



---

Publicly Accessible Penn Dissertations

---


Summer 8-12-2011

# Exploring the World of Helix Association: Disease Mechanism, Basic Folding and Novel Design

Yao Zhang

University of Pennsylvania, yaozhang@sas.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#)

---

## Recommended Citation

Zhang, Yao, "Exploring the World of Helix Association: Disease Mechanism, Basic Folding and Novel Design" (2011). *Publicly Accessible Penn Dissertations*. 973.

<http://repository.upenn.edu/edissertations/973>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/973>

For more information, please contact [libraryrepository@pobox.upenn.edu](mailto:libraryrepository@pobox.upenn.edu).

---

# Exploring the World of Helix Association: Disease Mechanism, Basic Folding and Novel Design

## **Abstract**

Helix association provides an efficient model for studying the fundamental principles behind protein folding. It also serves as a suitable template for the design of proteins with novel functions. This thesis begins by investigating the role of transmembrane helix association in protein folding, where a novel “protein-folding-centric” viral fusion model has been proposed here to explain the membrane-fusion process of paramyxovirus. Furthermore, the forces driving membrane helix association, which determine both affinity and orientation, have been quantitatively studied using a model membrane peptide MS1. Finally, two examples are discussed that illustrate the application of helix association in novel protein design. A pH-switchable drug delivery system for the endosomal escape of biomacromolecular therapeutics has been designed using the helix-association model. The sequence is designed to form a stable water-soluble helix bundle at pH 7.4 and to insert in membrane at lower pH to promote endosomal escape. The most successful sequence shows selective release for biomacromolecule (ATP and miRNA) at lower pH (pH 5.4). The assembly of the designed peptide has been studied in aqueous buffer, detergent micelle and model lipid bilayer using the most successful sequence. Also, the paradigm of helix association has been applied to the design of a membrane metalloprotein, which can serve as a template for further design of membrane metalloenzymes. In summary, the work in this thesis has established an efficient model for helix association that can be used to solve problems in both basic and applied research.

## **Degree Type**

Dissertation

## **Degree Name**

Doctor of Philosophy (PhD)

## **Graduate Group**

Chemistry

## **First Advisor**

William F. DeGrado

## **Keywords**

helix association

## **Subject Categories**

Biochemistry, Biophysics, and Structural Biology

**EXPLORING THE WORLD OF HELIX ASSOCIATION:  
DISEASE MECHANISM, BASIC FOLDING AND NOVEL DESIGN**

Yao Zhang

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

2011

---

Professor William F. DeGrado  
Supervisor of Dissertation

---

Professor Feng Gai  
Chair of Committee

---

Professor Gary A. Molander  
Graduate Group Chair

---

Professor Jeffery G. Saven  
Committee Member

---

Professor Ivan J. Dmochowski  
Committee Member

To my mom Yueqin Qin, my dad Xuewu Zhang  
and my husband Tong Gao

谨以此文献给我的母亲秦月琴，我的父亲张学武  
和我的爱人高侗

## ACKNOWLEDGEMENT

Life is magical. Ten years ago, when I was fighting for the college admission exam in a high school in central China, I never imagined that I would be standing on the other side of the earth working on my Ph.D thesis defense. This amazing journey was not easy. Without the support of these great people, I would not have enjoyed my Ph.D training so much and would not been able to achieve the accomplishments of which am proud.

I feel lucky to have has the opportunity to work on projects I love and work with people I respect. I especially want to thank my thesis advisor, Bill DeGrado. He introduced me to the world of protein science, and he encouraged me to explore new challenges in the field. With his mentoring, I have also developed collaborative and leadership skills, making me an effective team player and project manager.

I want to thank my thesis committee members: Feng Gai, Jeffery Saven and Ivan Dmochowski for generously sharing their knowledge and kindly granting me their support. The insightful discussions with them have shaped my ideas in research and sped up my progress. I especially remember the moment, five years ago when Dr. Gai called me and gave me an offer from Penn Chemistry. I hope he is still happy with his decision.

I would also like to thank those marvelous people in the DeGrado lab. The colleagues with whom I have worked, made me into the person I am now. In particular, I want to thank James Lear, Paul Billings, Ivan Korendovych, Alessandro Senes, Scott Shandler, Jason Donald, Dan Kulp, Bruk Mensa, Gevorg Grigoryan, Rudresh Acharaya, Geronda Montalvo, Michael Bryan, Yibing Wu, Yongho Kim, Lidia Cristian, Cinque Soto, Jade Qiu, Jenny Hu, Shao-Qing Zhang, Brett Hannigan and Amanda Reig. It has been a great pleasure to work with you guys and I have learned so much from our frequent interactions. I also want to thank David Tellers, Lin Guo and Kate Smith who have strongly supported my research outside of the DeGrado lab.

I want to thank my friends who have shared these wonderful years with me. Thank you, Lin, for picking me up at the airport when I arrived at Philadelphia and making my first meal in the U.S. You are like my older sister. Thank you, Yuting and Wulan. It is my honor to be one of the “PKU Three Girl Team” and have your support. Thank you, Yaoyao, Yanshan, Li, Yuanquan, Ang, Zhongqiang, Yi, for all the nice memories during weekends and holidays, which have kept me from succumbing to the pressures of work.

No doubt, the most important people to me are my family and I owe them so much. Thank you, Dad and Mom, for your unconditional love and support. Dad trained me to be a logical and critical person and mom always motivated me to be passionate in life. Their influence helped shape my personality, of which I am very proud. Particularly,

I deeply appreciate their respect and support in every big decision I have made. Thanks to their understanding, I am able to freely pursue my dream and explore life.

Speaking of exploration in my life, I would not be able to enjoy this journey without a very important person---my best friend, my soul mate and my husband Tong. With his love and support, I have been able to go through the most difficult times in my graduate training. Because of him, my life is full of happiness, and I have enjoyed every minute since we met. Honey, thank you for being my co-pilot.

# **ABSTRACT**

## **EXPLORING THE WORLD OF HELIX ASSOCIATION: DISEASE MECHANISM, BASIC FOLDING AND NOVEL DESIGN**

**Yao Zhang**

**William F. DeGrado, Thesis Advisor**

Helix association provides an efficient model for studying the fundamental principles behind protein folding. It also serves as a suitable template for the design of proteins with novel functions. This thesis begins by investigating the role of transmembrane helix association in protein folding, where a novel “protein-folding-centric” viral fusion model has been proposed here to explain the membrane-fusion process of paramyxovirus. Furthermore, the forces driving membrane helix association, which determine both affinity and orientation, have been quantitatively studied using a model membrane peptide MS1. Finally, two examples are discussed that illustrate the application of helix association in novel protein design. A pH-switchable drug delivery system for the endosomal escape of biomacromolecular therapeutics has been designed using the helix-association model. The sequence is designed to form a stable water-soluble helix bundle at pH 7.4 and to insert in membrane at lower pH to promote



endosomal escape. The most successful sequence shows selective release for biomacromolecule (ATP and miRNA) at lower pH (pH 5.4). The assembly of the designed peptide has been studied in aqueous buffer, detergent micelle and model lipid bilayer using the most successful sequence. Also, the paradigm of helix association has been applied to the design of a membrane metalloprotein, which can serve as a template for further design of membrane metalloenzymes. In summary, the work in this thesis has established an efficient model for helix association that can be used to solve problems in both basic and applied research.

