before and after the molecules associated together, \( S_{lv} \) is the area of the liquid-vapour interface and \( S_{sl} \) is the area of solid-liquid interface. For a flat surface, this free energy takes the following form

\[ G = \frac{4}{3} \pi r^3 \Delta G_v g(\cos \theta) + 4\pi r^2 \sigma g(\cos \theta) \]  \hspace{1cm} 1-9

in case of homogeneous nucleation, \( \theta = 180^\circ \), the total free energy change is

\[ G = \frac{4}{3} \pi r^3 \Delta G_v + 4\pi r^2 \sigma \]  \hspace{1cm} 1-10
Although people understand some of the basic thermodynamic principles controlling the crystallization of proteins, there is no comprehensive theory to guide our efforts. As an alternative to the predicted approach, success relies on trial and error, including variables such as concentration, pH, temperature and precipitants.

Biomolecules can also be organized into supramolecular structures with specific geometries in order to exert their biological function. In these cases, each biomolecule acts as a modular unit. It is expected that such structures generated by self-assembly are thermodynamically favourable for the biological systems. In such a complex, each subunit is specific not only in its own structure but also in the geometry of binding sites so that every molecule is oriented exactly as expected.

A more complicated example of self assembly would be the formation of a ligand-receptor complex. This process happens between specific interaction sites on the ligand and the receptor. For example, isoleucyl-tRNA synthetase can differentiate between similar amino acids, such as isoleucine and valine, because it contains two catalytic sites, one for activation and another one for editing. In the case when valine fits into the pocket, it can be cleaved by the additional editing site, which will reject the isoleucine. This double checking system ensures the identity of the molecules via the multiple interaction sites, and adjusts any mistakes that may have occurred. (Figure 1-4)
The ordered self-assembly of biomolecules has already been exploited as a novel technology. A good example is the assembly of single-stranded DNA into arbitrary two-dimensional shapes. The exquisite specificity of Watson-Crick base pairing (A-T, G-C) has been known for a long time. Generated from a sticky end, which is a short single strand overhang protruding from the end of a double-stranded helical DNA molecule, DNA fragments can self-assemble into two-dimensional structures, which are similar to Holliday junctions. (Figure 1-5)
Benefiting from the convenience of solid support-based DNA synthesis, a huge diversity of structures, both in shape (rectangular to hexagonal) and size (1 µm to 1 mm) has been created. These two-dimensional structures have been used as templates for the assembly of other materials, such as gold particles.

1.4 Crystallization of Membrane Proteins

The structure of membrane proteins has become an active area of research along with the rapid development of biomedical research, not only for understanding their fundamental biological functions, such as small molecule transport and cellular signalling pathways, but also as specific targets for drug design to cure certain diseases. For example, Losec® (omeprazole) is a proton pump inhibitor targeted at H\(^+\)/K\(^+\) ATPase in the treatment of dyspepsia and peptic ulcer disease.

An estimated 25-30% of the proteins encoded by the genome of an organism are membrane proteins. However, of the 42965 protein structures currently identified (http://www.rcsb.org/pdb/statistics/holdings.do), only 116 of these are membrane proteins. (http://www.mpibp-frankfurt.mpg.de/michel/public/memprotstruct.html). The major obstacle is crystallizing a membrane protein and solving its structure by X-ray diffraction.
By definition, all membrane proteins are amphipathic since they are naturally embedded in a lipid bilayer. Part of the molecular surface, which is in contact with the aqueous environment, is hydrophilic; the other part of the molecular surface is hydrophobic, which is in contact with the alkyl chains of the membrane bilayer. Member proteins will aggregate when removed from the native membrane bilayer and their hydrophobic surfaces are exposed. To keep membrane proteins soluble in solution, amphiphiles, such as detergents and artificial lipids, have been used to protect the hydrophobic surface from random aggregation.

The photosynthetic reaction centre of purple bacterium *Rhodopseudomomas virids* was the first membrane protein that has been crystallized. \(^{36,37}\) Using N, N-dimethyldodecylamine-N-oxide, Michel et al was able to obtain star-like crystals by vapour-diffusion method and the molecular structure was revealed to 2.9 Å resolution. (Figure 1-6) It is fascinating to see how four protein subunits (L, H, M and cytochrome) work together to form the reaction center and how the electron is transferred through the molecule. More importantly, this work illustrated the way to crystallize an amphiphilic membrane protein, an achievement that brought him the Nobel Prize in Chemistry in 1988.
In the presence of detergents, the experimental conditions used to crystallize membrane proteins are identical to those used for soluble proteins, e.g. an unsaturated protein solution is brought to supersaturation by precipitating agents such as salts and/or polyethylene glycols (PEGs). However, the success rate remains very low. A major problem remains in finding a suitable detergent and forming a stable protein-detergent-complex (PDC).

Detergents are amphiphilic molecules, consisting of a polar head group and a hydrophobic chain. They are able to solubilise the membrane protein by creating a structure similar to the natural lipid bilayer environment enclosing the membrane protein. Detergents are classified into four categories based on their structures: ionic, bile acid, non-ionic, and zwitterionic. Ionic detergents, e.g. sodium dodecyl sulphate (SDS), contain a positive/negative charged head group. They are very good in solubilising membrane proteins but could denature the protein to some extent. Bile acid detergents are ionic molecules with a rigid steroidal backbone. They form kidney-shaped aggregates unlike the spherical micelles formed by
traditional ionic detergents. Nonionic detergents contain an uncharged hydrophilic polyoxyethylene or glycosidic head group. They are mild detergents with the least effect on natural protein structure. A number of membrane proteins have been crystallized with the aid of n-octyl-\(\beta\)-D-glucopyranoside (OG) and n-dodecyl-\(\beta\)-D-maltoside (DDM), except that these detergents are relatively prone to hydrolysis. Zwitterionic detergents have a doubly charged head group. They combine the properties of ionic and non-ionic detergents and have more effect on protein structures than non-ionic detergents. There is a minimum concentration known as the critical micelle concentration (CMC). Above that concentration, detergents are able to form micelles in solution. At low temperature, detergent molecules form crystalline structures. At elevated temperature, detergents are solubilised into monomers. As the concentration of monomers increases, micelles will form and the solution appears cloudy.

However, selecting the right detergent is still a trial and error process for several reasons. First, the structure of the protein in question is unknown, so there is no way to predict which detergent can fit into the space between protein molecules. Second, the detergent molecules around the protein are only known to form a “belt-like” structure, and while computational simulations have been used to describe the arrangement of detergent around a protein,\(^{38}\) these are at best simple models.\(^ {39}\) Third, the interactions between the detergent and protein are very subtle. Small structural differences in the detergent can cause significant differences in the crystallization behaviour of these detergent-membrane protein complexes. For example, cytochrome C oxidase can be crystallized using maltosides. However, the crystal formed by n-dodecyl-\(\beta\)-D-maltoside (C12-maltoside) diffracted to a resolution of 8 Å. This can be improved to 2.6 Å when using C11-maltoside, while no crystal formed with C10-maltoside.\(^ {40}\) This suggests that more effort should be put into screening various detergents for crystallization.

An alternative approach is to reconstitute the membrane protein into a two-dimensional lipid bilayer so that proteins are maintained in their native environment.\(^ {41}\) The idea is that by using quasisolid lipid cubic phase, instead of micellar systems, the bilayer
provides nucleation sites for membrane proteins and supports growth by lateral diffusion of protein molecules in the membrane. Rather than X-ray diffraction, cryo-electron microscopy is often used to characterize the structure of two-dimensional crystals formed in this way. As an example, the structure of bacteriorhodopsin was determined to a resolution of 3.7 Å by cryo-EM. In this approach, electron density maps of the sample are generated as a function of sample tilt. Reconstruction of these density maps is then performed to create a three-dimensional representation of the protein of interest.

In a detergent-protein complex, the contacts between adjacent protein molecules are made by the polar surfaces of the protein protruding from the detergent micelles. These hydrophilic domains may be very small for some membrane proteins and the interaction from the polar surface may not be strong enough to stabilise the crystal packing. An alternative strategy to improve the possibility of crystallization is to enlarge the hydrophilic surface of membrane proteins by surface expansion, which can be achieved by attaching antibody fragments to the membrane protein. For example, monovalent antibody fragments, either Fv (\(\sim 28\) kDa) or Fab (\(\sim 56\) kDa) have been successfully used in crystallizing bacterial cytochrome c oxidase, and cytochrome bc1 complex.

1.5 Amphiphilic Molecules

Amphiphilic molecules consist of both polar/hydrophilic and nonpolar/hydrophobic segments. In water, they tend to form water-dispersed structures with clustered hydrophobic moieties to minimize their interaction with water, and their hydrophilic moieties oriented towards the water phase to satisfy specific hydration requirements. The driving force for the formation of these structures is the hydrophobic effect, which stems from the extensive hydrogen bonding that occurs between water molecules and expels molecules that are incapable of forming hydrogen bonds. Amphiphilic molecules can form different structures, depending on the chemical and physical properties of the solute, solvent, temperature, additives and salt concentration. These range from simple one-dimensional lamellar phases to two-dimensional hexagonal phases and more complex three-dimensional cubic phases. The
lamellar phase can be regarded as a critical transition conformation. Below lamellar phase, the polar/apolar interface curves away from the water phase, also called the type 1 phase. Above the lamellar phase, the polar/apolar interface curves towards the water phase, which is referred to as the type 2 phase or reversed phase. With increasing concentration, amphiphilic molecules transition phases from type 1 to type 2.

This phase change is also dependent on the molecular shape of amphiphilic molecules. Amphiphilic molecules with large hydrophilic and small hydrophobic groups tend to form the normal type 1 phase. Molecules with small hydrophilic and large hydrophobic groups tend to form the reverse type 2 phase. The ratio of the hydrophobic tail relative to the size of the hydrophilic head group can be used to predict the type of mesophase that these amphiphilic molecules will form. A shape factor, or critical packing parameter, has been defined as \( \frac{v}{a_{0}lc} \), where \( v \) is the hydrophobic chain volume, \( a_{0} \) is the optimal head group area and \( lc \) is the critical chain length of the hydrophobic chains. When the \( \frac{v}{a_{0}lc} \) is smaller than 1/3, normal type 1 micelles form; when the \( \frac{v}{a_{0}lc} \) smaller than 1/2, cylindrical micelles form; when the \( \frac{v}{a_{0}lc} \) is smaller than 1, lamellar phases form.

---

**Figure 1-7.** Amphiphiles with different shapes. The left-hand amphiphiles (A) tend to form normal phases \( (v/a_{0}lc < 1) \); the middle amphiphiles (B) tend to form lamellar phases \( (v/a_{0}lc = 1) \); the right-hand amphiphiles (C) tend to form reversed phases \( (v/a_{0}lc > 1) \).

The inherent biocompatibility of phospholipids, and their ability to self-organize into ordered structures such as spheres and tubules, has led to their use in the
compartmentalization and template synthesis of new biomimetic materials. They have been used in the biomineralization of bone mineral, marine shell and tooth enamel, \(^{50,51}\) as well as in controlled drug release applications. \(^{52}\) They are generally considered non-toxic, biodegradable and non-immunogenic, and have been shown to change the pharmacokinetics of a drug and lower systemic toxicity.

Phospholipids also play a critical role in certain types of disease formation. In studying the early onset of Parkinson’s disease, it was found that \(\alpha\)-synuclein (\(\alpha\)-Syn), a protein closely related with neurodegeneration, binds specifically to acidic phospholipid vesicles and this binding was significantly augmented in the presence of phosphatidylethanolamine. Atomic force microscopy images clearly revealed the \emph{in situ} formation and expansion of bilayer holes upon the addition of \(\alpha\)-Syn. Close inspection revealed the presence of small aggregates and putative fibrils, which could be either aggregates of \(\alpha\)-Syn or \(\alpha\)-Syn-lipid complex.\(^{53}\) (Figure 1-8) Lansbury Jr. et al. studied the aggregation of \(\alpha\)-Syn on brain-derived membrane fractions.\(^{54}\) They found that \(\alpha\)-Syn monomers form spherical protofibrils quickly and bound to brain-derived membrane fractions more tightly. After being incubated for a period of time, these spherical protofibrils were found to form annular structures, similar to the proposed ion channel model for the action of amyloid \(\beta\) protein (A\(\beta\)P), which is the major component of Alzheimer’s Disease (AD) senile plaques. It is interesting to notice that A\(\beta\)P retain a globular and nonfibrillar shape for an extended period on mica even at high concentrations, but readily reconstituted into donut-shaped structures protruding from the membrane due to oligomerization.\(^{55}\)
Figure 1-8. *In situ* tapping mode AFM images of lipid bilayer disrupted by wild-type α-synuclein. (A) Confluent bilayer of POPC/POPS (1:1) formed on mica before the addition of α-synuclein. (B) After adding the wild-type α-synuclein at 0.1 mg/ml in PBS, the bilayer gradually undergoes deformation, the aggregates of α-synuclein can be found in the zoomed image. (C) After 9 hours, protein aggregates and putative fibrils can be found, as indicated by the arrows. Scale bar is 2 µm. (Adapted from reference 53.)

1.6 **Elastin and Tropoelastin**

The physiological function of many tissues requires that they possess elastic properties. For example, inspiration is an active, energy-requiring process, whereas expiration is a passive process due to the elastic recoil of the respiratory tree. For vertebrates, elastin is the protein responsible for this elastic property, attributing for ~90% of the amorphous fraction of the elastic fibers. One thing that is fascinating about these elastic fibers is that they are superior to synthetic fibers. It has been proposed that the elastomeric force results not from stressing of chemical bonds, but rather from a decrease in the number of conformations of the linked peptide chains.

As the primary component of elastic, insoluble fibers, elastin is the organized assembly of tropoelastin, a soluble ~72 kDa precursor protein. Human tropoelastin consists of alternating hydrophobic and cross-linking domains generated from its 34 exons. The hydrophobic domains are made up of repeat segments containing glycine (G), proline (P) and valine (V). The cross-linking domains have a high lysine (K) and alanine (A) content. Tropoelastin rapidly assembles into elastic fibers on specific regions of the cell surface without
proteolysis. It is protected by the elastin-binding protein (EBP), which behaves like a molecular chaperone, intracellularly from degradation and coacervation. EBP primarily binds to the VGVAPG sequence in human tropoelastin. When the EBP/tropoelastin complex is secreted onto the cell surface, EBP is recycled through endosomal compartments to bind to newly synthesized tropoelastin.