# Table of Contents

|  |      |
|--|------|
| Dedication .....   | ii   |
| Acknowledgement .....  | iii  |
| Abstract .....   | vi   |
| Table of Contents .....  | viii |
| List of Tables .....   | xi   |
| List of Figures .....  | xii  |
| 1 Chapter 1: Introduction: The fundamentals of protein folding and the model of helix association..... | 1    |
| 1.1 The significance of protein folding .....  | 1    |
| 1.2 The impediments to studying protein folding .....  | 2    |
| 1.3 The helix-association-based model for studying protein folding.....                                | 3    |
| 1.4 The coiled-coil model.....   | 4    |
| 1.5 Case study of GCN4-p1: the folding in an aqueous system.....                                       | 6    |
| 1.6 The application of coiled coil-based design in aqueous systems .....                               | 10   |
| 1.7 Understanding membrane protein folding with coiled-coil-based design .....                         | 11   |
| 1.8 Case study of MS1: folding in micelles and bilayers.....   | 13   |
| 1.9 Membrane structural motifs .....   | 15   |
| 1.10 The application of sequence-based protein folding studies .....                                   | 16   |

|      |  |    |
|------|--|----|
| 1.11 | The exploration of helix association in both aqueous and membrane systems..  | 17 |
| 2    | Chapter 2: Transmembrane Orientation and Possible Role of the Fusogenic Peptide from PIV5 Virus in Promoting Fusion..... | 18 |
| 2.1  | Abstract .....   | 18 |
| 2.2  | Introduction .....   | 20 |
| 2.3  | Results .....  | 26 |
| 2.4  | Conclusion and Discussion .....  | 41 |
| 2.5  | Method .....   | 50 |
| 2.6  | Contributions.....   | 60 |
| 3    | Chapter 3: Experimental and computational evaluation of forces directing the association of transmembrane helices.....   | 61 |
| 3.1  | Abstract .....   | 61 |
| 3.2  | Introduction .....   | 63 |
| 3.3  | Results and Discussions .....  | 65 |
| 3.4  | Conclusion and Discussion .....  | 81 |
| 3.5  | Methods.....   | 82 |
| 3.6  | Contributions.....   | 90 |
| 4    | Chapter 4: pH-switchable peptide model for membrane insertion and pore formation.....                                    | 91 |

|     |   |     |
|-----|---|-----|
| 4.1 | Abstract .....  | 91  |
| 4.2 | Introduction .....  | 92  |
| 4.3 | Results .....   | 96  |
| 4.4 | Conclusion and Discussion .....   | 118 |
| 4.5 | Methods .....   | 124 |
| 4.6 | Contributions:.....   | 128 |
| 5   | Chapter 5: The design and characterization of a membrane metalloprotein ..... | 129 |
| 5.1 | Introduction .....  | 129 |
| 5.2 | Results .....   | 130 |
| 5.3 | Conclusion and discussion .....   | 136 |
| 5.4 | Methods.....  | 137 |
| 5.5 | Contributions.....  | 139 |
| 6   | Chapter 6: Conclusions and Discussions.....                                   | 140 |
| 7   | Chapter 7: References .....   | 143 |

## List of Tables

|   |     |
|---|-----|
| Table 2.1 Conservation of small residues (Ala, Gly) within four virus families.....   | 24  |
| Table 3.1 Association states of MS1 variants determined by AUC .....  | 70  |
| Table 3.2 $pK_{dim}$ and $pK_{ox}$ obtained from analysis of data in Figure 2A .....  | 70  |
| Table 3.3 Degree of association of air-oxidized MS1 variants in DPC micelles determined by analytical ultracentrifugation .....                               | 76  |
| Table 3.4 Energetic contributions for minima in energy landscape <sup>a</sup> .....   | 80  |
| Table 3.5. Crick parameters at the potential energy minimum.....  | 80  |
| Table 4.1 Amino-acid choices considered in design .....   | 99  |
| Table 4.2 The sequence of PSPF peptides .....   | 100 |
| Table 4.3 RBC Lysis assay of PSPF peptides.....   | 102 |
| Table 4.4 Trp fluorescence of PSPF- series peptides with various amounts of lipid vesicles.....   | 104 |
| Table 4.5 Apparent molecular weight and calculated oligomerization state based on size exclusion chromatography for PSPF- EKG and PSPF- DKG at both PHs ..... | 107 |
| Table 4.6 Analytical ultracentrifugation (AUC) sedimentation equilibrium for PSPF-EKG and PSPF- DKG at pH 5.5 and 7.4. ....                                   | 110 |
| Table 4.7 Fitting results for CD thermal denaturation of PSPF- EKG at both pH 7.4 and pH 5.5.....   | 113 |
| Table 5.1 The sequences of DF <sub>tet</sub> B, MS1 and MDF .....   | 130 |

## List of Figures

- Figure 1.1 A schematic coiled coil heptad repeat for a parallel dimer (A) and a parallel heptamer (B) The schematic map is made via program DrawCoil 1.0, developed by Dr. Gevorg Grigoryan(<http://www.gevorggrigoryan.com/drawcoil/>) ..... 5
- Figure 1.2 The sequence and sketch map of heptad repeat for GCN4-p1. The schematic map is made via program DrawCoil 1.0 ..... 7
- Figure 1.3 The sequence of MS1 and variants..... 14
- Figure 2.1 Sequence conservation suggests a continuous helix including heptad repeat A (HRA) and the fusion peptide. (A) Postfusion crystal structure of the soluble domain of closely related hPIV3 virus F protein <sup>135</sup>. Shown in magenta is HRA. Below HRA, in the postfusion membrane, is the predicted location of the fusion peptide. (B) Heptad repeat of the fusion peptide and HRA. The beginning of the crystallographic resolved region of HRA is shown in magenta. Heptad repeats of small residues in the fusion peptide are boxed. (C) Sequence entropy of the fusion peptide and HRA can be fit to a single sinusoidal function with period of  $3.47 \pm 0.02$  residues/turn ( $r = 0.51$ ). (D) Sequence entropy of the fusion peptide alone can be fit to a single sinusoidal function with a period of  $3.51 \pm 0.08$  residues/turn ( $r = 0.59$ ). ..... 23
- Figure 2.2 Analytical ultracentrifugation of fusion peptide wild-type (A, B) and mutant Q120A (C, D). Single species fitting of wild-type (A) and mutant Q120A (C) PIV5 fusion peptide suggests both associate as hexamers. The top of each panel shows the residuals of single species fitting. The species weight fraction is shown for wild-type (B) and mutant Q120A (D), indicating that hexamer is the dominant species composition for the wild-type, while oligomerization of the mutant Q120A requires significantly higher mole fractions..... 28
- Figure 2.3 Sedimentation equilibrium analysis of the wild-type (A) and mutant Q120A (B) fusion peptides in DPC micelles. The absorbance was measured at 280 nm. The data was fitted as a monomer-hexamer equilibrium, resulting in a  $pK_{\text{dissociation}}$  of 20.1 for the wild-type and a  $pK_{\text{dissociation}}$  of 10.2 for the mutant Q120A. The top of each panel shows the residuals of the fit. In this data analysis the concentration of peptide is the mole ratio of peptide/detergent and therefore is unitless. .... 29
- Figure 2.4 Analytical ultracentrifugation of C-term-TM peptide alone (A) and C-term-TM peptide-fusion peptide mixture (B). No significant curvature has been observed in AUC data of TM peptide alone (A) suggesting that C-term-TM peptide does not self-associate under these conditions. Single species fitting for a 1:1 mixture of C-term-TM peptide and fusion peptide (using the original sequence without a Trp label) suggests a strong interaction between these peptides..... 30

Figure 2.5 CD spectra of fusion peptide wild-type (A) and mutant Q120A (B). The spectra show that both wild-type and Q120A are predominantly  $\alpha$ -helical at a peptide:DPC ratio of 1/200. .... 32