Tropoelastin is interesting in that it is highly hydrophobic and yet soluble. The hydrophobic amino acids are separated by small cross-linking segments. Tropoelastin undergoes a reversible aggregation process, referred to as coacervation, when the temperature approaches 37 °C. In a recent study, it was found that soluble tropoelastin increases in size from 15 nm at 29 °C to ~5 µm at 37 °C with no detectable intermediates. When the temperature is reduced to 25 °C, the monomeric species return. SEM images clearly revealed the spherical shape of the tropoelastin droplets. In the presence of the cross-linking reagent, *Pichia pastoris* lysyl oxidase, these droplets clustered into larger structures and formed fibers after incubation.

Debelle et al. studied the secondary structure of human elastin by FT-IR and CD. Their results revealed that elastin contains a higher amount of β-strands (~ 35%) than α-helices (~ 10%). In the relaxed state, elastin tends to form an unordered structure as solvent water strongly interacts with hydration water. When elastin is stretched, the hydration water is excluded from the interstitial space so that elastin tends to maintain its original structure. Using molecular dynamics simulations, Li et al. confirmed that the stretched backbone of elastin is more dynamic in the extended state above the transition temperature, thus hydrophobic hydration is the main source of the entropy-based elasticity of elastin. Similar result has also been reported by Rauscher et al. from analysing the hydration and disorder of the peptide backbone, in which it is shown the rigid molecular structure of proline is important in keeping the elastomeric polypeptide chain from forming ordered stable structure, which could be the reason for amyloid-like fiber formation.
Using recombinant technology, a series of tropoelastin-like polypeptides, based on the human tropoelastin sequence, have been used as model systems to study the effect of composition on coacervation temperature and fibre structure. It is found that the mechanical properties of membranes formed from these fibres were very similar to the native ones. Due to their biocompatibility, these membranes can be used to reduce the adsorption of platelets in cardiovascular devices such as vascular conduits and arterial/venous catheters.

The temperature-dependent aggregation property of these tropoelastin-like polypeptides also makes them useful as potential local delivery vehicles for chemotherapy drugs and DNA. Linked with the silk-like segment GAGAGS, elastin-like polypeptides are able to release active DNA for two weeks after being injected into a tumour grown in a mouse model.

We recently reported the temperature-dependent ordered self-assembly of a series of tropoelastin-like peptides on hydrophilic mica and hydrophobic HOPG. EP 20-24 (EP I) corresponds to exons 20-21-23-24 from human tropoelastin cDNA; EP 20-24-24 (EP II) corresponds to exons 20-21-23-24-21-23-24 from the same source. In EP 20-24-24-24-24 (EP IV), exons 21-23-24 are repeated four times. Exon 22 is missing between exon 21 and exon 23 because it has never been expressed. These EP peptides contain one or more 7-fold repeats of highly hydrophobic PGVGVA segment, which is encoded by exon 24. They also contain the AAKAAAKA segment from exons 21-23, which is believed to participate in the assembly of EP peptide. The coacervation temperature of these peptides is inversely related with molecular mass. The smallest EP I has the highest coacervation temperature while EP IV has the lowest coacervation temperature. At the same temperature, the concentration required for EP I to coacervate is much higher than that required for EP IV. (Table 1-1)

We reported that these EP peptides formed amorphous aggregates on mica but ordered hexagonal arrays of well-defined single fibers on HOPG. We believe that this is due to hydrophobic interactions between the molecules and substrate. We also suspect that rather than an epitaxial relationship between the fibres and the HOPG surface, the EP molecules self-assemble into a close-packing motif on the hydrophobic HOPG surface and this in turn leads to the formation of EP fibrils on the pre-layer. Since the EP peptides can be synthesized recombinantly, it is an excellent model system to study the effect of hydrophobic content and sequence on protein self-assembly.

1.7 Atomic Force Microscopy (AFM)

The invention of scanning tunnelling microscopy (STM) in 1982, which generates sample topography by measuring the quantum tunnelling current between the conductive tip and the surface of an object, has led to the development of an entire family of scanning probe microscopes (SPMs). The SPM maps a sample surface by measuring the near-field interactions (chemical, electronic, physical) between the raster-scanning probe tip and the sample. Within this family, the atomic force microscope (AFM) is a particularly important tool for studying biological samples.
AFM generates a sample surface topography by measuring the inter-molecular forces, e.g. Van der Waals’, electrostatics and hydrogen bonding, between the scanning tip and subject surface. These interactions can be represented by a simple mathematical model, Lennard-Jones potential, in which the energy between two particles at a certain distance is described as:

\[
W(r) = -\frac{A}{r^6} + \frac{B}{r^{12}}
\]

where \( r \) is the separation of two atoms, the inverse sixth-power of \( r \) represents the attractive interaction and the inverse twelfth-power represents the repulsive interaction; \( A \) and \( B \) are interaction constants, \( A \approx 10^{-77} \text{ Jm}^6 \), and \( B \approx 10^{-134} \text{ Jm}^{12} \). The interaction force between two molecules is:

\[
F = -\frac{dW(r)}{dr} = -\frac{6A}{r^7} + \frac{12B}{r^{13}}
\]

From this distance and interaction relationship, when an AFM tip is approaching to the sample surface, it moves from the non-contact regime to the contact regime, and the interactions change from attractive to repulsive. (Figure 1-9)
AFM usually contains four main parts. Although there could be slight differences between vendors, a typical layout includes a tip attached to a flexible cantilever, a laser signal and a photodetector sensing the movement of the tip, a piezoelectric scanner controlling the X-Y-Z position of the sample, and a feedback loop controlling the movement of these three parts (Figure 1-10). While the tip raster scans the sample surface, the laser beam is projected off the back of cantilever onto the photodetector. The tip deflection caused by tip-sample interaction is reflected as lateral or vertical shift of the laser spot on the photodetector. This shift is used to adjust the X-Y-Z position of the scanner and the position of the sample on top of it, to maintain the distance between the sample and the scanning tip at the prescribed setpoint value, which corresponds to the specific interaction force between the tip and sample. The X-Y-Z position of the sample generated from this process is recorded as the topography of the sample, also known as the height image. Meanwhile, the bending of the tip can also be used to generate the deflection image, which is the error signal image coming from the photodiode detector’s differential signal. Deflection images collected with high feedback gain essentially is the derivative of the height, which provides a sensitive edge-detection technique.
Figure 1-10. Schematic of AFM components. The laser beam is projected onto the back of the tip and reflected onto the photodiode detector. The feedback loop controls the movement of the piezoelectric scanner, on which the sample is positioned. Right-hand side is the image of Multimode SPM (Veeco, Santa Barbara, CA) equipped with an "E" scanner, with a lateral scan area of 14.6 × 14.6 µm.

With little sample preparation and no need to stain the sample, the AFM is able to produce high-resolution images of sample topography in aqueous/liquid environment, with nanometer resolution laterally and Angstrom resolution vertically. This unique ability enables the AFM to capture real-time images and provides a true three-dimensional surface profile of the sample.

Over time, the AFM has undergone numerous modifications and new imaging modes have been developed. For example, it was found that the contact mode AFM, the primary imaging mode in which the tip is held in constant contact with the sample surface, could cause sample damage. Hansma et al. invented a new way of oscillating the tip at a high frequency over the sample within a very close distance, instead of physically touching it. (Figure 1-11) Dampening of the amplitude of the oscillating tip due to the interaction between the tip and the subject is used to generate the topography of sample. Known as intermittent contact or tapping-mode AFM, this technique enables the imaging of soft samples, e.g. DNA molecules adsorbed on mica. Additionally, the contamination of the tip is dramatically reduced due to the high-frequency oscillation of the tip and minimal interaction force.
The invention of AFM has enabled many interesting studies of dynamic biological process, such as short-term and localized effects of amyloid β peptides (AβP) on endothelial cells. It is found that cells treated with AβP<sub>1-42</sub> undergo significant morphological changes compared with untreated cells within 10-15 minutes, including somal (cell body) shrinkage and plasma membrane blebbing. The morphology change is more obvious at the periphery of the cell than in the central region, and was seen as retraction of the cell body and loss of cytoskeletal structure. (Figure 1-12) These changes are commonly used as indicators of cellular degeneration. These data demonstrated that AβP<sub>1-42</sub> is the most effective in inducing degeneration of the cells, compared to AβP<sub>1-40</sub> and AβP<sub>25-35</sub>. It is also found that these effects could be prevented by anti-AβP antibody and removal of extra-cellular calcium, which suggests that AβP forms a calcium-permeable channel that destroys the cell. This work provides strong proof for the hypothesis that the soluble AβP may be the toxic agent responsible for Alzheimer’s and Parkinson’s disease.
Figure 1-12. Continuous change of endothelial cell morphology captured by *in situ* AFM. A series of AFM images clearly show the cell body is retracting after treated by AβP$_{1-42}$ (a-c). This effect can be prevented by anti-AβP antibody even after the same exposure time, there is little change on cell morphology as shown in images d-f. (Adapted from reference 73)

Its high resolution imaging ability has enabled the AFM to investigate changes in the structure of cell membrane associated with exocytosis. Jena et al. captured the “pits” and “depressions” at exposed apical regions on isolated pancreatic acinar cells. The size of the pits range from 500 nm to 2000 nm, containing depressions from 100 nm to 200 nm. Stimulated by 20 µM Mas7, an active Mastoparan analog that induces the secretion of amylase from zymogen granules, the diameter of the pits increased by 35% after 5 minutes and the pits recovered after 30 minutes. This experiment is important not only for successfully resolving a physiological phenomenon on a single cell that is unable to be resolved by other techniques.
such as optical or electron microscopy, but also for proving that AFM is suitable for imaging very soft living biological samples.

The AFM is a particularly useful tool for studying protein crystallization. Ward and Yip used AFM to reveal the packing difference between human, porcine and bovine ultralente insulin.\textsuperscript{75} Although all three forms have R3 space group symmetry, the hexamers from human and porcine insulin are arranged in the normal close packed hexagonal arrays in the (001) plane, while the bovine insulin exhibits bundles of parallel corrugated cylinders, which represents the morphology of close-packed stacks of insulin hexamers on either the (010) or (110) surface. The reason for the difference lies in the fact that there are more hydrophobic domains on the upper and lower surface of (donut-shaped) bovine insulin hexamer than the human and porcine insulin. This difference in the packing motif explains the difference in the stability of insulin crystals, e.g. why the human and porcine insulin are relatively easy to be dissolved by solvent than the bovine insulin, and why bovine insulin has a more desirable basal time action.

AFM image studies have provided a large amount of useful information on the fundamental mechanisms of crystal growth.\textsuperscript{76,77} Besides the ability of directly visualizing various growth pattern of protein crystallization, such as screw dislocation, complex dislocation and two-dimensional dislocation, AFM has been used to study the nucleation and growth mechanisms at the single molecule level.\textsuperscript{78} In studying the crystallization of thaumatin, a single protein molecule’s attachment onto the step edge and the nucleation of new molecular rows was analyzed. It was found that molecules incorporate into the crystal individually through a rapid nucleation process, instead of diffusing along the surface until they associate with an existing growth step or attach to other molecules to form nuclei. It is also found that the thermal fluctuations are not sufficient to affect the detachment of molecules from a step edge, even at low concentrations. With its high resolution imaging ability, AFM is able to discriminate enantiomorphs from the height difference in the fungal lipase crystals, which was not resolved by X-ray diffraction.\textsuperscript{79}
AFM has also been used to measure interaction forces between molecules. The AFM cantilever can be viewed as a spring with a spring constant of $k$. According to Hooke’s law, the force exerted on the spring can be calculated from the deflection of the tip ($\Delta Z$) multiplied by the spring constant $k$. (Figure 1-13)

$$F = -\Delta Z \cdot k$$

Figure 1-13. Schematic of AFM force spectroscopy experiment depicting the deflection of cantilever and the analogy to the extension of a spring under an applied force.

This is a very sensitive approach that has been used to measure the interaction forces between molecules, e.g. van der Waals, electrostatic and solvation forces, in the nN to pN range under a wide variety of conditions. For ligand-receptor complexes, the recognition force measurement has been used to examine the thermodynamics and kinetics of protein association at the single molecule level. $^{80,81}$ AFM has also been applied to study insulin self-association forces, revealing that insulin unbinding is a combination of molecular disentanglement and dimer dissociation processes. $^{82}$

Despite the power of AFM, there are some key challenges that remain. AFM is a surface technology and thus unable to image molecules floating in solution. All samples are necessarily bound to a surface, so they may behave differently than in solution, possibly reflecting non-native sample-substrate interactions, especially in the case of protein adsorption.
The shape of the AFM stylus will affect the images, often generating an apparent lateral size that is much larger than the actual value. (Figure 1-14). There are a number of ways to compensate for this. One theoretical approach used the measured width at half-height of the sample, $W_m$, and the radius of the sample $r$ can be calculated as:

\[ W_m = \frac{2[R(1-\sin\theta)+r]}{\cos\theta} \]  \hspace{1cm} \text{for } \sin\theta > \frac{R}{R+r}  \tag{1-14} \]

\[ W_m = 2\sqrt{r(r+2R)} \]  \hspace{1cm} \text{for } \sin\theta < \frac{R}{R+r} \]

where $R$ is the radius of the tip, and $2\theta$ is the cone angle of the $5−10$ nm radius tip. For example, if the measured middle width is 20 nm and a tip cone angle is 70°, from the above equation, the actual radius of the sample is ~ 4.9 nm.

In an AFM single molecule force spectroscopy experiment, it is typically assumed that the pulling direction is parallel to the extension of the molecule, so that the force measured is the same as the force applied on the molecule. However, this is not necessarily the case. The
force measured through the deflection of the AFM cantilever, \( F_z \) (force in the Z direction), is only part of the force applied to the molecule. This problem is more significant for short and rigid molecules rather than long and flexible molecules like titin. Using a home-made AFM, Marszalek et al. measured the effect of pulling geometry on the force-extension relationship of double-stranded \( \lambda \)-phage DNA.\(^{84}\) It is found that increasing the pulling angle can significantly decrease the force associated with the characteristic stretching transition.

1.8 Dynamic Light Scattering (DLS)

Light scattering is a common everyday phenomenon, such as a light beam coming through the window into the dark room, which is reflected from tiny dust particles suspended in the air.\(^ {85}\) Most of the scattered light has the same wavelength as the incident light and this is called elastic (or Rayleigh) scattering. Static light scattering (SLS) measures the angular dependence of the time-averaged intensity change \( (I) \) in terms of wave factor \( q \) from scattered light of sample solution, which is proportional to \( S(q) \) – the structure factor and \( P(q) \) - the form factor.\(^ {86}\) In order to relate the intensity of the scattered light to the properties of molecules in solution, background scattering must be subtracted and the latter should be normalized by the incident light. After subtraction and normalization, the scattered light from the molecules in solution can be expressed as the excess Rayleigh ratio \( (R_\theta) \):

\[
R_\theta = \frac{I - I_0}{I_{st}} \left( \frac{n_o}{n_{st}} \right)^2
\]

\( I, I_0, I_{st} \) are scattered intensity of solution, solvent and standard, respectively, \( n_o \) and \( n_{st} \) are reflective index of solvent and standard, respectively. Hence,

\[
R_\theta = KMcP(q)S(q)
\]

where \( M \) is the molecular weight of the molecules in the solution, \( c \) is the molecular concentration and \( K \) is the system specific constant given by:
\[
K = \frac{4\pi n_0^2}{N\lambda^2} \left( \frac{dn}{dc} \right)^2
\]

with \(N_A\) is the Avogadro’s number, \(dn/dc\) is the reflective index of molecule.

When the molecule concentration is low so that the inter-particle interactions are very small, the structure factor \(S(q)\) can be neglected. When the molecule size is very small in comparison to the incident wavelength, the form factor tends to 1. To measure \(R_g\) (gyration radius), e.g. to see the difference in the angular dependence of scattered light intensity, the molecule size should be > 10% of the wavelength of the incident light beam. (e.g. for \(\lambda = 514\) nm, the molecule size should be bigger than 51.4 nm).

Moreover, the excess Rayleigh ratio can be used to estimate the weight-average molar mass \((M_w)\) of molecules from equation (2) when \(S(q)\) and \(P(q)\) approach 1, so that

\[
R_\theta = K \sum_i c_i M_i = K \sum_i n_i M_i^2
\]

By measuring the scattered light intensity at different angles and different sample concentrations, the excess Rayleigh ratio could be extrapolated to \(\theta = 0\) and \(c = 0\). The offset value is the equivalent to \(1/M_w\). This diagram is also referred as a Zimm plot.

However, a small amount of the scattered light has a shifted frequency from the incident light, which could be caused by the real-time (Brownian) motion of molecules, also known as the Doppler effect. This quasi-elastic scattering can be used to measure movement of the molecules. If the incident light has a narrow frequency bandwidth, smaller than the frequency shift, this shift can be measured by a specialized interferometer and spectrum analyzer using a technique known as dynamic light scattering (DLS).

In DLS, a very sensitive photomultiplier and a very small aperture are used to measure the number of photons of scattered light during a series of short time intervals. The intensity autocorrelation function is given as:
\[ G^{(2)}(\tau) = A(1 + \beta |g^{(1)}(\tau)|^2) \quad 1-19 \]

\( G^{(2)}(t) \) is the unnormalized intensity time correlation function; \( A \) is the background, \( \beta \) is a coherence factor and \( g^{(1)}(\tau) \) is the normalized electric field autocorrelation function. For a monodisperse solute undergoing Brownian motion, \( g^{(1)}(\tau) \) is a single exponential:

\[ g^{(1)}(\tau) = \exp(-\frac{\tau}{\tau_0}) \quad 1-20 \]

where \( \tau_0 \) is the characteristic correlation time of the scatterer motion. For simple translational diffusion, the decay rate is:

\[ \frac{1}{\tau_0} = Dq^2 \quad 1-21 \]

where \( D \) is the translational diffusion coefficient. The scattering wave vector magnitude, \( q \), is:

\[ q = \frac{4\pi n \sin(\theta / 2)}{\lambda} \quad 1-22 \]

where \( n \) is the reflective index of the liquid, \( \lambda \) is the wavelength of the laser, and \( \theta \) is the scattering angle. For a diluted solution, \( D \rightarrow D_0 \), where \( D_0 \) is given by the Stokes-Einstein equation:

\[ D_0 = \frac{kT}{6\pi \eta R_H} \quad 1-23 \]

\( k \) is the Boltzmann constant, \( T \) is the absolute temperature, \( \eta \) is the viscosity of the solution, \( R_H \) is the hydrodynamic radius of the measured particle.
Figure 1-15. Schematic of difference in the incident light and scattered light in the light scattering measurement. It is shown that the constructive interference and high scattered intensity found at lower angles. Strong destructive interference and low scattered intensity happens at higher angle, 90°.

DLS is a powerful technique able to detect particles ranging in size from several nanometers to several microns. (Figure 1-16)
Figure 1-16. A typical dynamic light scattering measurement of 100 nm polystyrene latex beads. The effective diameter is 93.6 nm and the polydispersity is 0.139, which indicates the particle size is monodisperse in solution.

However, compared with AFM, the applications of DLS are limited as it is very sensitive to impurities, such as large particles, contaminants, and air bubbles. The particle size measured is the hydrodynamic radius, which is generated by particle motion, including the water molecules surrounding the particle so that it may appear larger than expected. For samples having a similar reflective index as the background, it is difficult to differentiate them from the background, so a higher concentration is needed to obtain a stable DLS signal.

AFM is an excellent technology in providing direct evidence of sample association on surface at high resolution. As a complementary technique, DLS can provide insights about the similar process happening in solution. Together, AFM and DLS are particularly useful for
examining protein self-assembly at surfaces and in solution. By applying these two techniques to study the self-assembly of EPs, we will develop an understanding of the mechanisms of EP self-assembly at the single molecule level.
2 Multi-Layer and Nanoconfined Self-Assembly of EP on HOPG

We previously reported the temperature-dependent ordered self-assembly of tropoelastin-like peptides (EPs) on HOPG. It was found that the EP molecules first formed a thin layer on HOPG surface. Subsequent molecules from the solution absorbed onto this thin layer in a hexagonal pattern, forming fibrils with intersection angles of 60° or 120°. (Figure 2-1(A) and (B)) These fibrils were able to further extend into parallel aligned fibers covering most areas of the HOPG surface. This process is likely a nucleation process akin to the Stranski-Krastanov crystal growth model. Compared with the amorphous aggregates of EP that form on hydrophilic mica under same condition (Figure 2-1(C) and (D)), we believe this process is driven by hydrophobic interactions between the EP molecules and the HOPG surface.
Figure 2-1. *In situ* TMAFM images of the rapid nucleation and growth of EP I fibrils on HOPG at 50 °C. Note that the elapsed time between images (A) and (B) is ~2 h. (C) and (D) show the amorphous aggregates of EP I and EP II on mica at room temperature. (Reproduced with permission from *J Am Chem Soc* 2002, 124, 10648. Copyright © 2002 American Chemical Society.)
2.1 Ordered Self-Assembled Multilayers of EP Fibrils on Graphite Surface

Since the EP peptides self-assembled into multilayer fibers in a hexagonal pattern on HOPG, this suggests an interesting approach for investigating the possibility of creating ordered three-dimensional EP arrays.

Using the optimized temperature control system (see Appendix), the desired experimental temperature is easily stabilized with an error of less than 1 °C, so that we are able to observe the self-assembly of EP over extended periods. It is interesting to find the EPs form multilayers of ordered fibrils which was not observed before. At 50°C, EP I molecules assembled into hexagonal pattern as we reported before. We found this ordered structure developed on top of the first ordered layer. To clarify these layers at different levels, we named the first amorphous layer on the HOPG surface as layer A, the first ordered layer as layer B, and the second ordered layer as layer C, and the third ordered layer as layer D. (Figure 2-2)

Figure 2-2. Schematic illustration of EP multi-layers self-assembled on HOPG surface. A is the first amorphous thin film layer, B is the first ordered layer of fibers in hexagonal pattern, C is the second ordered layer and D is the third ordered layer.
At 50 °C, the ordered layer B formed completely on top of the amorphous layer A, followed by the hexagonally patterned ordered layer C. A two-dimensional fast Fourier transform (FFT) analysis revealed the characteristics of the alignment of those fibrils within layer B and layer C. An FFT is a discrete Fourier transform algorithm used to discover periodic features within an input signal or image. It breaks down a signal into constituent sinusoids of different frequencies, from which it is able to find out a particular frequency with outstanding strength contained within the signal. For a two-dimensional signal that is separable, e.g. a topography image containing a countable subset, it is possible to calculate a two-dimensional FFT as two successive one-dimensional Fourier transforms. The two-dimensional FFT has been used as an image processing tool for image filtering, image reconstruction and image compression.

Using two-dimensional FFT analysis, it was found that the orientation of fibrils in the C layer were rotated ~30° relative to the ones in the B layer, but remained hexagonally arranged. The section analysis of these fibrils revealed that they are ~0.65 nm thick. (Figure 2-3)
Figure 2-3. Hexagonal multilayer assembly of EP I on HOPG at 50°C. Image (B) was captured four hours later than (A), in which two distinctive layers, layer B and layer C, can be clearly identified. (C) The close-up view from image (A), the arrows indicate the fiber orientation. (D) 2D Fourier power spectrum of (A), solid arrow refers to layer B and dotted arrow refers to layer C. (E) and (F) are the section analysis of the height of fibers seen in image (A) and (B).
Detailed characterization of these fiber multilayers provided very interesting information about their hexagonal arrangement. The radius of outer pattern in the 2D Fourier power spectrum is $14.7 \pm 0.4$ nm; the intersection angle is $60^\circ \pm 2.0^\circ$, shown as the solid arrow in Figure 2-3(D). The radius of the inner pattern is $18.9 \pm 0.4$ nm; the intersection angle is $60^\circ \pm 2.5^\circ$, shown as the dashed arrow in Figure 2-3(D). The average length of the second layer fibers are $100.6 \pm 25.0$ nm. Since the FT spectrum extracts the periodicity of the signal, the radius of the pattern represents the distance between fibers. If we divide the 100.6 nm by 14.7 nm, it equals $\sim 7$, which means every fiber within the C layer crosses approximately seven B layer fibers.

Close inspection of the fibers revealed that they are oriented independent of the step edge of the graphite surface, as shown in Figure 2-3(C). The absence of edge-directed growth or nucleation suggests that there is no preferential stabilization of EP by the HOPG steps. These multilayer fibrillar structures were captured assembling from molecules of smaller units. From the sequential images shown in Figure 2-3(B) it can be seen that molecules are adsorbed onto the B layer to form those fibrils. We also can see a few molecules sitting on top of the C-layer fibrils. Height analysis shows they are 0.64 nm thick, suggesting that these may be part of the D layer described earlier.

The growth of the C layer occurs via independent nucleation events. In Figure 2-4, there are three separate domains oriented in the same direction expanding on the supporting layer. It is clearly shown, as indicated by the circles in Figure 2-4(A) and (B) that the area of the C layer increases after three hours. This demonstrates the possibility of forming another complete ordered EP fiber layer on top of the existing B layer. The orientation shift between two neighbouring layers is a consequence of the close packing between two layers. The multilayer formation is also due to the increased assembly kinetics at elevated temperature.
This continuous growth of multilayer assembly of fibrillar structure is a novel phenomenon that has not yet been reported for EP peptides. We think it could be interesting to apply it to further investigations of the crystallization of these peptides.

2.2 **Trial of Crystallizing EP on Graphite Surface**

Using the hanging droplet method, approximately 10 µl of the EP solution was placed on the surface of freshly cleaved HOPG, inverted over a 500 µl reservoir of coacervation buffer and sealed with vacuum grease. The sample was left in place at 50 °C for EP I, 40 °C for EP II. Normally after 3 days, little particles were found on the graphite. (Figure 2-5, 2-6)
Figure 2-5. Optical microscope image of particles formed on HOPG at 50 °C from 25 μM EP I solution at pH 7.5.

Figure 2-6. Optical microscope image of particles formed on HOPG at 40 °C from 25 μM EP II solution at pH 7.5.
However, the size of particles is too small to be detected by X-ray. To obtain a larger particle, the concentration of EP II solution was increased to ~125 µM. Using the same setup, we obtained a particle larger than 100 µm. (Figure 2-7)

Figure 2-7. A larger particle formed on HOPG at 40 °C from 125 µM EP II solution at pH 7.5. Inserted is the X-ray diffraction pattern of this particle. The diffraction ring at 3.86 Å is from the ice within the particle.

Although the particle seen in Figure 2-7 is quite large (~ 100 µm in diameter), it did not generate a diffraction pattern, which suggests that the molecules within the particle are not arranged in an ordered pattern. To improve the packing of molecules, a better understanding of the nucleation events that precede fibril formation is required.

2.3 The Growth of EP Fibers within Corrals on Graphite Surface

Nano-confinement via thermally etched mono- or multilayer pits, on the (0001) basal plane of HOPG has been used to study the epitaxial ordering of liquid crystal 4’-octyl-4-
biphenylcarbonitrile (8CB) and the electrochemical nucleation of type I and type II form of bis(ethylenedithiolo)-tetrathiafulvalene triiodide, (ET)$_2$I$_3$. It is found that the pit edge perturbs the ordering of the thin liquid crystal layer and that nucleation of (ET)$_2$I$_3$ is inhibited within pits with diameters less than 100 nm. It is suggested that the pit edge serves as surface barrier preventing the diffusion of molecules from surrounding area, which provides us with an ideal template to study the nucleation and growth of EP fibrils.

To study the growth of EP peptides within confined areas, we needed to find the appropriate conditions for forming corrals of a desired size in HOPG (Grade SPI-3, Structure Probe, Inc., West Chester, PA). We treated HOPG at 700, 800 °C for 7, 6 and 5 minutes, separately, and measured the size of the resulting pits. It is found that the corrals formed after a 6 minute exposure at 800 °C were ~ 435 nm in diameter (Figure 2-8(A)). The corrals formed at other conditions are either too big or too small. For example, annealing at 800 °C for 7 minutes resulted in extended crevices on the HOPG surface (Figure 2-8(B)). In all the experiments described here, the HOPG surface was heated for 6 minutes at 800 °C.