Figure 2.6 Attenuated total reflection IR spectroscopy (ATR-IR) of fusion peptide wild-type (A) and mutant Q120A (B) in phospholipid (POPC) bilayers. The sharp peak at  $1656\text{ cm}^{-1}$  is indicative of alpha helical secondary structure. The TM orientation is demonstrated by the much greater intensity of the  $1656\text{ cm}^{-1}$  amide I bond for parallel ( $0^\circ$ ) versus perpendicular ( $90^\circ$ ) polarized incident light (relative to the membrane normal). . 33

Figure 2.7 Computational prediction of the hexameric helical bundle. (A) Heat map shows predicted energies of coiled-coil models with different radii and phases. Regions in black are predicted to be more energetically favorable. (B) Backbone atom RMSD of five coiled-coil models selected from (A) and one antiparallel model generated using an analogous procedure. Parallel models are labeled according to their phase angle. .... 37

Figure 2.8 Final structures of the MD simulations of different hexamer models and coiled coil parameters as a function of time. (A-F) Final backbone structures in cartoon representations of five parallel models of the hexamers corresponding to the phase angles,  $\phi$ , of  $40^\circ$  (A),  $43^\circ$  (B),  $88^\circ$  (C),  $192^\circ$  (D),  $300^\circ$  (E), and of the antiparallel model (F). The heavy atoms of the Gln120 side chains are drawn in blue. (G-L) Coiled coil parameters as a function of time of the six models, in the same order as in (A-F); the deviations from the initial values of the pitch angle  $\alpha$  and of the phase angle  $\phi$  ( $\Delta\alpha$  and  $\Delta\phi$ ) are plotted in green and red, respectively; the deviation of the radius ( $\Delta r$ ) is plotted in blue. .... 38

Figure 2.9 Computational model of the PIV5 F fusion peptide hexameric bundle. (A) The Q120 residues form hydrogen bonds with one another as well as waters on the interior. (B) Side view shows the bundle oriented with the N-terminal end (which presumably faces the cellular interior) up. Water is shown in blue. Not shown for clarity are the phospholipids as well as one helix closest to the viewer. .... 39

Figure 2.10 Density of oxygen atoms from water molecules through the pore of the hexamer bundle. From left to right, data from the following models are plotted:  $\phi = 40^\circ$ ,  $43^\circ$ ,  $88^\circ$ ,  $192^\circ$ ,  $300^\circ$ , and the antiparallel model. The density is expressed in number of atoms per  $\text{\AA}$ , and plotted as a red line along the horizontal axis. For comparison, the density profile from the  $\phi = 40^\circ$  model is also shown as a dashed blue line in the other models (duplicated for the antiparallel model, which is symmetric with respect to the membrane plane). For those models featuring pore-lining Gln120 side chains, the region occupied by their nitrogen and oxygen atoms is highlighted in green. .... 40

Figure 2.11 Provisional model of PIV5 fusion. (A) Schematic diagrams of the limiting extremes of lipid-centric and pin-prick fusion. (B) Shown is a model of the conformational change of the F protein (FP) hexamer (6HB) from a pre-hairpin, extended intermediate (left) to a point of membrane apposition (middle) and finally to the postfusion state (right). Proposed conformations of the FP 6HB are shown in the insets

along with 90° rotations. Note the increased tilt of the FP moving from the extended intermediate to the point of membrane apposition as well as the recruitment of lipid headgroups to the nascent pore. FPs are shown in red and blue, C-term-TMs are shown in magenta and yellow. The C-term-TM in the middle image contains two trimeric structures<sup>138</sup>..... 48

Figure 3.1 Helical wheel (A), computational model showing side and top view (B), and sequence of MS1 variants (C). MS1, -Gly, -Ala, -Val, and -Ile are N-terminally Cys-modified. -Gly-(Ct) is C-terminally Cys-modified. The variable **a** positions are shown in green, and the Leu at 'd' in red. All peptides had an N-terminal acetyl group..... 67

Figure 3.2 CD Spectra of MS1 variants (units are deg cm<sup>2</sup> dmol<sup>-1</sup> res<sup>-1</sup>). The spectra show that all peptides are predominantly  $\alpha$ -helical at a peptide/detergent ratio of 1/1000 or 1/50 (matching the concentration range for the thiol-disulfide exchange equilibria). The magnitude of the ellipticity at 222 nm is the same within the experimental error, which derives primarily from the concentration determination used to compute the mean residue ellipticity. Buffer conditions and other methods are given below. .... 68

Figure 3.3 Sedimentation equilibrium analysis of the reduced MS1 variants in DPC micelles: (A) MS1-Gly, (B) MS1-Ala, (C), MS1-Val, (D) and MS1-Ile. The absorbance was measured at 280nm. The top panels shows the residuals of single species fitting to data at respectively 30, 35, 40, 45, 48 KRPM. (E) MS1-Ala is fitted with monomer-dimer equilibrium, resulting in  $pK_{dim}$  equal to -1.3. In this analysis the concentration of peptide is mole ratio of peptide/detergent, and hence is unitless. (F) Monomer/dimer species distribution of MSI-Ala as a function of peptide/DPC ratio based on fitting in (E). Buffer conditions and other methods are given below..... 69

Figure 3.4 Simulation of disulfide exchange equilibrium related to various dimerization affinities ( $K_{dim}$ ) and redox ratios ( $K_{ox}$ ). The figure shows theoretical fractions of crosslinked dimer as a function of peptide/DPC ratio at (A) varying  $pK_{dim}$  ( $pK_{ox}=0$ ), and at (B) varying  $pK_{ox}$  ( $pK_{dim}=0$ ). Note that the shapes of the curves vary depending on both parameters. Thus, the figures demonstrates that thiol disulfide equilibrium method is sensitive to both the free energy of association as well as and redox potential for a given peptide..... 71

Figure 3.5 (A) Analysis of the fraction of crosslinked dimer as a function of peptide/DPC ratios for each MS1/variant. The theoretical curve describes the least-squares fit to scheme 1 (Table 2). The standard errors in the experimental points are similar for each peptide and are indicated for MS1-Ile (Others are not so shown for clarity). (B) HPLC chromatogram after redox equilibration of the C-terminal and N-terminal Cys-modified MS1-Gly mixture indicates that MS1-Gly prefers an anti-parallel orientation. The other peaks are glutathione adducts. .... 72

Figure 3.6 Analytical ultracentrifugation of approximately 80% disulfide-bonded MS1-Gly (A) and MS1-Ile (B). The greater degree of curvature in panel A vs. B is indicative



of greater oligomerization. (C) Oligomerization of MS1-Gly and MS1-Ala via formation of antiparallel dimers. .... 75

Figure 3.7 Sedimentation equilibrium analysis of the MS1 variants in DPC micelles: (a) MS1-Ala and (b) MS1-Val. Absorbance is measured at 280nm. Top panels shows residuals of single species fitting to data at respectively 30, 35, 40, 45, 47 KRPM. Peptides were first air-oxidized in buffer. The percentage of air oxidized dimer was determined by analytical HPLC to be (75±5)%. Buffer conditions and other methods are given below..... 76

Figure 3.8 (A) Sampling Crick parameters for parallel ( $r, \theta$ ) and anti-parallel ( $r, \theta, z$ -translation) dimers. (B) Energy landscape showing the difference in computed Leonard-Jones energy  $E_{LJ}$  (for the dimer versus two monomers) of MS1-Gly in a parallel orientation. The minimum in the surface has a helical phase ( $\theta$ ) of 154°, allowing packing of the Gly residues at the helix/helix interface as in Figure 1B. .... 78