![Figure 2-8. Corrals formed on HOPG after thermal-etching treatment. (A) The HOPG surface heated at 800 °C for 6 minutes, (B) The HOPG surface heated at 800 °C for 7 minutes.](image)
The procedure of imaging fiber formation was similar to that described earlier. The EP I peptide solution was put on heat-treated HOPG surface and sealed in the contact/tapping liquid cell. The temperature was raised to 50°C to image fiber growth, both within and outside the corrals. (Figure 2-9)

![Figure 2-9](image)

Figure 2-9. Two images of EP I peptide grown within and outside the HOPG corrals at 50°C. Image (B) is captured 28 minutes after image (A). (Height scale: 20 nm)

To calculate the fiber growth rate, we manually selected individual fibers outside and inside the pits from a series of consecutive images (Figure 2-9), which were captured in the same scanning direction and at the same scanning rate. All the fibers grew in a highly oriented hexagonal pattern. Within the pit, no fibers nucleated from the edge. All of them nucleated independently at isolated sites within the pit and grew from both ends. In this case, surface stabilization of the EP fibril nuclei on the HOPG surface was stronger than edge-directed stabilization, a process similar as the multilayer fiber growth described earlier.
The observed growth rate of the fiber in the pit reached a maximum of 13 nm/min, with an average rate of 9.5 nm/min. For the fibers growing in the open area, we observed a maximal rate of ~8 nm/min with an average rate of 5 nm/min. The growth rate for the fiber inside the pit is faster than that on the surrounding area, which can be explained by the surface discontinuity created by the pit edge that prevents the diffusion of molecules from the terrace to the pit, and the fact that there are fewer nucleation sites within the limited area. More molecules can assemble at one nucleation site and thus the fibers grow faster.

The size and shape of the pits do affect fiber formation. There are certain pits that cannot be filled with hexagonal fibers even after an extended incubation time. (Figure 2-10) Usually they are of smaller size, and they remain as defects on the surface after other pits have been completely filled.

Figure 2-10. A small pit remains visible after surrounding area has been covered by hexagonal fibers. The height and width of the pit can be seen from the section diagram.

In Figure 2-10, there is a pit surrounded by ordered fibers but remained unfilled even after a two-hour incubation. The depth of the pit is about 3.5 nm with a corresponding width.
of ~ 78 nm. This absence of fibers within this pit is either because the 2D space within the pit is not large enough for fiber formation, or there is not sufficient material within the pit to form the fiber. Within the pits with larger surface areas, the growth of fibers is not restricted, even though they are much deeper. (Figure 2-11)

![Figure 2-11. Ordered EP fibers formed within deep pits. The fibers can be clearly identified from the amplitude image (left), while the height image (right) only reveals that it is a very deep pit. The section diagram shows this deep pit has a flat bottom.](image)

In Figure 2-11, the depth indicated by the white arrow (♦) is about 37 nm; the depth marked by the green arrow (♦) is about 14 nm. Both of them are much deeper than the pit shown in Figure 2-10. However, the fibers still grow in a highly ordered pattern, as can be
found in the amplitude image. This demonstrates that pit depth does not affect fiber nucleation.

Comparing a series of pits, it appears the smallest size/diameter capable of supporting fiber formation is 121 nm. However, the structure of the pit is more critical than the size of open surface area, especially with regards to the curvature within the pit. In a pit having a diameter of 342 nm, only individual globular particles of EP I are found within the area even after 3 hours of incubation. (Figure 2-12)

Figure 2-12. Within thermally-annealed pits having multiple (0001) layers of narrow step edges, EP peptide molecules were unable to form extended fibers as found on the open area. Image (B) was taken 3 hours later than image (A) at the same position. No fibers formed within the lower area, except the individual globular aggregates. (C) The section diagram of the pit showing the curvature.
The formation of pits on thermally treated HOPG is generated by the annealing of defects within the graphite structure. The pit shown in figure 2-12 is 8.4 nm high, likely consisting multiple (0001) layers with narrow step edges, in which the layer by layer distance is 3.4 Å. The sizes of individual globular particle found on the curved surface provide very useful information about critical size for the peptide fiber formation. The height of the globular particles is ~ 3.5 nm, the observed width is ~ 22 nm. Assuming a nominal AFM tip radius of ~ 10 nm, the calculated diameter of the globular particle is ~ 9.6 nm, so that the volume of this semi-sphere is about 42.2 nm³. There were no long fibers found within this region. The measured fiber height is ~ 1 nm; the width is ~ 20 nm and the estimated width ~ 4 nm. Assuming the width is the radius of one single peptide molecule, the volume calculated from the same semi-sphere model is ~ 8.4 nm³, which means a single globular sphere could contain 4 - 5 single peptide molecules. The length of the shortest fibers observed was ~ 32 nm, most likely it is caused from a single sphere spreading into a short fiber on HOPG surface. Assuming a protein mass density of ~ 1.35 g/cm³ and an EP I MW of 10010 Da, this suggests that these aggregates contain about 3.4 EP molecules. This is in good match with the volume determination.

We propose that these globular spheres are the smallest stable EP aggregates. They are the nuclei for further fiber growth when the substrate is large enough. If the available space on the graphite terrace is too small to accommodate more molecules, they will only form these globular particles or may not even be able to stay on the substrate, as shown in the schematic. (Figure 2-13)
2.4 Conclusion

Elastin is a structural protein. The presence of mainly non-charged amino acids (>95% for EP I) in its sequence suggests that there are no significant interactions between EP molecules that would dominate their assembly. In solution, molecules tend to aggregate by hydrophobic interactions. This is the main reason why crystallization of these peptides is difficult.

Our observations of the multilayer assembly of the EP peptides is the first time that a clearly resolved layer by layer growth of biomolecules on ordered substrates has been reported. The observed domains are similar to the nuclei seen in protein crystallization. Compared with normal growth pattern of crystals, in this process, the number of molecular interactions involved in stabilizing a molecule is less. It is also very interesting to note that the growth of the upper layer is mainly mediated from the ordered pre-layer.
3 Self-Assembly of EP-Detergent Complex at Solid-Liquid Interface

The structure of EP remains unknown, as are the mechanisms of its self assembly, due to its tendency to aggregate into amorphous structures in solution. Prompted by reports of the detergent-aided crystallization of membrane proteins, we believe an interesting approach would be to use detergents to protect the hydrophobic surface of EP by forming an EP-detergent complex thereby preventing EP aggregation, and facilitating its ordered self-assembly.

The structure of EP is similar to that of a membrane protein. Both have large hydrophobic surfaces that need to be protected to prevent random aggregation. There are two possibilities for arranging membrane proteins with detergents. The “Type I” arrangement involves embedding the proteins in 2D planes and stacking these planes. The other way, known as “type II”, is to crystallize membrane proteins within detergent micelles. The 3D crystal is established through interactions between the polar contacts of the hydrophilic part of the membrane protein extending out of the detergent micelles. (Figure 3-1)

However, the structural character of EP is still quite different from membrane proteins, from the fact that it does not have many polar amino acids so that hydrophilic interactions between molecules are weak. We decided to test the association of EP with a series of detergents that have been used successfully in the crystallization of membrane proteins.
3.1 Ordered Self-Assembly of Detergent on HOPG

The first experiment carried out was the trial of mixing 5 µM EP II with 1 CMC (250 mM) n-Hexyl-β-D-glucopyranoside (HG), and studying its aggregation on HOPG as a function of temperature. It is surprising to find that this mixture formed hexagonally arranged fibers on HOPG, at room temperature. When the sample is heated up to 35 °C, the number of fibers decreased and finally dissolved away from the surface. When the temperature returned to room temperature, the fibers spontaneously re-assembled in their original pattern. (Figure 3-2)
Figure 3-2. Self-assembly of EP II with Hexyl-β-D-glucopyranoside (HG) on HOPG surface as a function of temperature. (A) The image captured right after the mixture was deposited on the HOPG at room temperature. Self-assembled fibers in a hexagonal pattern can be clearly identified besides the big particles. Upon heating to 35°C, the fibers start to dissolve, as shown in image (B) and (C). Upon cooling, the fibers reformed on the HOPG surface, image (D).
This process is dramatically different from the aggregation of pure EP fibers. EP self-assembly occurs at elevated temperature because the water shell protecting the hydrophobic domain is destroyed, allowing the molecules to aggregate through hydrophobic interactions. The mixture of EP with HG behaves in an opposite way, forming ordered structures at room temperature and dissolving at high temperatures. We believe this is caused by the addition of detergents, which cover the hydrophobic surface of EP. The interaction between EPs is thus dominated by hydrogen bonding among the hydrophilic glucose head groups of the detergent molecules. Due to their amphiphilic character, detergents connect EP with hydrophobic HOPG via its ordered alkyl chain array, so the EPs are still able to form ordered fibers on HOPG surface. (Figure 3-3(B))

![Figure 3-3. Illustration of the self assembly of EP and its complex with detergent on HOPG at various temperatures.](image)
(A) The adsorption of pure EP molecule on HOPG surface via hydrophobic interaction. White part is the hydrophobic domain; red part is the hydrophilic domain. (B) The assembly of EP molecules with detergents on HOPG surface through the hydrophobic tail of the detergent molecules. The dot line indicates the detergents embedded between the EP and HOPG.

This experiment proves that the approach of using detergents to improve the self-assembly of EP is feasible. Using the same composition of EP and HG, a particle of crystalline appearance has been observed together with hexagonal arrays of fibers on graphite surface at 10 °C. (Figure 3-4)
Figure 3-4. Deflection image (A) and height image (B) of EP II and HG on graphite surface at 10 °C. Besides the large particle with angular edge, the ordered ~ 170 nm long fibers can be clearly seen on the surface of graphite from the deflection image. Image size: 2 × 2 µm.

It is interesting to find that the detergents did not interrupt the epitaxial registration of EP with the underlying graphite surface, even after forming the EP-detergent complex. It is known that detergents of different structures have different properties, so their affinity with proteins must be different. We are interested in testing the association of EP with different detergents to see if they are able to form ordered three-dimensional structure.

The detergents we chose are n-Hexyl-β-D-Glucopyranoside (HG), n-Octyl-β-D-Glucopyranoside (OG), n-Decyl-β-D-Maltopyranoside (DM) and N-Decyl-N,N-Dimethyl-3-Ammonio-1-Propanesulfonate (Anzergent® 3-10). The reason for selecting these molecules is that OG, HG and DM are nonionic detergents. They are mild protein crystallization reagents and do not alter a protein’s native structure. HG and OG share the same head group, but have different alkyl chain lengths; DM has a double-sugar head group. Anzergent® 3-10 is used as a comparison to the above three.
Table 3-1. Chemical structure and properties of detergents used for screening the crystallization of EP.

<table>
<thead>
<tr>
<th></th>
<th>Name</th>
<th>Property</th>
<th>Structure</th>
<th>Alkyl Length</th>
<th>Estimated Length</th>
<th>CMC (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n-Hexyl-ß-D-Glucopyranoside (HG)</td>
<td>Nonionic</td>
<td></td>
<td>6</td>
<td>10.88 Å</td>
<td>250</td>
</tr>
<tr>
<td>2</td>
<td>n-Octyl-ß-D-Glucopyranoside (OG)</td>
<td>Nonionic</td>
<td></td>
<td>8</td>
<td>13.40 Å</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>n-Decyl-ß-D-Maltopyranoside (DM)</td>
<td>Nonionic</td>
<td></td>
<td>10</td>
<td>15.91 Å</td>
<td>1.8</td>
</tr>
<tr>
<td>4</td>
<td>N-Decyl-N,N-Dimethyl-3-Ammonio-1-Propanesulfonate (Anzergent®3-10)</td>
<td>Zwitter-ionic</td>
<td></td>
<td>10</td>
<td>17.78 Å</td>
<td>40</td>
</tr>
</tbody>
</table>

First we studied the adsorption and assembly of detergents on HOPG from room temperature to 4 °C. We found that they all readily formed linear hemi-cylindrical structures similar to that reported for other amphiphiles. All detergent samples were prepared at their critical micelle concentration (CMC). 200 µl solution was injected into the contact/tapping liquid cell on HOPG. After half-an-hour incubation at 4 °C in the AFM fluid cell, images were acquired by tapping mode AFM and analyzed by two-dimensional FFT image analysis. As mentioned previously, two-dimensional FFT analysis is an image processing tool able to characterize periodic structures embedded within an image. After decomposition, the DC value (the image mean) is displayed as the center of the frequency domain. We used this method to find the periodicity of the ordered structures and estimate the spacing of the hemi-cylindrical micelles. The results are summarized in the following Figure 3-5.
Figure 3-5. AFM images of ordered structure of HG, OG, DM and Anzergent adsorbed on HOPG at 4 °C. All these detergents form linear hemi-cylindrical structures parallel to each other. Inserted: Periodicity revealed by Fast Fourier transform analysis, showing the spacing of the ordered structures. Image size 250 × 250 nm.

We can see the periodicity of the radius, as determined by the fast Fourier transform, increases from HG to OG to DM to Anzergent® 3-10, representing the increasing spacing of
the adsorbed epitaxial array. This is in good agreement with the increasing alkyl chain length in each detergent. Some questions however still remain. For example,

* Detergents with shorter chains, like HG which has only has 6 C atoms, can also form the hemicylindrical structure.

* The width of the rods is four times larger than a single monomer and is two times longer than the hemicylinder model.

* The height of the different rods is all about the same, ~0.15 nm. Compared with the hemicylinder model, the height of the rod should be about the length of a monomer.

* The FFT spectrum suggests that there is also two-dimensional order.

Close inspection reveals some interesting details of the detergent array on HOPG. (Figure 3-6) The detergent molecules are able to form separate domains oriented at either 60° or 120° to each other. Within each domain, the arrays are aligned parallel. At the edge of the domain, the arrays belonging to different domains are clearly separated, which means the surface arrangement of detergent molecules on graphite is more affected by the underlying graphite surface.
Figure 3-6. Self assembly of 1 CMC Zwitterionic ® 3-10 on HOPG at 4 °C. In image (A), three separate domains set side by side on graphite surface with different orientation and clear boundary, as indicated by the red line. (B) The close-up image of the distinctive end of each single array of detergents within different domains. The intersection angle between domains is ∼ 120°. The alignment of detergent molecules is indicated as perpendicular to the long axis of the rods.

Compared with the adsorption of long-chain hydrocarbons on HOPG revealed by STM, we believe the reason for the hexagonal arrangement of detergent molecules is epitaxial registration of the detergent’s methyl groups into the hexagonal hollows on the HOPG surface. The surfactant molecules adsorb initially as a monolayer in registry with the graphite surface in a tail to tail configuration. Subsequent packing of additional surfactant molecules results in the formation of a hemi-cylindrical structure through hydrophobic interactions with the existing surfactant monolayer.

3.2 Complex of EP Peptides with Detergents on HOPG at Lower Temperature

Observation of fiber formation from the EP II - HG complex at room temperature on HOPG, let us to investigate the possibility of forming similar structures of EP-detergent complexes. This was achieved through a series of combinations of 25 µM EP I and EP II with
1 CMC HG, OG, DM and Anzergent® 3-10 to find the optimum conditions for fiber formation.

It is known that too little detergent is not enough to cover the hydrophobic surface of proteins, and too much detergent could cause repulsive force between proteins. The CMC is the concentration at which a detergent is able to form micelles. Each detergent has their own CMC values, which is a consequence of their distinct molecular structures and capabilities for self-association. It is generally accepted that in membrane protein crystallization trials, the starting concentration of detergent is selected at 1 CMC or a little higher. We selected 1 CMC as the concentration of these four detergents and studied their association with EP. The molar concentration corresponding to the CMC of HG, OG, DM and Anzergent® 3-10 are different, so that the molar ratio of protein to detergent is different, 1:10000 for EP to HG, 1:800 for EP to OG, 1:72 for EP to DM and 1:1600 for EP to Anzergent® 3-10. Although the ratio of EP to DM is ~ 2 fold lower than the one of EP to HG, comparing with the minimum requirement on the ratio of protein to detergent, which is 1:10, it is expected there will be enough detergents to cover the surface of EP.

The mixing procedure was as follows: a 25 µM EP and 1 CMC detergent solution were mixed together and sealed in contact/tapping liquid cell, and in situ tapping mode AFM imaging was conducted over a temperature range running from room temperature to 4 °C. (Figure 3-7 and 3-8)
Figure 3-7. Images (1 × 1 µm) of the structures formed from the mixture of 25 µM EP I with 1 CMC detergent on HOPG surface at 4 °C. While EP I - HG complex and EP I - DM complex only formed an amorphous layer on the HOPG surface, the EP I - OG complex and EP I - Zwitterionic complex formed linear arrays which are similar to pure detergent, as shown in the inserted image (250 × 250 nm).
Figure 3-8. Images (1 × 1 µm) of the structures formed from the mixture of 25 µM EP II with 1 CMC detergent on HOPG surface at 4 °C. Besides the EP II - HG fibers seen previously, only EP II - Anzergent® 3-10 complex formed fiber-like structure. EP II - OG complex formed linear arrays similar to pure detergent (inserted image: 250 × 250 nm) and EP II - DM complex formed an amorphous layer.
The number of ordered fibers formed from the mixture of EP and detergent was small. This was not very surprising as it is known that a structural match between protein and detergent is necessary for fibre formation.

3.3 Substrate Facilitated Self-Assembly of EP and Its Complex with Detergent

The next step was to test whether the detergents could facilitate the formation of EP peptide fibers. The well-defined two-dimensional arrays of the detergents were first formed on HOPG and the EP peptides were introduced on top of these detergent-modified HOPG surfaces at 4 °C. In situ AFM was used to image the adsorption of EP peptides on top of these templates. (Figure 3-9)

![Figure 3-9](image_url)

Figure 3-9. The morphology of 25 µM EP I on a pre-formed OG monolayer at 4 °C after a 2-hour incubation. The linear alignment of OG can be clearly identified from image (A), as indicated by the red arrow. (B) EP I form globular structure on the pre-layer instead of those fibers in the hexagonal pattern. (Image size: 1 × 1 µm)

Our initial studies were inconclusive as the EP I peptide failed to form fibrils on the detergent-modified HOPG surface. Only large amorphous EP peptides were observed, even
after heating in solution, conditions that had previously been shown to foster EP peptide coacervation. (Figure 3-10)

We also tried to study whether it will be better to use a mixture of EP I and OG instead of the pure EP I. Results show that this complex only forms amorphous particles on the OG pre-layer. (Figure 3-11)
Figure 3-11. An amorphous particle formed from the mixture of 25 µM EP I and 1 CMC OG on a pre-layer of OG on top of the HOPG.

The reason for the difficulty in forming an ordered EP I layer on the OG pre-layer could be that the hemi-cylindrical structure of the OG micelle is not atomically flat. However, we believe that this effect is more a consequence of the hydrophilic surface of the hemi-cylindrical OG pre-layer and its incompatibility with the hydrophobic EP. (Figure 3-12) This effect is similar to that proposed for the formation of the amorphous EP aggregates on mica, as we have previously reported.
Figure 3-12. Illustration of the effect of surface property on the adsorption of EP peptides. EPs are able to form fibers in an ordered pattern on HOPG due to hydrophobic interaction. When the surface is covered by the detergents, similar structures cannot form due to the loss of proper interaction.

An alternative approach involved using the ordered array of EP as a template to study its interaction with OG. In the next experiment, we formed an ordered layer of EP I fibers on HOPG at high temperature (45°C), injected 200 µl of a 1 CMC OG solution onto the EP I pre-layer and allowed it to cool to room temperature to study the interaction between EP I and OG and the structure change of the EP I fibers. (Figure 3-13)
Figure 3-13. Structure change of EP I fibers formed on HOPG before and after the injection of OG. (A) 25 μM EP I self-assembled into normal fibers in a hexagonal pattern on HOPG at 45 °C, alone with some big aggregates. After the injection of 1 CMC OG, image (B) shows the big aggregates and part of the fibers were washed away. (C) When the solution was cooled down to room temperature, the remaining fibers stayed on the surface but no specific structure was formed.
It is clear from the result shown in Figure 3-13 that the detergent simply dissolved the existing fibers after injection. Compared with inhibited aggregation of EP in solution observed by dynamic light scattering, we believed in this case, the detergent breaks up the hydrophobic interaction between the EP peptide and HOPG.

From the structure of EP, we know it is soluble at room temperature because water forms a protective shell around the peptide, preventing hydrophobic interactions between EP molecules. If the detergent is mixed at this stage, it needs to compete with the stable existing water structure, which could be kinetically slow or unfavourable from a free energy perspective. The question becomes how to attach the detergent onto the hydrophobic surface of EP successfully.

We proposed that mixing the EP peptide with the detergent at elevated temperatures may facilitate formation of an EP-detergent complex. Because under those conditions, the protective water shell would be disrupted enabling the detergent to interact with EP. Given that EP will self-aggregate at elevated temperatures, we used a modified approach that involved injecting EP into a heated detergent solution. In this way, we hypothesized that the detergent molecules would compete successfully for the exposed hydrophobic surface of the individual EP molecules.

The next experiment we did involved injecting 200 µl of a 5 µM EP I solution into 1 CMC OG pre-heated at 40 °C and cooled down the solution to 4 °C. (Figure 3-14)
From the AFM images, we can see there are two types of fibers formed on HOPG simultaneously. All of them are arranged in an hexagonal pattern on the HOPG surface, but there are clear differences between the fiber types. One type of fiber is longer, thinner and arranged in a pattern similar to the fibers formed by EP I in the absence of detergent. Another type of fiber is shorter, thicker and arranged in parallel bundles. From Figure 3-14(A), in a randomly selected 1 x 1 µm area, there are 12 long fibers with an average length of 308 ± 51 nm. The second type of fiber is about 1 nm thicker than the first one (as shown from the section analysis). Within the same 1 µm² region, there is a higher density of these fibers, ~ 128 shorter fibers with an average length of 68 ± 2 nm. When the solution is cooled down to room temperature, more and more thick fibers appeared on the surface.

Figure 3-14. 5 µM EP I formed fibers on HOPG surface after injected into 1 CMC OG at 40 °C (A). The section line shows the difference in height of two types of fibers, the thick fibers (as indicted by the arrow) are ~ 1 nm taller than the thinner fibers. (B) After ~ 1 hour, when the solution was cooled down to room temperature, more and more thick fibers appeared on the surface.
temperature after 1 hour, in another $1 \times 1 \, \mu\text{m}$ area from Figure 3-14(B), there are 10 long fibers; their average length is $285 \pm 36 \, \text{nm}$. Compared with 187 shorter fibers having an average length of $82 \pm 2 \, \text{nm}$, it is clear that the number/density of the shorter fiber increased. This suggests that there are two different types of fibre species: the uncomplexed EP I and an OG-EP I complex. The shortest fiber observed was $\sim 34 \, \text{nm}$ in length, which may correspond to the nucleus for fiber growth.

When this sample was cooled down to 4 °C, we found a two-dimensional ordered structure besides the fibers (Figure 3-15). This crystal kept on growing during a 1-hour period. Figure 3-16 shows that it increased in area from $278 \, \text{nm}^2$ to $867 \, \text{nm}^2$. However, in the first 26 minutes its growth rate was $17.2 \, \text{nm}^2/\text{min}$, in the next 22 minutes it decreased to $5.8 \, \text{nm}^2/\text{min}$, which could be due to the consumption of available molecules from solution. A close-up view reveals a multilayer structure comprised of uniform $\sim 4.6 \, \text{nm}$ steps.
Figure 3-15. At 4 °C, a two-dimensional ordered structure was found from the mixture of 25 μM EP I and 1 CMC OG (A) and it continued to expand during a 1-hour period (B). Image (C) is the close-up view, showing the typical crystalline structure. The section diagram of the designated area in image (C) shows the step height is ~ 4.6 nm.
Figure 3-16. Crystal size of EP I with OG increase during a 1-hour period.

Because EP I is soluble at low temperature and OG forms hemi-cylindrical linear arrays on HOPG, we believe this combination does form an OG - EP I complex. Compared with the usual hexagonally arranged linear fibers formed by EP I alone, it is clear that the molecules are arranged in an ordered three-dimensional array rather than simply in two dimensions.

3.4 Particle Size Change Measured by DLS.

DLS has been widely applied to study the early stages of molecular assembly and association in solution. For example, McPherson et al. used DLS to characterize the crystallization of canavalin at different protein concentrations, salt concentrations and pH. Their results revealed the formation of a ~ 75 nm trimer, which was considered as the critical species for crystallization. Besides the application of DLS in protein crystallization, this technique has also been used for the drug discovery process. For example, in high-throughput screening studies, often many hits are identified that do not have the correct structure-activity relationship. They can inhibit several unrelated enzymes, but lose function when the concentration of model enzymes is 10 times higher, despite of a 1000-fold excess of
inhibitor. By using DLS and SEM, Brian Shoich et al found the active form of these promiscuous inhibitors was an aggregation of many individual molecules. They will lose their enzyme-inhibiting activity if aggregation is inhibited.

DLS has also been used for studying molecular association, which is important for physiological function. T-cell receptor multimerization is the initial step of the T-cell activation. Davies et al revealed that the T-cell receptor interacts with an antigenic peptide bound to major histocompatibility complex (MHC) to form supramolecular structure at the concentration near the dissociation constant of the binding reaction. The size of the oligomers is concentration-dependent and contains two to six ternary complexes. This research not only provided direct evidence for the specific signalling of T-cells, but also demonstrated the DLS is a very sensitive tool in studying molecular association since all the particle size changes were below 10 nm.

To compare the self-assembly of EP and detergent on HOPG imaged by AFM, we also measured the particle size of the EP-detergent complex as a function of temperature by dynamic light scattering. The procedure involved tracking changes in particle size while the sample from 25 °C to 50 °C, followed by cooling to 5 °C. First we studied the interaction of OG with EP I and EP II. (Figure 3-17)
Figure 3-17. Aggregate size of EP I and EP II with 1 CMC OG as a function of temperature. The orange curve indicates the increased size of EP I and OG complex upon heating from 25 °C to 50 °C. The blue curve represents the size change of the same composition upon cooling to 5 °C. Another group of size change curves were shown for the combination of EP II with OG, the red line is the size heating curve and the deep blue line is the cooling curve.

From the result, we can see the mixture of 25 µM EP I with 1 CMC OG does not aggregate until the temperature is heated to 50 °C, at which point the aggregate size increases from several nanometers to about 1433 nm. The mixture of EP II with OG starts to aggregate at 35 °C, and grew in size from 3264 nm to 4521 nm at 50 °C. We find the high temperature aggregate has a smaller polydispersity and narrower size distribution, compared to the normal aggregates formed by EP alone. (Figure 3-18)
Figure 3-18. Particle size measurement of 25 µM EP I and 1 CMC OG at 50 °C obtained by dynamic light scattering.

Similar result has been obtained from the mixture of 25 µM EP with 1 CMC DM under the same experimental condition. (Figure 3-19)
Particle size of DM with EP I and EP II

EPI + DM

EPII + DM

0

500

1000

1500

2000

2500

3000

3500

4000

4500

5000

0 102 0 30 405 0 60

Temperature

Particle size (nm)

EPI+DM_heat

EPI+DM_cool

EPII+DM_heat

EPII+DM_cool

Figure 3-19. Aggregate size of EP I and EP II with 1 CMC DM as a function of temperature. The red line and orange line are the heating curves of EP II and EP I with DM, 25 °C to 50 °C. The deep blue line and the blue lines are the cooling curves, from 50 °C to 5 °C.

The mixture of 25 µM EP I and 1 CMC DM aggregates at 50 °C. The size of the aggregates is ~ 1808 nm. The mixture of 25 µM EP II and 1 CMC DM aggregates at around 35 °C showing a size of 3707 nm, and increases to 4286 nm at 50 °C.

Comparing these two experiments we find the aggregation temperature of EP with OG and EP with DM to be same; 50 °C for EP I and 35 °C for EP II. The sizes for the aggregates are also similar. There is a 23% difference for the EP I and only 5% difference for the EP II. We believe this means that the numbers of molecules within the aggregates is close and their structures are similar.

The same experiment has also been carried on the mixture of 25 µM EP with 1 CMC zwitterionic detergent Anzergent® 3-10. (Figure 3-20) However, this process looks very
different. There is no consistent trend in particle size upon heating and the aggregates never exceeded 2000 nm in size. This suggests the aggregation of EP has been mainly inhibited by Anzergent® 3-10.

The aggregate size is 151 nm for EP I with Anzergent® 3-10 at 50 °C, and 941 nm for EP II with Anzergent® 3-10. There is a sudden particle size change at 25 °C for the mixture of Anzergent® 3-10 with EP II and 5 °C for EP I, but the mechanism is unknown. It may be due to a change in the interaction between Anzergent® 3-10 and EP. Similar processes were recorded from the mixture of EP I and EP II with HG. (Figure 3-21)
The particle size of HG with EP II is always below 10 nm. The particle size of EP I with HG at 50 °C is 156 nm. It is interesting to note that, at 5 °C the solution of EP I with HG is cloudy and the particle size increases to ~ 1131 nm with a narrow size distribution. Crystallization trials of this combination at the same condition are under investigation.

The reason for the different behaviour of HG and Anzergent®3-10 with EP, as compared to OG and DM, probably can be attributed to the different structure and interactions between the detergents and protein. For example, in the crystallization of cytochrome c oxidase using maltoside with different alkyl chain length, C12-maltoside enabled structure determination to 8 Å while a C10-maltoside did not result in the formation of any crystals, and C11-maltoside allowed for a structure resolution to 2.6 Å. This result clearly

---

**Figure 3-21.** Aggregate size of EP I and EP II with 1 CMC HG as a function of temperature. The orange line is from the mixture of EP II with HG, from 10 °C to 50 °C. The red and blue lines are the heating and cooling curves of EP I with HG.
indicates that detergent selection is critical for crystallizing proteins. A small difference in detergent structure could cause significant differences in the crystallization process.