Figure 3.9 Electrostatics contribute to the orientation preference of MS1 variant models. The difference between electrostatic energy in the parallel and anti-parallel models correlates with inter-helical distance (average of parallel and anti-parallel distances). The glycine model shows the largest electrostatic energy difference between parallel and anti-parallel orientations..... 89

Figure 4.1 The desired free energy diagram of the designed peptide as a function of pH. Lowering pH should destabilize water-soluble bundle state and stabilize first membrane-associated monomeric state and then, in a concentration-dependent manner, the membrane-inserted channel state..... 94

Figure 4.2 The design concept illustrated using one of the designed sequences (PSPF-DKG). Hydrophobic residues are either lining the bundle the core in the water-soluble state (A), or are facing the lipid membrane in the membrane channel state (B). Dotted circles illustrate potential hydrogen bonding in the channel state. Heptad positions in both panels are labeled according to the water-soluble state. The amino acid choices at each position are shown in Table TX1..... 97

Figure 4.3 Correlation between ATP-release by PSPF peptides and the degree of lipid engagement as assessed by the fractional change of Trp-fluorescence signal upon addition of 200 uM lipid vesicles..... 104

Figure 4.4 Size exclusion chromatography of PSPF-EKG and PSPF-DKG at each pH. Both PSPF-EKG and PSPF-DKG elute as a single species corresponding to the oligomerization of hexamer at pH 7.4 (B). PSPF-EKG elutes as a single-species peak with a significant shoulder at pH5.4 and the major peak corresponds to a formation of hexamer. PSPF-DKG almost fails to elute at PH 5.5 (A). The salt concentration has been increased to 2M and the shoulder of elution peak still exists at pH 5.5 (C, D). .... 106

|  |     |
|--|-----|
| Figure 4.5 AUC sedimentation equilibrium of PSPF-EKG at pH 5.5 (A) and 7.4 (C). Single species fitting of PSPF-EKG suggests a hexameric association state at both pH 7.4 (A) and pH 5.5 (C). For each peptide and pH condition, the top plot shows the single species fitting with residuals above while the below plot shows the species weight fraction. Then the data has been fit with a monomer-hexamer equilibrium model at both pHs. The dissociation state and dissociation energy is shown in Table 6. The weight fraction distributions have also been plot for pH 5.5 (B) and pH 7.4 (D). ..... | 108 |
| Figure 4.6 AUC sedimentation equilibrium of PSPF-DKG at pH 5.5 and 7.4. Single species fitting of PSPF-DKG suggests it associates as a hexamer at pH 7.4 and reaches an apparent molecular weight of approximately 24,000 at pH 5.5. For each peptide and pH condition, the top plot shows the single species fitting with residuals above while the below plot shows the species weight fraction. ....  | 109 |
| Figure 4.7 Circular dichroism of PSPF-EKG suggests an alpha-helical secondary structure at both pHs. ....  | 111 |
| Figure 4.8 Thermal denaturation of PSPF- EKG at pH 7.4 (A) and 5.5 (B). The raw data are fit to the Gibbs-Helmholtz Equation (C, D). ....  | 112 |
| Figure 4.9 The single-species fitting of AUC sedimentation in detergent micelles for PSPF-EKG at pH 7.4 (A) and 5.5 (C). Species weight fraction of PSPF- EKG at pH 7.4 (C) and pH 5.5 (D) as the data were globally fit to a monomer-trimer equilibrium as an example. ....   | 115 |
| Figure 4.10 ATR-IR of PSPF-EKG in phospholipids (POPC) bilayers. The peak at 1656 $\text{cm}^{-1}$ is indicative of alpha helical secondary structure. The orientation is demonstrated by the ratio of peak area of the 1656 $\text{cm}^{-1}$ amide I bond for parallel ( $0^\circ$ ) versus perpendicular ( $90^\circ$ ) polarized incident light (relative to the membrane normal). ....   | 117 |
| Figure 4.11 Model of PSPF-EKG membrane insertion and pore formation upon pH decrease. ....   | 122 |
| Figure 5.1 Frequently observed His/Glu chelated metal system in both natural and designed proteins <sup>229</sup> (A). Four-helix bundle scaffold for MDF (B). ....  | 131 |
| Figure 5.2 CD spectra of MDF (A) and iron-bound MDF (B) in DPC micelles. The spectra show that both apo and bound forms are predominantly $\alpha$ -helical at a peptide:DPC ratio of 1:25. ....   | 132 |
| Figure 5.3 Sedimentation equilibrium analysis of MDF (A) and Co(II)-bound MDF (B). in DPC micelles. The absorbance was measured at 300 nm. The single-species fitting suggests a tetramer conformation for both apo and bound forms The Top of each panel shows the residuals of the fit. ....   | 132 |
| Figure 5.4 (A) Co(II) spectrum of MDF, (B) Titration of MDF with Co(II) .....  | 133 |

Figure 5.5 (A) Oxidation of Fe (II) in presence of MDF, (B) The time course of the intensity increase at 320nm was fit based on Scheme 2. .... 135

# **Chapter 1: Introduction: The fundamentals of protein folding and the model of helix association**

## **1.1 The significance of protein folding**

Proteins carry out vital functions in every living creature.<sup>1</sup> The proper folding of proteins into their specific three-dimensional structures is the only way that the correct function can be achieved.<sup>1</sup> The ability to interpret the relationship between the primary sequence of protein and its folding will not only help understand the pathology of protein misfolding-related diseases such as Alzheimer's,<sup>2</sup> cystic fibrosis,<sup>3</sup> and prions disease;<sup>4</sup> but also create new functions such as novel enzymes, biomaterials and delivery agents. After decades of effort, protein folding is subject to the attentions of multiple disciplines, such as biology, chemistry, physics, mathematics, and engineering; and is one of the most fundamental processes studied by science.

While the mechanism of protein folding is complicated, the final folded status, including secondary, tertiary and quaternary structure, is simply determined by the primary sequence of a protein.<sup>5</sup> Nevertheless, the process that translates from sequence to the final folded structure is still mysterious. One strategy to investigate this mystery is to identify the sequence motif that determines specific interactions. With the golden rules identified, we can predict unknown protein structures/functions, also design novel protein-based functions, which are two major goals in the field of protein folding research.

## 1.2 The impediments to studying protein folding

The major challenge in the field of protein folding originates from the complexity of the systems that proteins are involved in. The major contributors to this complexity would be folding hierarchy, environmental factors, and folding dynamics.

Folding takes multiple steps. It has been proposed that the sequence first searches into a specific secondary structures, such as alpha-helix, beta-sheet, etc.<sup>6</sup> Then the single chain further forms tertiary structures driven by forces such as local hydrophobic interaction, hydrogen-bond formation and ligand binding.<sup>7</sup> Sometimes in order to obtain a more advanced function, a group of tertiary structures need to interact with each other and further form quaternary structures.<sup>7</sup>

Secondly, the sequence of amino acids alone usually does not completely determine the folding pathway, although it determines the final fold. Environmental factors such as pH,<sup>8</sup> ligand binding,<sup>9</sup> thermal changes,<sup>10</sup> energy transfer,<sup>11</sup> etc., frequently play key roles which triggers/assists the folding and also add to the variables in the equation of folding.

A third hurdle to understanding protein folding comes from the dynamics of the folding process itself. Regardless of reaching equilibrium or not, the system usually

involves transitions among multiple components in a dynamic way.<sup>12</sup> Taking the dynamics into consideration, the system becomes even more complex and mysterious.