3.5 Conclusion

Detergents readily form hemi-cylindrical structures aligned on the (0001) surface of HOPG. It is believed that epitaxial registration directs the parallel alignment of each detergent molecule. The resulting structure can be used to convert the hydrophobic HOPG surface to a hydrophilic surface. The two-dimensional crystalline structure formed by mixing hot OG with EP supports the hypothesis that detergent can improve the assembly of hydrophobic EP molecules. The 1-nm-thicker fibers, rather than the fibers of normal thickness, formed from the optimized procedure suggest that the detergent has formed an EP-detergent complex.
4 Molecular Structure Investigation of EP

4.1 FT-IR Spectrum of EP Secondary Structure

The ordered self-assembly of EP on HOPG is determined by molecule-molecule and molecule-substrate interactions. To understand and control this process, it is necessary to reveal the mechanisms behind these interactions. However, the molecular structure of EP is unknown, because of the difficulties in crystallizing this protein. Due to their long hydrophobic segments, EP molecules tend to aggregate in solution at elevated temperature, instead of forming ordered crystalline structures. An alternative approach to study the molecular structure of protein is using nuclear magnetic resonance (NMR). NMR is able to determine the protein structures in solution and in the solid state. However, its current application is limited to small proteins (less than 30 kDa), because the chemical shift with structure and conformation is different for large proteins than small organic molecules, and the interpretation of a large protein’s NMR spectrum is very complicated. Circular dichroism (CD) has been used to detect the secondary structure of proteins. CD measures the optical activity of a protein in the range between 190 nm and 230 nm, which is dominated by the backbone conformation. At elevated temperatures, EP forms amorphous aggregates so that the solution become cloudy, which limits the application of CD.

There are other approaches that can be used to detect protein structures. FT-IR is a popular technique in determining the secondary structure of proteins. IR measures the absorption energy of vibrating chemical bonds, from either the stretching motion or the bending motion.

FT-IR is different from traditional dispersive IR, which uses a diffraction grating to separate the wavelengths of light in the spectral range and direct it to the sample. FT-IR uses an interferometer to collect the spectrum. The light from the source is split by the beam
splitter into two parts. One part is transmitted to a moving mirror; another part is reflected to a fixed mirror. The moving mirror moves back and forth at a constant velocity, according to the precise laser wavelength in the system. The two FT-IR beams reflected from the mirrors are re-combined at the beam splitter. Because they have traveled a different distance, an interference pattern is created since some wavelength combine constructively and some destructively. This interference is called an interferogram. This interferogram goes through the sample and reaches the detector. The detector senses the signal and sends it to the computer. The computer uses Fourier transform to convert it back to normal single beam spectrum. (Figure 4-1)

![General diagram of interferometer used for FT-IR.](image)

FT-IR reads information about the sample’s absorbance at every wavelength in the infrared range simultaneously, so that it is much faster than traditional IR. The signal to noise
ratio is much higher for FT-IR than dispersive IR because there is less reflection loss caused by
the mirror surface. Virtually all modern infrared spectrometers use the FT design instead of the
dispersive method. Due to its high sensitivity, FT-IR has been used to determine the
secondary structure of proteins.

Typically there are nine bands associated with the amide bond of proteins. They are
referred as A, B, and I – VII in order of decreasing frequency. The amide I band (80% C=O
stretch, near 1650 cm$^{-1}$), amide II band (60% N – H bend and 40% C – N stretch, near 1550
cm$^{-1}$), and amide III band (40% C – N stretch, 30% N – H bend, near 1300 cm$^{-1}$) are used to
identify protein secondary structure. From proteins and peptides of known structures and
based on theoretical calculations, specific frequency regions have been assigned to particular
secondary structural elements and types. For antiparallel $\beta$-sheet structure, there are two amide
I absorption, one is from 1615 to 1640 cm$^{-1}$; another weak band is at 1670 - 1690 cm$^{-1}$. The
amide I band of parallel $\beta$- sheet is located from 1626 -1640 cm$^{-1}$. Antiparallel $\beta$- sheet also has
a strong amide II band at 1510 - 1530 cm$^{-1}$; an amide II band of parallel $\beta$ - sheet is found at
1530 - 1550 cm$^{-1}$. For $\alpha$ - helix, there is a strong amide I band at 1650 - 1655 cm$^{-1}$, and a peak
around 1548 cm$^{-1}$ for the amide II band. (Table 4-1)
Amide I Band (cm$^{-1}$) | Amide II Band (cm$^{-1}$)  
--- | ---  
α-Helices | 1650 - 1655 | ~ 1548  
Antiparallel β-sheets | 1615 - 1640 | 1510 - 1530  
| | 1670 - 1690 |  
Parallel β-sheets | 1626 - 1640 | 1530 - 1550  

Table 4-1. Typical peak frequency in FT-IR spectrum of amide I and II band from the secondary structure of proteins.

Subtle changes in IR spectra provide useful information about the secondary structure of the protein. For example, if there is a hydrogen bond forming between carbonyl and amine group, which is parallel to the C=O stretching and N-H bending, the amide I band could move to a lower energy / wavenumbers, because the hydrogen bond makes it easier to stretch the carbonyl oxygen toward the hydrogen. Meanwhile, the amide II band could move to a higher energy because the hydrogen of N-H is fixed in place due to the hydrogen bond.  

The secondary structure of tropoleasin-like peptides has been shown to contain large amounts of β-strands. From the ordered self-assembly of EP on HOPG, we believe there could be a conformational change in EP after absorption onto the HOPG surface. To obtain direct evidence of the secondary structure change of EP after absorption, we decided to use FT-IR to compare the frequency shift of chemical bonds within these peptide samples in monomer status and after adsorption onto HOPG. The FT-IR spectra of monomeric, non-aggregated EP were collected in transmission mode on a Nexus 670 spectrometer (ThermoNicolet, Madison, WI). (Figure 4-2)
Figure 4-2. Infrared transmission-absorbance spectra of EP in KBr pellet in air collected at room temperature. The absorption peaks at ~ 1630 cm\(^{-1}\) and 1550 cm\(^{-1}\) belong to amide I (*) and amide II band (**).

Not unexpectedly, the spectra of pure EP I, II and IV are almost identical. The amide I band is at ~ 1630 cm\(^{-1}\), the amide II band is at 1551 cm\(^{-1}\). Compared with the normal wavelength position of protein secondary structure, EP likely has a \(\beta\)-strand. Details in the spectra change after the formation of thin film provide us interesting information about the structural change after deposition.

The spectra of the thin film of EP adsorbed on HOPG were collected using a Continuum FT-IR microscope (ThermoNicolet, Madison, WI). The EP sample solution was placed on the HOPG surface and a thin film was allowed to form at 40°C. The spectra were collected from the average of 128 scans at a 4 cm\(^{-1}\) resolution, on the sample surface between the buffer salt particles. (Figure 4-3 and Figure 4-4) in air using reflectance mode.
Figure 4-3. Image shows the location on the EP I thin film where the IR spectra were collected.
Figure 4-4. Infrared spectra of EP thin films, from 2000 to 1000 cm\(^{-1}\), adsorbed on HOPG, collected on Nicolet\textsuperscript{TM} Continuum\textsuperscript{TM} infrared microscope, using reflectance mode. Amide I (*) and amide II (**) bands are indicated on graph.

It is found that after deposition, the amide I bands from EP I, EP II and EP IV film formed on HOPG remained largely the same. However, the amide II band shifts to the higher energy level / larger wavenumbers. (Table 4-2) It is clear that there is some secondary structural change for EP molecules after absorption onto the HOPG surface.
Given the general rule of spectral shifts in the amide bond position, these results indicate that there is hydrogen bond formation between N-H and C=O after the deposition of EP, so that the bending of N-H becomes more difficult and wavenumbers increases by 4 cm\(^{-1}\).\(^{101,102}\) This result is also consistent with the hypothesis that the aggregation of EP arises from the hydrophobic interaction at elevated temperature. We believe after deposition, hydrogen bonds form between N-H bending and C=O stretching, altering the N-H vibrational mode.

### 4.2 Unfolding of EP by Single Molecule Force Spectroscopy

Single molecule force spectroscopy by AFM has been used to measure the intramolecular interaction force of biomolecules with the goal of understanding their structure and function. For example, the mechanism involved in the elastic deformation of polysaccharides is poorly understood. The force-extension curve of a single dextran molecule between AFM tip and substrate revealed a force plateau at 250 ± 30 pN, and the slope increases after this plateau, which indicates that the molecule switches to a stiffer conformation to resist the applied tension. From the normalized force versus elongated polymer length data, the segment elasticity is 670 pN/Å in the low force regime and 1700 pN/Å in the high force regime.\(^{105}\) From \textit{ab initio} calculations, this observed entropic elasticity was attributed to the chair-boat conformation change of the glucopyranose ring.\(^{106}\)

<table>
<thead>
<tr>
<th></th>
<th>Amide I (cm(^{-1}))</th>
<th>Amide II (cm(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EP I</strong></td>
<td>powder</td>
<td>1630.35</td>
</tr>
<tr>
<td></td>
<td>HOPG</td>
<td>1630.09</td>
</tr>
<tr>
<td><strong>EP II</strong></td>
<td>powder</td>
<td>1630.39</td>
</tr>
<tr>
<td></td>
<td>HOPG</td>
<td>1631.25</td>
</tr>
<tr>
<td><strong>EP IV</strong></td>
<td>powder</td>
<td>1630.66</td>
</tr>
<tr>
<td></td>
<td>HOPG</td>
<td>1632.10</td>
</tr>
</tbody>
</table>

Table 4-2. Wavenumbers of amide I and amide II Bands in FT-IR Spectrum of EP I, II and IV in KBr pellet and on HOPG surface.
Recent studies have examined the force-extension characteristics of DNA, which is a remarkably flexible molecule that can be extended from a twisted state (B-form) 1.7 times longer to a stretched state (S-form). By single-molecule AFM force spectroscopy, a 1.5 μm λ-DNA molecule undergoes the B-S transition under the application of 65 pN of force. A second force threshold at 150 pN appears to induce a nonequilibrium transition, which has been attributed to a split in the DNA double strand. This is similar to the melting transition of DNA. Using single poly(dG-dC) and poly(dA-dT) DNA strands, unzipping of these hairpins suggests that the base pair-unbinding forces for G-C and A-T are 20 ± 3 pN and 9 ± 3 pN respectively. These molecular interactions are very important in understanding DNA replication and transcription.

AFM force spectroscopy is also able to manipulate the structure of proteins. A good example was a recent study of bacteriorhodopsin. By applying forces of 100-200 pN, researchers were able to use the AFM tip to unfold the seven closely-packed α-helices within the membrane-associated molecule. Comparing the native protein and cleaved protein (at E-F loop), it has been suggested that in bacteriorhodopsin, helices G and F as well as E and D unfold in a pairwise fashion. Native proteins have complicated three-dimensional structures, so their force spectra can be quite complex. To solve this problem, recombinant proteins have been used to study the structure of specific domains. (Figure 4-5) Recombinant proteins derived from titin containing repeating Ig4 and Ig8 domains were studied under force extension conditions. It is found that there are always four or less equally spaced peaks in case of Ig4 and no more than eight peaks for Ig8. All force curves exhibit a characteristic sawtooth pattern with a 25-nm periodicity at 150-300 pN, which is very similar to native titin. Similarly, single molecule force spectroscopy has been performed on the extracellular matrix protein tenascin, which contains repeating fibronectin Type III domains.
Figure 4-5. Well-defined recombinant titin segments produces force extension curves from the unfolding of individual Ig domains. (A) is from the stretching of Ig8, (C) is from the unfolding of Ig4, the superimposed force curves in (B) and (D) reveals the 25 nm pattern, (F) demonstrates the unfolding of individual Ig domains. (Adapted from reference 109)

However, the force measured from breaking a pair of bond of the same type could be different, depending on the orientation of the molecule relative to the direction of force applied. For example, for the same type of hydrogen bonds between β-strands, if their direction is perpendicular to the applied force, such as the I27 domain from human cardiac titin, the measured force is 204 ± 26 pN, which is attributed from the simultaneous rupture of a cluster of hydrogen bonds from the “shear topology”. While the orientation of hydrogen bonds is parallel to the direction of the applied force, such as C2A domain from synaptotagmin, the measured force is only ~ 60 pN because each bond breaks sequentially in a “zipper” configuration. (Figure 4-6)
Since it is feasible to target the AFM tip precisely onto the sample surface, the ordered hexagonal array of EP peptide molecules provides a good template to pull out a single molecule or fiber and study its unfolding process. We applied this approach to EP peptides to reveal their unfolding pathway and possibly details of its secondary structure.

For single molecule force spectroscopy of EP peptides, triangular-shaped cantilever with 200 µm extension was employed for both imaging and force measurement. Its apparent spring constant is 0.12 N/m, as measured by resonance-frequency-shift method. This cantilever is capable of detecting forces in the pN range when the vertical resolution of AFM is within Å range. To demonstrate the single force curve obtained is from the unfolding of a single protein, we targeted the AFM tip on top of a single fiber, carried out the force curve twice and compared the images before and after the single force spectroscopy. (Figure 4-7)

**Figure 4-6.** Schematic diagram for different patterns in breaking hydrogen bond between β-strands. (A) The “zipper” model of H-bond breakage, in which each bond breaks sequentially. (B) The “shear pattern” model of H-bond breakage, in which the orientation of H-bond is perpendicular to the applied force.
Figure 4-7. Single molecule force spectroscopy on a single fiber formed from EP I. (A) the entire fiber before the force spectroscopy. (B) The dissected fiber after the force spectroscopy. Insert: two force curves recorded between these images, in which the sawtooth pattern can be clearly identified.

Comparing these two images, it clearly shows that before the single molecule force spectroscopy experiment, there is an intact fiber from EP I formed on top of ordered first layer. After two force curves were executed exactly on top of this fiber, in which a sawtooth pattern was clearly identified, this fiber has been dissected into three segments. This experiment demonstrates the sawtooth pattern within the force curve is generated from the unfolding of a single EP molecule. We applied this approach to investigate the unfolding pathway of EP I and EP II molecules.
The EP I and II molecules were pulled at a rate of 250 nm/s. More than 100 force curves were collected, and each single unfolding peak within the curve was picked manually. It is interesting to find that, by comparing the local maximum, there are well-organized specific peaks having similar unfolding length from unfolding pathway, which are in agreement with the repetitive amino acid sequences of EP I and EP II. (Figure 4-8)

For EP I, there are three peaks exhibiting certain unfolding lengths were more frequent and are distinctive from other peaks, as shown by the arrows in Figure 4-8. Peak 1 is at the length of ~ 2.6 nm, peak 2 is at the length of ~ 5.0 nm and peak 3 is at the length of ~ 8.0 nm. If the length of a single amino acid is ~ 0.36 nm, the number of amino acids contained within peak 1, 2, 3 are correspondingly 7, 14 and 22. It is interesting to notice that they are all multiples of the length of seven amino acids. Compared with the repetitive segment

Figure 4-8. Histogram distribution of peaks found in the single molecule force spectroscopy of EP I pulled at 250 nm/s, arrows indicate the characteristic unfolding length as identified from the local maximum.
-PGVGVA- within known sequence of EP I, we believe these peaks represents the extension of these folded segments.

By using Fast Fourier Transform (FFT) analysis of unfolding peaks, the characteristic unfolding length for EP I is found at 2.59 nm, corresponding to the ~7 amino acids, which is the same as the results found by using the local maximum method. There is another peak found at 3.29 nm, which is believed coming from the unfolding of longer segments, e.g. the multiples of 7 amino acids, and the irregularity within the unfolding pathway. (Figure 4-9)

Figure 4-9. Fast Fourier transform of the frequency of the unfolding length of EP I reveals the well-defined spacing of 2.59 nm, which is believed from the repeat -PGVGVA- segment.
It is also worth noting the longest peak found within EP I is ~16.6 nm, corresponding to the length of 46 amino acids. From the total 118 amino acids within EP I, the estimated fully extended should be ~ 42.5 nm. The difference between these two length scales indicates that these force-extension curves likely are originated from the unfolding of single EP molecules.

For EP II, there are six distinctive peaks found in the unfolding pathway, as indicated by the arrows in the Figure 4-10. They are at length of 3.1 nm, 4.9 nm, 7.6 nm, 10.0 nm, 14.6 nm and 20.2 nm, corresponding to 9, 14, 21, 28, 41 and 56 amino acids. It is interesting to find that once again these peaks represent the repetitive segments of seven amino acids, -PGVGVA-. Compared with the longest unfolding length of EP I, it is 27.0 nm for EP II, representing 75 amino acids. From the total 200 amino acids within EP II, the estimated
extension of the entire molecule is \(~72\) nm, which is longer than the detected extension of EP II molecule.

Figure 4-11. Fast Fourier transform of the frequency of the unfolding length of the EP II reveals the well-defined spacing of 2.46 nm, which is in match with the extended length of -PGVGVA- segment.

Compared with local maximum approach, the Fast Fourier transformation also revealed a similar unfolding length for EP II, as it did for EP I. The first characteristic peak occurs at 2.46 nm, which corresponds to \(\sim 7\) amino acids. These single molecule force spectroscopy we obtained provide very useful information about the local secondary structure and the folding mechanism of EP.
4.3 Molecular Dynamics Simulation EP Using CAChe

We also used molecular dynamics simulations to predict the structure and conformational changes in EP to explain the self-assembly of EP on HOPG at elevated temperatures. The program we used was CAChe (CAChe Group, Fujitsu America, Inc., Beaverton, OR). It was designed to estimate the reactivity of chemical bond formation from calculating the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of molecules. Since version 5.0, it has also been improved to estimate the free energy change of a large amount of atoms using molecular dynamics simulation, such as ligand docking.

CAChe is not a mainstream tool for molecular dynamics simulation. However, it is still capable of estimating the conformation change of biomolecules under a defined molecular mechanics (MM3) force field, which was developed for calculating the structures and energies, including conformational energies, rotational barriers and heat of formation for large hydrocarbons. We use this software to estimate the minimum free energy of EP and its conformation change at changed temperature, to provide theoretical understanding for the self-assembly process.

We built EP I according to its amino acid sequence in extended conformation, minimized its conformational free energy by using molecular dynamic simulation in vacuum, then placed it in a 24 Å water droplet and simulate its conformational change at 25 °C and 50 °C. (Figure 4-12).
To find the difference in molecular conformation of EP I at 25 °C and 50 °C, we used an online software, DETAREA 1.1 (http://pauli.utmb.edu/cgi-bin/get_a_form.tcl), Sealy Centre for Structural Biology and Molecular Biophysics, University of Texas Medical Branch), to calculate the molecular surface area of EP I. It is found that, for EP I at 25 °C, the polar area is 20.68 nm², the apolar area is 57.87 nm²; for EP I at 50 °C, the polar area is 20.21 nm²,
the apolar area is 57.76 nm². We can see the surface area change mainly comes from a decreasing in polar area, while there is limited change for the apolar surface area. Using CAChe, we estimated the number of water molecules surrounding the EP molecule. It is found there are 219 water molecules within 2 Å range of EP I at 25 °C, while there are only 167 water molecule within the same distance from EP I at 50 °C. It is clear that at low temperature, the surface of EP I is more hydrophilic than at elevated temperature, and there are more surrounding water molecules.

Based on the protein structure file generated from this molecular dynamics simulation, we have also tested the feasibility of continuing this dynamics simulation of the temperature dependent conformational change of EP I in solution using Gromacs. After 10 ns simulation, it was found that the radius of gyration of EP I changed from 1.42 nm to 1.20 nm at 50 °C, and from 1.46 nm to 1.27 nm at 25 °C. From the same starting structure, the hydrophobic surface of EP I decreased to 38.15 nm² at 50 °C, while it changed to 41.74 nm² at 25 °C. This result indicates the size of EP I molecule is smaller at 50 °C than at 25 °C in solution, which is probably due to the hydrophobic effect at elevated temperature. (Figure 4-13)
Figure 4-13. Diagram of radius of gyration change of EP I in solution at 25 °C (A) and at 50 °C (B). Below are the solvent accessible surface area of EP I at 25 °C (C) and at 50 °C (D), the black line is for the hydrophobic surface.
Both the dynamics simulation results from Gromacs and CAChe are very similar to the conformational change of repetitive (VPGVG)\textsubscript{18} simulated at different temperatures by Valerie Daggett. In this work performed on a 7421 Da polypeptide, it was found that the number of hydration water molecules decreases as the temperature increases, from 437 at 25 °C to 400 at 42 °C. The solvent accessible surface area decreases as well, from 60.08 nm\textsuperscript{2} at 25 °C to 57.00 nm\textsuperscript{2} at 42 °C. The radius of gyration of the molecule decreased from 15.6 Å to 13.3 Å.

This molecular dynamics simulation supports the hypothesis that water molecules form a shell around the EP peptide to protect the hydrophobic interaction at low temperature. At elevated temperature, this water shell is interrupted and the hydrophobic surface is exposed so that elastin-like peptides are able to aggregate or self assemble due to the hydrophobic interaction.

These FT-IR spectra of the secondary structure of EP, AFM single molecule force spectroscopy of the unfolding of EP and the molecular dynamics simulations of EP at different temperatures indicate the repetitive -PGVGVA- segment within EP forms a β-strand at elevated temperature. We believe this is the reason why EP aggregates in solution and assembles into an ordered structure on hydrophobic HOPG.
5 Conclusion and Discussion

5.1 Discussion

We have studied the self-assembly of EP on HOPG and in solution as a function of temperature using atomic force microscopy and dynamic light scattering. Compared with other native or specifically designed molecules, additional information can be found in understanding the thermodynamics and kinetics of the molecular self-assembly process.

Molecular self-assembly is the spontaneous association of building blocks/modules with unique structure into well-defined higher-ordered forms. During this process, different kinds of noncovalent or weak interactions, such as van der Waals forces, Coulombic interactions, hydrophobic interactions and hydrogen bonds are involved. These weak interactions allow for reversibility and optimization during assembly. The energy barrier to break an undesired weak interaction is much lower than that of strong covalent bonds. Meanwhile, the overall free energy for the entire molecule is sufficient to ensure the molecule is in its native desired conformation.

Zhang et al. has used this principle to design short peptide segments to self-assemble into nanofibers.\textsuperscript{118-120} Peptides of alternative hydrophilic and hydrophobic amino acid residues readily form β-sheet structures with two distinctive surfaces in water as expected. (Figure 5-1) The hydrophobic surfaces shield themselves from water and the complementary charged amino acid residues on the hydrophilic surface form ionic bonds similar to molecular “Lego”. Depending on the sequence of the amino acids on the hydrophilic side of molecule, these peptides can self-assemble into interwoven matrices, which is excellent for cell proliferation due to the high water content in the surrounding environment. However, this type of self-
assembled system is not appropriate for synthesizing functional materials with specific three-dimensional structures.

Figure 5-1. Molecular structure of a series of peptides capable of self-assembling in aqueous solution, RAD16-I, RAD16-II, EAK16-I and EAK16-II. (A) They contain both hydrophobic and hydrophilic surfaces, which function as pegs and holes to facilitate the further extension of these “peptide Lego”s. (B) AFM image of nanofibers formed from RADA16-I, which has been used as scaffold for cell culture and regenerative medicine. (Adapted from reference 120)

In order to form ordered three-dimensional structures of large molecules, the correct orientation of interaction forces on the surface of the molecules is required. However, it is recognized that the strong attractive forces are not ideal in this process. In such a situation, particles rapidly and strongly adhere to each other so that they easily form gel-like aggregates. To form an ordered crystal, the ideal approach is to have very small repulsive interaction forces on the particle surface and restrict the free volume under a threshold, which allows their re-arrangement during the self-assembly process in case any errors occur. However, this volume-restricted method always results in simple closely-packed lattices, which limits the size of a crystal and their application. By manipulating the electrostatic forces between large particles through introducing tethered charge groups, the structure formed can be precisely
adjusted as desired. Interestingly, similar ideas have been used in the crystallization of large biomolecules for more than 20 years.

It is known, for the crystallization of membrane proteins, detergents not only protect the hydrophobic surface of the proteins, but also provide additional interactions for orienting the proteins in the same direction. However, for a given protein, how it interacts with detergents could be very different. There are several aspects that need to be considered include the strength of the interaction between the head groups, the length and structure of the hydrophobic tail, and the association of the detergent with protein. For example, non-ionic detergents are widely used not only because they have little effect on the protein structure, but also because their mild interaction could prevent rapid precipitation. Lipids are amphipilic molecules as well, but they are rarely used to solubilise membrane proteins and can only reconstitute proteins into two-dimensional crystalline structures, because of their bulky tail structure.

Elastin-like peptides, the recombinant tropoelastin-like peptides containing characteristic repeat segments composed of alternating hydrophobic and cross-linking domains, is a good model system for studying the self-assembly of large biomolecules and their association with detergents. It is known that at elevated temperature, the structure of the water shell around the hydrophobic segment can be destroyed so that EP molecules are able to aggregate by the hydrophobic interactions. Since the coacervation temperature of EP decreases with an increased number of hydrophobic segments, and the free energy to break the water shell structure remains the same, this suggests that the sum of the number of available hydrophobic segments is reflected in the apparent coacervation temperature of a single molecule. In aqueous solution, EP randomly associates into amorphous aggregates. On HOPG surface, EP self-assembles into multilayer ordered fibers in the hexagonal pattern.

There are two reasons for the formation of these particular structures. The first is that in solution, the EP molecules have 360° accessibility to the existing nucleus. However, fibril
growth on the HOPG surface is necessarily constrained because of the supporting substrate. Another reason is that the interaction forces available for directing the self-assembly of EP molecules are relatively weak. In order to manipulate EP self-assembly, one needs to either apply external forces to lower the free energy barrier or provide appropriate templates. We have shown that heterogeneous nucleation of EP can be controlled by the nature of the supporting substrate (HOPG vs. mica) and this process appears to involve the formation of an intermediate EP layer that facilitates sequential fiber formation on top of the atomically smooth HOPG surface.

It is interesting to find that the EP molecules directly associate with existing fibers without the process of navigating the surface to find the right position. It is believed that, before association, there is a dense phase of EP close to the HOPG surface. When the solution temperature changes, EP molecules undergo a conformational change and then bind to the surface. This process is different from the normal nucleation and crystallization of proteins, in which molecules use the existing step edges to lower the free energy barrier. We believe one reason is that the interaction among EP molecules is relatively weak, from its small percentage of polar amino acids. Another reason is the step edge of HOPG is not compatible with EP molecular structure. There are two types of edge in a single-layer graphite sheet, zigzag and armchair edges. They are almost in linear shape in the length scale of 100 nm or longer, but they do have slight irregularity at atomic scale. (Figure 5-2) Chemically, the edge of HOPG layer is different from the basal plane, at where the conjugated π- structure is disrupted, the carbon atom has been oxidized to form a carboxylic acid group.
The failure to crystallize EP on HOPG surfaces indicates that hydrophobic interactions alone are not sufficient enough to direct its orientation in three dimensions. The introduction of additional interaction sites on the molecule surface is preferred. Due to their characteristic amphiphilic structure, detergents are helpful not only for protecting the hydrophobic surface from random aggregation, but also for directing the orientation of molecules through the additional hydrophilic interaction site. However, there is no general principle for selecting the right detergent and optimizing its association with a given protein, even for proteins with known structure. Most cases rely on trial and error. A direct consequence is that small differences in the composition and mixing procedure could result in dramatically different results. This could result in either missing a potentially promising condition or wasting time on misleading protein-detergent combinations.

We have studied the assembly of EP I and EP II with HG, OG, DM and zwitterionic detergents. This complete process has been summarized in one diagram as shown in Figure 5-3. In the presence of detergents, the hydrophobic interaction of EP at elevated temperature was dramatically suppressed. It is clear that the coacervation temperature of EP increases in the presence of any of these four detergents. It is also interesting to note that EP is still able to
form ordered structures for certain types of detergents, as can be seen from the narrow distribution of particle size in solution and the hexagonal array of fibers formed on HOPG at low temperatures. Of these detergents, OG formed a well-structured three-dimensional complex with EP that was never observed before. Compared with other detergents tested, it is understandable from its non-ionic head group and medium-sized hydrophobic tail. DM is also a potential candidate, which is in agreement with its increasing use in the crystallization of membrane proteins. However, several promising experiments were difficult to repeat. We believe the reason is that the self-assembly of EP with detergents is a sensitive process, which may include the injection and mixing of sample solution, and the temperature cooling rate. Further investigation of these relating aspects is necessary to optimize the process.