Facing the complexity of the protein folding systems, a simplified model is called for, which should allow us to study the key problems in a detailed manner in a manageable system and also serve as a scaffold to design more advanced systems.

### **1.3 The helix-association-based model for studying protein folding**

Helix association is the major part of the model of peptide interaction. Statistics suggests that the predominant structural motif in proteins is alpha-helix in both aqueous and membrane systems.<sup>7</sup> The association of helices is a major part of protein folding and is thermodynamically linked to protein folding. Helix association has been widely used as a model to study protein folding because helix association sufficiently represents the key steps for protein folding in most cases and it also significantly simplifies the system.

In terms of sufficient representatives, successful examples of helix association have been developed in the field of disease mechanism exploration, such as the six-helix coiled coil in HIV gp41,<sup>13</sup> the tetramer M2 proton transporter in influenza virus,<sup>14</sup> and the heterotetramer in SNARE complex.<sup>15</sup> Most importantly, inhibitors designed upon these

segment has successfully blocked folding,<sup>16</sup> supporting their uses as sufficient representatives.

Helix association model is also simplified by the relatively short length of the studied peptides. Compared to proteins with more than a few hundred residues, a 20-50 residue peptide (helix) is often relatively easier in its synthesis, purification and assay development, which allows more advanced experimental study. This short length also saves cost for computational modeling/simulation, both in terms of coding time and calculation circles. Altogether, peptides (helix) have been considered as a practical model for studying protein folding.

#### **1.4 The coiled-coil model**

Coiled-coil helix interactions are one of the most studied types of helix interactions. Coiled coils represent a ubiquitous and important structural motif in helices and also in proteins, with 10% estimated population in proteins.<sup>17</sup>

In coiled coils two to seven alpha helices coil with each other into a strand of rope.<sup>18</sup> Usually this super-twist is left-handed. Because of the geometry of the left-handed coiled coil, the orientation of the amino acids repeats every seven residues with respect to the protein interface. That is why coiled coils are characterized by a so called heptad

repeat in their primary sequences  $(abcdefg)_n$ ,<sup>19, 20</sup> in which the **a** and **d** positions are presented at the interface and play essential roles for bringing helix association (Figure 1).

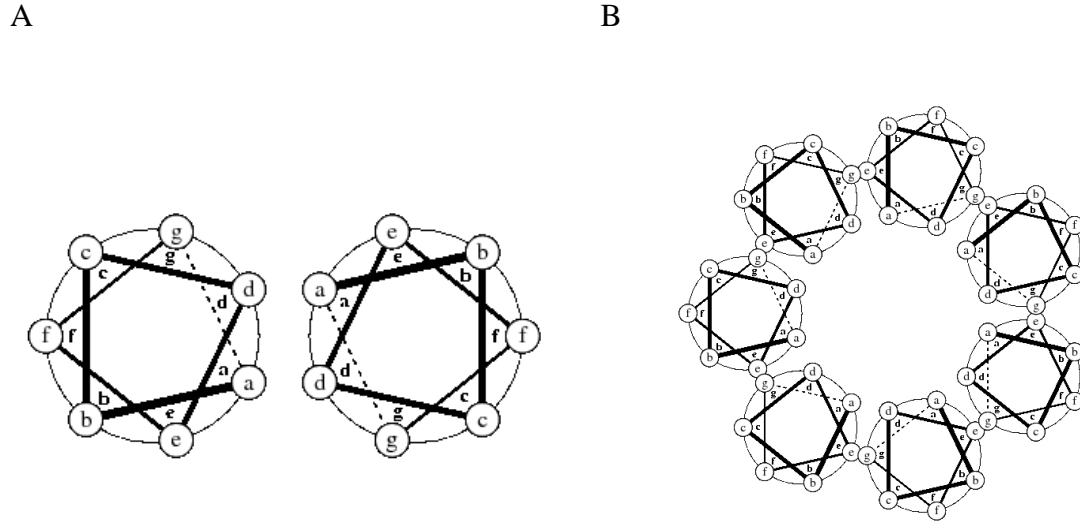


Figure 1.1 A schematic coiled coil heptad repeat for a parallel dimer (A) and a parallel heptamer (B) The schematic map is made via program DrawCoil 1.0, developed by Dr. Gevorg Grigoryan (<http://www.gevorggrigoryan.com/drawcoil/>).

One of the most studied left-handed coiled coils is the water-soluble parallel dimer GCN4-p1, the coiled-coil domain of the yeast transcription activator GCN4.<sup>21</sup> Mutations in GCN4-p1 can derive coiled coil variants that are trimers or tetramers,<sup>22, 23</sup> and parallel or anti-parallel,<sup>24</sup> which will be discussed later in the introduction. Recently a *de novo* designed left-handed parallel seven-helix coiled coil has been characterized with the 1.25 Å crystal structure, which reveals a large tubular channel in the interior.<sup>25</sup>



Like left handed coiled-coils, right-handed coiled coils have also been discussed in both natural and designed system, such as a parallel tetramer with 11-residue repeats from the surface layer of the protein tetrabrachion<sup>26</sup> and a parallel tetramer with 15-residue repeats from human vasodilator-stimulated phosphoprotein (VASP).<sup>27</sup> These right-handed coiled coils represent the versatility of the coiled coil and play important roles in biological functions and design purposes.

Coiled coils have been extensively studied in the past few decades. The accumulated information and regularity of from coiled coil make them a preferred model system for study/design of helix association. In the following part of the introduction, case studies will illustrate the application of the coiled-coil model in research of both aqueous (water-soluble) system and membrane system.

### **1.5 Case study of GCN4-p1: the folding in an aqueous system**

GCN4-p1 is a parallel, left-handed water-soluble homodimer. Crystal structure suggests that GCN4-p1 adopts a coiled-coil conformation, with the side chains in the interface packing into the classical “knob-into-holes” model.<sup>28</sup> As seen in the sequence of the primary coiled coil region, four leucines are aligned as a twisted ladder along the **a** position of the helix, forming the canonical “leucine zipper” and promoting dimerization (Figure 1.2). This leucine zipper echoes nicely with the previous research that hydrophobic interaction is the driving force for water-soluble protein folding.<sup>29, 30</sup>

A. Peptide abcdefg abcdefg abcdefg abcdefg abcdefg  
 GCN4-P1 MKQLEDK VEELLSK NYHLENE VARLKKL VGER

B.

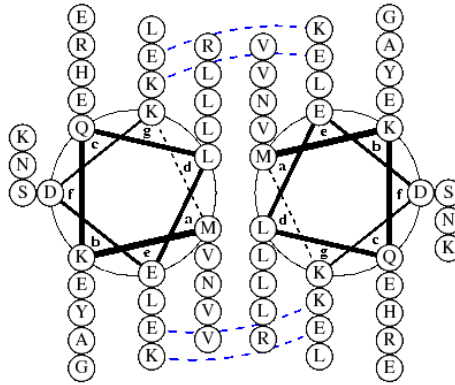


Figure 1.2 The sequence and sketch map of heptad repeat for GCN4-p1. The schematic map is made via program DrawCoil 1.0

Not limited in GCN4-p1, on average 80% of the **a** and **d** position in natural water-soluble coiled coils are occupied by apolar residues.<sup>31</sup> Also, inserting unnatural amino acids with stronger hydrophobicity, such as trifluoroleucine or trifluorovaline, leads to higher stability for coiled coils.<sup>32, 33</sup>