Figure 5-3. Illustration of the schematic showing the temperature-dependent ordered self-assembly of EP and its complex with detergent in solution and on HOPG surface.
AFM is a powerful tool for characterizing the initial stages of self-assembly, including direct observation of nucleation and growth. We believe this approach can be used for the screening of protein crystallization. Using conventional crystallization screening methods, it is easy to miss the nucleus if the particle size is not big enough. To avoid this, the protein concentration needs to be increased. However, it could cause amorphous aggregates if the precipitation happens too quickly. AFM is able to find the crystalline structure at its initial stage while the concentration of solution does not need to be very high. This can save the consumption of protein and the problem of random aggregation.

5.2 Future Recommendation

We have studied the in situ self-assembly of tropoelastin-like peptides in solution and on HOPG surface. However, many details about this process remain to be revealed. The main obstacle is the unknown molecular conformation of EP and its large hydrophobic segment. To solve these problems, other useful approaches can be applied.

From published results it has been shown that the cross-linking segment from exon 21, 23 within elastin-like peptides, which is likely to form α-helical structure, will also affect the self-assembly characteristics of EPs other than the hydrophobic domain. By substituting the flexible “hinge” -GVGTP- segment with a rigid “unhinged” all-alanine segment, between the two long -AAAAAKAAK- domains, the percentage of the α-helical structure within the entire EP molecule is increased. It is interesting to find the coacervation temperature of these “unhinged” EP I and EP II are much lower than those flexible “hinged” EPs, which has been attributed to the more rigid rod-like structure. It is known that the self-assembly of molecules is closely related with their molecular structure. These “unhinged” EP molecules will be very interesting for us to further carry out studies on the self-assembly of these proteins under similar conditions. We would test to see whether they are able to form similar ordered fibers on HOPG as the native EPs, and whether the morphology, e.g. the width of the fiber, will change. Such insights will provide additional information on the relationship between the self-assembly behaviour of EP molecules with their molecular structures.
Molecular dynamics simulation is a powerful supplement to experimental approaches, especially for avoiding practical difficulties, saving cost and revealing molecular details. We used the computational chemistry package, CACHe™ to simulate the folding of EP as a function of temperature. The observation of destruction of the hydration shell is consistent with theoretical predictions; additional insights into the free energy change are necessary. Furthermore, simulating the assembly of EP molecules in solution is beyond the capability of CACHe™. For dedicated molecular dynamics simulation, Gromacs is a powerful software for its optimised algorithm and faster calculation speed. Rauscher et al. have studied the hydration and backbone flexibility of polypeptides using Gromacs and its effect on forming either elastomeric or amyloid fibrils. 65 It is suggested that prolines and glycines are the two key components. Rigid prolines induce polyproline II (PPII) structures and intrinsically reduce the ability to form hydrogen-bonded turns, and plastic glycines determine the extent of backbone hydration in monomers. There is a threshold for the composition of proline and glycine, above which amyloid fibril formation is impeded and polypeptides tend to form elastomeric fibers. This molecular simulation is in agreement with experimental results,127 and provides a unified model for understanding protein aggregation from the aspect of the hydration and conformational change of the backbone structure. Gautam Pennathur et al. has used Gromacs to simulate the folding of β-hairpins for periods of up to 250 ns.128 It is shown that hydrophobic interactions are necessary for initiating the folding of peptide, and the peptide is stabilized by hydrogen bonds. Without hydrophobic interactions, the peptide fails to fold. The introduction of salt-bridges compensates for the hydrophobic interaction at some extent. A problem with traditional MD is that the simulated system tends to be trapped in one of the many local minimum energy states. With optimized replica exchange molecular dynamics (REMD) technique, copies (replicas) of the system are simulated independently under the same condition, and at regular intervals, exchange configuration between thermally adjacent replicas.129 This technique enables a replica to escape local energy minima by exchanging configuration with a replica at higher temperature. Using this method, the folding behaviour
and helical conformation of amine-functionized \( m \)-poly(phenyleneethynylene) (\( m \)-PPE) was simulated more effectively than in the traditional way.\(^{130}\)

On base of the molecular structure of EP molecules generated from CAChe, we can simulate the folding of EP in solution at different temperatures and salt concentrations. With selected combination of hydrophilic and hydrophobic segments, this series of simulations could provide insights for the effect of the hydrophobic segment on the conformational change of a single protein, and its relationship with surrounding water molecules. We can also simulate the assembly of several EP molecules at coacervation conditions, which is able to reveal the minimum amount of molecules required to form a nucleus, on which extended fibrils can form.

In the last ten years, solid-state nuclear magnetic resonance spectroscopy (ssNMR) has become a powerful technique in estimating the structure and interaction of proteins. For NMR, the nuclear spin is detected within a magnetic or electric field. The interaction between atoms is different due to spatial proximity and chemical interactions so the nuclei absorb energy at different frequencies. With limited mobility, the anisotropic interaction has a large effect on the behaviour of nuclear spins. For the traditional solution-state NMR, the anisotropic interaction is averaged by the Brownian movement of the molecule so that the spectrum is narrow, e.g. 0.001 - 0.1 ppm. This approach is difficult for molecules in restricted environments such as crystalline structures or micelles, because the spectrum is wider. By using magic angle spinning (MAS), 54.7°, the nuclear dipole-dipole interaction averages to zero and the chemical shift anisotropy averages to a non-zero value. The quadrupolar interaction is only partially averaged by MAS leaving a residual secondary quadrupolar interaction, which cause the signal to become much narrower.

Because ssNMR is able to determine interatomic distances and place constraints on backbone and side-chain torsion angles of molecules, besides its wide application in studying the structural and topological properties of membrane proteins in oriented membranes,\(^{131}\) it
has been used for developing models of amyloid fibrils. By using TEM and ssNMR, it is found that Aβ_{1-40} fibrils growing from quiescent peptide solution is different from those growing from solution that is gently agitated by rotary motion. Fibrils from quiescent condition exhibit a periodic twist, have maximum diameters of 9 ± 1 nm and do not adhere to each other. Fibrils from agitation motion have diameters of 5 ± 1 nm, they lack a twist but adhere to each other. ssNMR data suggest that the β strand segments in quiescent Aβ_{1-40} fibrils include residues 10 to 14, 16 to 22, 30 to 32, and 34 to 36, whereas those in agitated Aβ_{1-40} fibrils include residues 10 to 22, 30 to 32, and 34 to 36. The data for agitated fibrils is consistent with recent models for the molecular structure of Aβ_{1-40} protofilament, while the data for quiescent Aβ_{1-40} fibrils indicate a different structure in both molecular conformation and supramolecular organization. This result shows that the growth condition is also critical for the fibril formation, which may depend on the interface between peptide sample solution with air and container surface, but not solely by amino acid sequence and thermodynamic control.

This technology could be very useful in investigating the possible structural differences between the ordered EP fibrils formed on HOPG surface and the amorphous thin film formed in solution. Our results show the fibril growth on HOPG surface is generated from the alignment of individual folded EP molecules. In solution, the film is generated from the cluster of globular spheres of large size. We can use ssNMR to study the structure of EP in these two different morphologies, e.g. folded or extended, so as to find clues to control the assembly of EP.

Additionally, we have used FT-IR to study the secondary structure of EP before and after the adsorption on HOPG surface. In the presence of detergents, the organization of detergent-EP complex is unknown. We can use FT-IR to study the self-assembly of detergent, those linear hemi-cylindrical arrays, on HOPG surfaces and its association with EP. From these experiments, we expect to detect the structure difference of EP before and after its
association with detergents and the interaction between these two species so as to find out
clues about how they are organized.

5.3 Conclusion

This study of the self-assembly of EP and its complex with various detergents at
elevated temperature, either on HOPG or in solution, has provided useful information about
the heterogeneous nucleation of molecules on surfaces and the association of amphiphilic
molecules with proteins with large hydrophobic segments. A two-dimensional crystalline
structure of EP with detergent, which was never observed before, was produced by an
optimized mixing procedure. Besides its direct application in the crystallization of membrane
proteins, we believe it is also very useful in controlling the self-assembly of large particles,
which is critical for the synthesis of nanomaterials with novel properties.
6 Appendices

6.1 Atomic Force Microscopy Imaging

All AFM imaging was conducted on a Digital Instruments Nanoscope IIIa Multimode scanning probe microscope (Digital Instruments/Veeco, Santa Barbara, CA) equipped with an “E” scanner having a maximum lateral scan area of 14.6 x 14.6 µm. Both contact and tapping mode AFM imaging was employed, depending on the study. For tapping mode studies, 120 µm long oxide-sharpened silicon nitride V-shaped cantilevers (Type DNP-S, Veeco Instruments), with nominal spring constant of ~ 0.38 N/m (provided by the manufacturer) were used. For contact mode studies, 200 µm long oxide-sharpened silicon nitride V-shaped cantilevers were employed. In contact mode, all images were collected at 90 deg to the fast scan axis to deconvolve topographical data from frictional / torsional data. The AFM cantilevers were irradiated with UV light prior to use to remove any adventitious organic contaminants. A contact/tapping mode fluid cell was sealed against the supporting substrate freshly cleaved mica or HOPG (Grade SPI-3, Structure Probe, Inc., West Chester, PA) with a silicone O-ring. The fluid cell was fitted with inlet and outlet tubing to allow exchange of solution in the cell during imaging. All AFM images were captured as 512 x 512 pixel images at scan rates of between 2 to 3 Hz using a tip oscillation frequency of 6-10 kHz, in the case of tapping mode imaging. All images were plane-fit and zero-order flattened prior to analysis. All analyses were performed using the Nanoscope software (version 4.42r9)
6.2 FT-IR Reflectance and Absorbance Spectroscopy

Infrared spectra of EP were collected using Nicolet™ Continuum™ infrared microscope and Nexus 670 spectrometer (ThermoNicolet, Madison, WI) equipped with liquid nitrogen cooled MCTA detector and KBr beamsplitter. The spectra of EP monomers were obtained by grinding the pure EP sample with potassium bromide (KBr) powder, collected within the Nexus 670 spectrometer using traditional transmission mode, from the average of 64 scans at a 4 cm\(^{-1}\) resolution in a spectral range from 4000 cm\(^{-1}\) to 650 cm\(^{-1}\). The spectra of the thin film of EP adsorbed on HOPG were collected by using the reflectance mode on the Continuum microscope with the Reflachromat™ objective. The EP thin film was allowed to form at 40°C on HOPG. The IR beam was projected on the sample surface at 90° and spectra were collected on the thin film formed on the surface between the buffer salt particles. The spectra were collected from the average of 128 scans at a 4 cm\(^{-1}\) resolution in a spectral range from 4000 cm\(^{-1}\) to 650 cm\(^{-1}\). All the spectra were collected and base line corrected by using the OMNIC E.S.P 5.2a software (ThermoNicolet, Madison, WI).

6.3 Dynamic Light Scattering

DLS was performed on BI-200SM Goniometer (Brookhaven Instruments Corporation, NY) equipped with a BI-9000 AT digital autocorrelator and 514 nm Argon laser operated at 50 mW (Lelxel Laser Inc., Fremont, CA). Scattered light was collected at 90° using a 200 µm pinhole. The sample solution was heated from 25 °C to 50 °C, then cooled down to 5 °C, measuring particle size change during this process at each interval of 5°C. The effective diameter of sample is obtained by the cumulant analysis tool of the 200-channel BI-9000AT digital autocorrelator, using BIC Dynamic Light Scattering Software.

6.4 Molecular Dynamics Simulation

The sequence of EPs were generated from CAChe WorkSystem Pro Version 6.1.12.33 (CAChe Group, Fujitsu America, Inc., Beaverton, OR) using amino acid sequence-building function. The folding conformation of the extended long-chain amino acids is first optimized
by using the “conformation from dynamics” under the MM3 force field in vacuum for 100 ps. The folded structure of EPs were simulated in a 24 Å water sphere under the same dynamics simulation for 100 ps, which generates lower energy conformation and trajectory file as a function of simulation time. The further dynamics simulation of the generated EP structure is extended for 10 ns by using GROMACS 3.3.1, with the GROMOS-96 43a1 force field. The system was equilibrated with 50 ps position restraints on the protein atoms to allow the solvent to relax. All systems were run with a 2-fs time step, a non-bonded cutoff of 0.9 nm for Van der Waals interactions. The radius of gyration and hydrophobic surface of the protein are estimated during the period of 10 ns.

6.5 Customized Automatic Temperature Control Circuit

We selected OMEGA CSC32 bench top controller (OMEGA Engineering, Inc., Stamford, CT) as the central unit for it has two parallel outputs, one is for heating and another is cooling, which are adjusted according to the feedback from the temperature sensor. The major problem remains in the connection of these two power sources into the single input of the Igloo Koolmate box, which is the used as the temperature chamber driven by semiconductive heat-transfer Peltier. The working mechanism of the Peltier is the current drives mobile charged carriers from the heat side to the cold side so as to heat up the sample, or vice versa, which means the positive electrode of one input has to be connected with the negative electrode of another input. To solve this problem, we designed a control circuit to separate these two inputs. The working mechanism is when the heating input is in use, the magnetic coil relay cut off the cooling input loop when there is no current, so that they do not interact each other. The circuit is shown as in Figure 6-1.
Figure 6-1. Circuit diagram of the automatic switch introduced between the Igloo Koolmate and the Omega temperature controller. By using this unit, the Koolmate is able to work as either a cooler or heater.
References:

(1) Rajagopalan, R. *Current Opinion in Colloid & Interface Science* 2001, 6, 357.
(3) Osborne, T. B. *American Chemical Journal* 1892, 6, 662-689.


Bellingham, C. M.; Woodhouse, K. A.; Robson, P.; Rothstein, S. J.; Keeley, F. W. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 2001, 1550, 6-19.


Ke, C.; Jiang, Y.; Rivera, M.; Clark, R. L.; Marszalek, P. E. Biophys J 2007, 92, 176-8.

Young, R. J.; Lovell, P. A. Introduction to Polymers (2nd Edition); Chapman & Hall.


(122)Kalsin, A. M.; Fialkowski, M.; Paszewski, M.; Smoukov, S. K.; Bishop, K. J. M.;
Surface Science 2005, 241, 43.
23757-65.
(127)Miao, M.; Bellingham, C. M.; Stahl, R. J.; Sitarz, E. E.; Lane, C. J.; Keeley, F. W. J Biol
(128)Seshasayee, A. S.; Raghunathan, K.; Sivaraman, K.; Pennathur, G. J Mol Model 2006, 12,
197-204.
(131)Bechinger, B.; Aisenbrey, C.; Bertani, P. Biochim Biophys Acta 2004, 1666, 190-204.