Hydrophobic interactions not only affect coiled coils' stability, but also determine their specificity. A series of mutational studies have been done with GCN4-p1. The coiled coil stays as a dimer with "I" and "L" in **a** and **d** positions respectively; it becomes a trimer with "I" in both **a** and **d** positions; and it is turned into tetramer with "L" and "I" in **a** and **d** positions respectively.<sup>22, 23</sup>

Hydrogen bonding is another major force driving the folding of coiled coils in aqueous systems. The polar residues in the core positions play a key role in determining specificity, supported by the fact that polar residues comprise 20% population of all the residues in the **a** and **d** positions of natural coiled coils.<sup>31</sup> If the asparagine in the **a** position in GCN4-p1 is altered into a valine, the stability of the coiled increases due to the increase in hydrophobic packing, however, the predominance of dimeric form disappears.<sup>34</sup>

Ionic interaction between the **e** and **g'**, as well as **e'** and **g** (Figure 1.2B), also plays important roles in determining the stability and specificity for left-handed coiled coils, which are defined as the canonical *i* to *i*+5 salt bridge in heptad-repeat systems.<sup>20</sup> As shown in Figure 1.2, in GCN4-p1 the glutamic acid and lysine, in the **e** and **g'**, as well as **e'** and **g** positions respectively, form a favorable salt bridge and stabilize the homodimer. The formation of a salt bridge highly depends on sequence, space, geometric preferences and environment factors.<sup>35</sup> For example, the salt bridge contributes about 1 to 2 kJ/mol to coiled coil stability, depending on salt concentration.<sup>36-39</sup>

Interestingly, replacing the charged residue in **e** or **g** positions with apolar residues results in an anti-parallel tetramer,<sup>40</sup> while replacing both the charged residues in **e** and **g** positions with the small apolar residue alanine results in a parallel seven helix-coiled coil.<sup>25</sup> This discloses the importance of packing at the **e** and **g** positions for water-

soluble coiled coils, which apparently play very interesting roles in determining the specificity of coiled coil and deserves further exploration.

Specificity regarding oligomerization has been discussed above. Many investigations into the specificity of interaction orientation have also been done in the past few decades. Since the majority of coiled coils appear to form parallel relative helix alignments in the earliest studies, the design of anti-parallel coiled coils or helix bundles has become an interesting current topic.<sup>24</sup> The first reported *de novo* designed anti-parallel coiled coil was restrained by an interior disulfide bond.<sup>41, 42</sup> A major advance was when glutamic acid and lysine were engineered to interact only when the helices adopt an anti-parallel orientation in a heterodimer,<sup>28</sup> and an asparagine pair in the adjacent helices was engineered to interact only when the helices adopt anti-parallel manner in homodimer.<sup>43</sup>

The successful design of anti-parallel trimer coiled coils<sup>44, 45</sup> and anti-parallel three-helix bundle<sup>46</sup> are mainly based on the modification of the homotrimeric parallel coiled coil by the manipulating Coulomb interactions from charged residues. Also, tryptophan residues have been proposed to contribute to the formation of an antiparallel trimer since three indole sidechains would be poorly accommodated in a parallel coil.<sup>44, 47</sup> Moreover, effort has been made in modifying interactions between **b** and **e** position, as

well as **g** and **c** positions, as these interactions significantly contribute to the stabilization of anti-parallel four stranded coiled coil.<sup>48</sup>

## **1.6 The application of coiled coil-based design in aqueous systems**

In recent decades, coiled-coil-based design has succeeded in diverse fields, thanks to the gradually accumulated knowledge of sequence-determined stability and specificity. Here I will highlight a few examples in aqueous systems.

Biofuel cell cathodes have been designed based on a coiled coil dimer, taking advantage of a coiled-coil's physical self-assembly functionality.<sup>49</sup> An antibody screening assay against HIV-1 has been designed based on the coiled-coil region of the transient viral entry intermediate (the prehairpin intermediate) formed by the HIV-1 gp41 protein.<sup>50</sup> A vaccine has been designed for prototypical Severe Acute Respiratory Syndrome (SARS), the design of which is based on the C-terminal coiled coil of the virus' spike protein.<sup>51</sup> A novel protein fiber has been designed based on a leucine zipper coiled coil. This self-assembled coiled coil can super assemble into a fiber that measure about 45 nm across<sup>52</sup> and could serve as a perfect scaffold for future tissue engineering.

Computational calculations, modeling and simulation have boomed in the field of rational design. Nowadays scientists not only use computational programming to pack the

designed structure in order to achieve the sequence with the most stable target structure,<sup>53</sup>  
<sup>54</sup> but also utilize negative design to improve final specificity.<sup>55</sup>

Undoubtedly, coiled coils provide efficient scaffolds for future design and present great potential for diverse applications.

### **1.7 Understanding membrane protein folding with coiled-coil-based design**

As remarkable as the contribution the coiled-coil-based model has made to the aqueous system, it also has had great importance in the research of membrane protein systems.

It is predicated that about 30% of all the genes encode for membrane proteins and 60% of drugs in current market target membrane proteins.<sup>56</sup> However, only 1% of structures in PDB database correspond to membrane proteins. The lack of knowledge about membrane proteins is mainly because of the difficulties associated with the relevant studies. From an experimental point of view, first of all most membrane proteins are hard to express, synthesize and purify compared to water-soluble proteins due to their high hydrophobicity. Secondly, membrane proteins have to be solubilized in the lipid bilayer or detergent during studies, thus some established assays in water-soluble systems can't be applied to the membrane target. Moreover, it is hard to find an assay which can create

a detectable thermodynamically reversible equilibrium for membrane protein folding. Computational modeling requires simulation of the lipid or detergent environment, which creates challenges in molecular dynamics simulations.

Despite the difficulties of working with membrane proteins, significant progress has been made in the membrane field in the past decade. There are two major classes of structural motifs in membrane proteins: all-beta sheets and all-alpha helices. The proteins in cytoplasmic membranes are essentially all in the all-alpha class, whereas the outer membranes of bacteria and mitochondria also are rich in all-beta. However, most pharmacologically relevant membrane proteins are all-alpha helices therefore the majority of work on membrane protein has focused on the transmembrane (TM) helix.<sup>57</sup>

In a well-accepted model, membrane helix association involves two kinetically separate steps: In the first step the TM helical region is inserted into lipid bilayer; in the second step the TM helices interact with each other and form bundles<sup>57, 58</sup>. This thesis will focus on the second step and explore the folding mechanism for membrane helix association and relative designs.

Where the issue of folding is concerned, it is always required to identify an unfolded state and a folded state in the beginning of study. In our helix-association model,

the unfolded state would be the single helix which retains the helical secondary structure; and the folded state is the associated helix bundle.<sup>59</sup> Mirroring the trajectory of water-soluble systems, the coiled coil model has been widely used in understanding the driving force for membrane helix association and helix-association-based membrane protein design. This thesis will highlight the successful design and characterization of the model peptide MS1, a putative membrane coiled coil.

### **1.8 Case study of MS1: folding in micelles and bilayers**

MS1 is designed based on the classic water-soluble coiled coil GCN4-p1. The design maintains the core residues in **a** and **d** positions of GCN4-1:<sup>60</sup> “V” “N” “V” “V” in consecutive putative **a** positions and “L” in all the putative **d** positions (Figure 1.3). Fluorescence and sedimentation analytical ultracentrifugation assays suggest that MS1 adopts an equilibrium of a monomer-dimer-trimer in detergent micelle N-tetradecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (C-14 betaine). The maintained asparagine has been shown to play a crucial role in the self-association of MS1: the mutation of asparagine into valine (N14V) eliminates the ability to oligomerize.<sup>60</sup> Mutation of asparagine into other polar residues with two polar atoms such as glutamine, aspartic acid and glutamic acid, retains the ability to form stable trimers, while the mutation of asparagine into polar residues with one polar atom such as serine and threonine, results in a much weaker tendency to associate.<sup>61</sup> It is proposed that the inter-molecular hydrogen-



bond network formed by the asparagine is the major driving forces for the helix-trimerization in MS1.<sup>62</sup>

With MS1, the role of hydrophobic effect in membrane helix association has also been probed. The mutation L10A (Figure 1.3) probes how hydrophobicity in putative **d** positions affects that association of membrane helices. Interestingly this alteration does not change the oligomerization state. Moreover, the variant forms significantly tighter trimers than the wild type (by -0.5 kCal/mol monomer), which suggests that hydrophobic interactions are not the major driving forces for membrane helix association. Instead, small residues are preferred in the helical interface<sup>60, 63</sup>, which will be further discussed in Chapter 3.

| <b>Peptide</b> | <b>abcdefg</b> | <b>abcdefg</b> | <b>abcdefg</b> | <b>abcdefg</b> | <b>abcdefg</b> |
|----------------|----------------|----------------|----------------|----------------|----------------|
| <b>GCN4-P1</b> | <b>MKQLEDK</b> | <b>VEELLSK</b> | <b>NYHLENE</b> | <b>VARLKKL</b> | <b>VGER</b>    |
| <b>MS1</b>     | <b>BQLLIA</b>  | <b>VLLLIIV</b> | <b>NLILLIA</b> | <b>VARLRYL</b> | <b>VG</b>      |
| <b>N14V</b>    | <b>BQLLIA</b>  | <b>VLLLIIV</b> | <b>VLILLIA</b> | <b>VARLRYL</b> | <b>VG</b>      |
| <b>L10A</b>    | <b>BQLLIA</b>  | <b>VLLAIIV</b> | <b>NLILLIA</b> | <b>VARLRYL</b> | <b>VG</b>      |

Figure 1.3 The sequence of MS1 and variants

A parallel project carried out at the same time also emphasizes the importance of polar residue in membrane helix association. The TM part of this parallel design is also based on GCN-p1, which maintains the residues in the putative **a** position (VNVV) and replaced the rest of the residues with leucine. Thermodynamic studies with this design suggest that helix association is driven by the polar asparagines at the putative **a** position, independent of the rest of the hydrophobic leucine sequence in both detergent micelles and biological membranes.<sup>64</sup>

The importance of polar residues in membrane protein folding also echoes the fact that mutations of the polar residues in membrane proteins are usually the most common disease initiators for membrane protein related diseases.<sup>65, 66</sup> Interestingly, this important driving force in some cases is relatively modest. Using bacteriorhodopsin as the target systems, double-mutation cycle analysis suggests that the average contribution of the eight hydrogen-bonding interactions in bacteriorhodopsin is only about 0.6 kcal/mol. The authors proposes that this important while modest hydrogen-bonding might result from the geometry constrains or evolutionary pressures.<sup>67, 68</sup>

## **1.9 Membrane structural motifs**

As the exploration of membrane protein extends, structural motifs/rules about membrane protein folding have gradually been revealed. Discovered in glycophorin A,<sup>69-</sup><sup>74</sup> the GxxxG motif has been shown to drive membrane helix dimerization by promoting van der Waals interaction with an ideal geometric fit and weak C $\alpha$ -H hydrogen-bonding.

Extensive mutagenesis studies suggest that the motif LIxxGVxxGVxxT is crucial for the dimerization of glycoporphin A.<sup>69, 70</sup> Further statistics suggests that this role of glycine can also be extended to other small residues such as alanine and serine. Therefore this motif has been updated to the [Sm]xxx[Sm] motif that drives TM helix association, where “[Sm]” here refers to small apolar residues.<sup>74</sup>

As more crystallographic structures of membrane proteins are deposited to the PDB library, the analysis of the membrane helix interactions have been categorized into two groups.<sup>75</sup> For the right-handed membrane helix interactions, GxxxG or [Sm]xxx[Sm] is the most popular structural motif. For the left-handed membrane helix interactions, small residues, such as glycine or serine, repeating in every seven residues, are commonly seen in the packing interface. Heptad repeats of serine, which is called a serine zipper, have been used to design anion channel peptide<sup>76</sup> and an anti-parallel membrane helix bundle.<sup>77</sup> Also, hydrogen bonds via polar residues are frequently found to mediate the association in left-handed membrane helices.<sup>78</sup>

### **1.10 The application of sequence-based protein folding studies**

The ultimate goal of sequence-based folding studies is to predict the protein structure based on the primary sequence, and to design protein sequence which can obtain specific functions.

The accumulated understanding of membrane protein folding provides support for structural prediction of membrane proteins.<sup>79-81</sup> The knowledge of first principles and prediction has also promoted the rational design of functional membrane proteins, such as activator<sup>82</sup> for platelet integrins, a drug delivery agent,<sup>83</sup> and a diporphyrin-binding TM electron transfer protein.<sup>84</sup> In both prediction and design, computational calculation and modeling have played significant roles. Adding explicit terms and modeling realistic membrane phases has significantly increased the success rates of the designs.<sup>80,84</sup> More detailed examples will be discussed in the following chapters.

### **1.11 The exploration of helix association in both aqueous and membrane systems**

This thesis focuses on the application of helix association in both fundamental and applied research, such as: discovering the role of helix interaction in diseases, especially viral fusion/entry (Chapter 2), exploring the driving forces for membrane helix association (Chapter 3), designing a pH-switchable helix-association based drug delivery agent (Chapter 4), and designing a membrane metal-binding four-helix bundle (Chapter 5).

## **Chapter 2 Transmembrane Orientation and Possible Role of the Fusogenic Peptide from PIV5 Virus in Promoting Fusion**

### **2.1 Abstract**

Membrane fusion is required for diverse biological functions ranging from viral infection to neurotransmitter release. Fusogenic proteins increase the intrinsically slow rate of fusion by coupling energetically downhill conformational changes of the protein to kinetically unfavorable fusion of the membrane-phospholipid bilayers. Class I viral fusogenic proteins have an N-terminal hydrophobic fusion peptide (FP) domain, important for interaction with the target membrane, plus a C-terminal transmembrane (C-term-TM) helical membrane anchor. The role of the water-soluble regions of fusogenic proteins has been extensively studied, but the contributions of the membrane-interacting FP and C-term-TM peptides are less well characterized. Typically, FPs are thought to bind to membranes at an angle that allows helix penetration but not traversal of the lipid bilayer. Here we show that the FP from the paramyxovirus PIV5 fusogenic protein, F, forms an N-terminal TM helix, which self-associates into a hexameric bundle. This FP also interacts strongly with the C-term-TM helix. Thus, the fusogenic F protein resembles SNARE proteins involved in vesicle fusion by having water-soluble coiled-coils that zipper during fusion and TM helices in *both* membranes. By analogy to mechanosensitive channels, the force associated with zipping of the water-soluble coiled coil domain is expected to lead to tilting of the FP helices, promoting interaction with the C-term-TM helices. The energetically unfavorable dehydration of lipid headgroups of opposing

bilayers is compensated by thermodynamically favorable interactions between the FP and C-term-TM helices as the coiled-coils zipper into the membrane phase, leading to a pore lined by both lipid and protein.

## 2.2 Introduction

The basic mechanisms of viral membrane fusion have been studied extensively, but major gaps remain in our understanding of the relative roles of lipidic intermediates and viral fusogenic proteins in lowering the energy barrier for the overall process<sup>85-88</sup>. The most common mechanistic hypothesis concerning enveloped viral fusion is that fusogenic proteins primarily serve to bring the target cell and viral membranes into proximity. Fusion occurs in a multi-step process, in which the virus first binds to a specific receptor; this event and/or other environmental cues then cause a conformational change in the protein, leading to a metastable state with an exposed hydrophobic fusion peptide (FP) that binds to the target membrane. Once engaged with the bilayer a second energetically favorable conformational change in the fusogenic protein then exerts a force pulling the FP towards the viral membrane, in effect reeling the host and viral membranes together.

The conformational changes involved in the water-soluble portions of viral fusogenic proteins have been largely elucidated, but the roles of the membrane-binding FP and the C-terminal transmembrane (C-term-TM) anchor are less clear. After the crystal structure of the pre-fusogenic form of influenza hemagglutinin (HA) was solved<sup>89</sup>, experimental studies suggested that its FP inserted into the bilayer<sup>90</sup>. The FP helix was thought to bind sufficiently deeply to act as a hydrophobic wedge that not only served as an anchor but also destabilized the bilayer and facilitated fusion. Many biochemical, biophysical and mutagenesis studies on the fusion proteins of the influenza virus, HIV

and other viruses are consistent with, and have added considerable detail to, this initial suggestion<sup>91-118</sup>. However, a number of intriguing findings suggest that the FPs and C-term-TM helices might play additional, more specific, roles in bilayer fusion than mere membrane binding and disruption.

Surprisingly subtle mutations in the C-term-TM sequence of fusogenic proteins can be quite deleterious to their ability to induce fusion, while retaining normal processing and the ability to change conformations<sup>119-122</sup>. Also, replacing C-term-TM helices of fusogenic proteins with lipid anchors results in a loss of fusion<sup>123-126</sup> and, surprisingly, FP sequences often show greater conservation than might be expected from the functional requirements for membrane binding<sup>127</sup>. Moreover, the very strong conservation of polar and small residues at regularly spaced intervals as found in GXXXG<sup>70, 74, 128</sup>, glycine zippers<sup>129</sup>, and GAS motifs<sup>75</sup> (Table 2.1), is intriguing. These patterns are known to stabilize TM helix association and also figure in the helix-helix packing of proteins that undergo large-scale conformational changes in response to lipidic environment, such as in mechanosensitive (MS) channels<sup>130</sup>.

Thus, we investigated whether FPs might adopt TM helical rather than surface orientations, and how this might relate to their mechanisms of action. If indeed FPs adopt a TM orientation, then one might envision a mechanism akin to SNARE proteins<sup>131-134</sup>, in which both the target and vesicular proteins have TM helices that associate as



membrane fusion progresses. Accordingly, we investigate herein a class I fusogenic protein, F, from parainfluenza virus 5 (PIV5), which, like influenza virus HA and HIV gp41, has an N-terminal FP in the mature, cleaved protein<sup>87</sup>. High-resolution structures are available for both the pre- and post-fusion forms of the ectodomain of the F protein<sup>135-137</sup>, and the structure and function of its C-term-TM domain has been extensively investigated by scanning mutagenesis and Cys crosslinking<sup>138</sup>. The sequence of the FP suggests that it also forms a TM helix<sup>56</sup>. Moreover, it has a single polar Gln residue, a residue known to promote helix-helix interactions in membranes<sup>61</sup>, plus glycine and alanine residues in a heptad repeat pattern (Figure 2.1B), known to stabilize TM helix-helix assembly<sup>59, 78, 139</sup> and pore formation<sup>129</sup>. The sequence variability of the FP across homologous viruses shows a heptad repeat in phase with the heptad repeat of the long water-soluble coiled-coil, which directly follows it (Figure 2.1C, 1D), interrupted only by a highly conserved region (residues 112-117) that is constrained by packing in the pre-fusion trimer<sup>136, 140</sup>. Consistent with this, the postfusion structure of PIV5<sup>135</sup> showed that the C-terminus of the fusion peptide is helical. Here, we show that this FP adopts a TM orientation in phospholipid membranes, specifically oligomerizing into a homohexameric bundle (6HB) and it also associates with the C-term-TM domain in micelles. Computational studies suggest that conformational changes involving zippering of the water-soluble coiled-coil in the ectodomain drive changes in helix-crossing angles that may lead to an initial heteromeric contact or “pin-prick” between the FP and the C-term-TM leading to a fusion pore possibly lined by both protein and lipid.

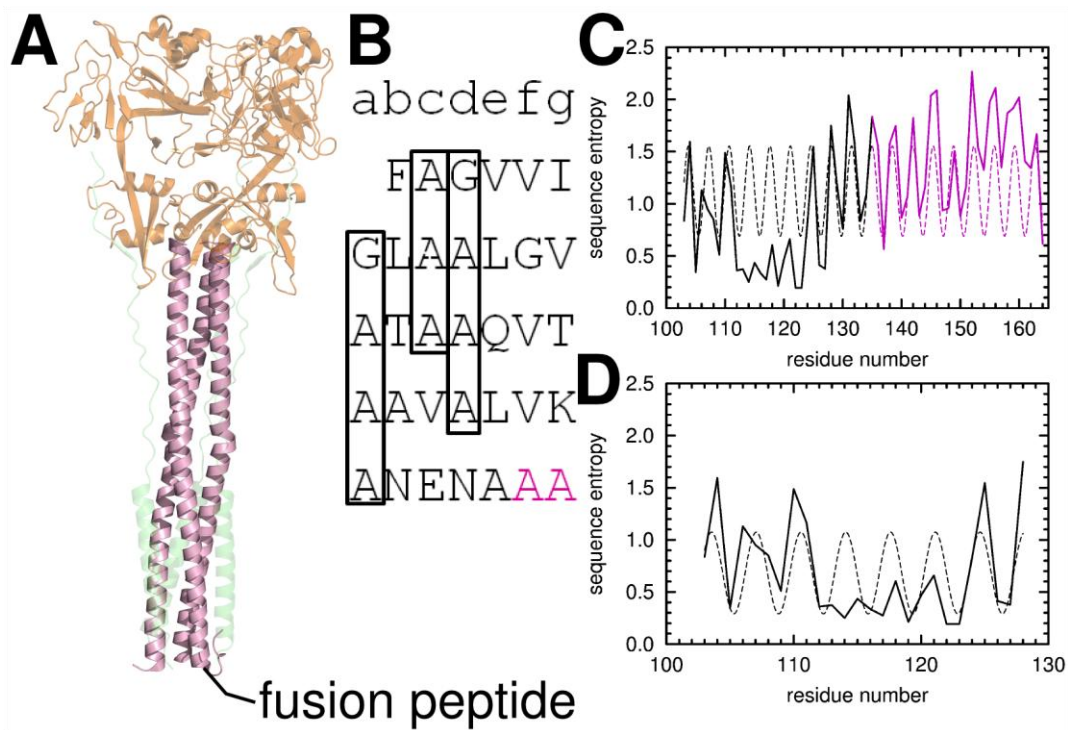


Figure 2.1 Sequence conservation suggests a continuous helix including heptad repeat A (HRA) and the fusion peptide. (A) Postfusion crystal structure of the soluble domain of closely related hPIV3 virus F protein<sup>135</sup>. Shown in magenta is HRA. Below HRA, in the postfusion membrane, is the predicted location of the fusion peptide. (B) Heptad repeat of the fusion peptide and HRA. The beginning of the crystallographic resolved region of HRA is shown in magenta. Heptad repeats of small residues in the fusion peptide are boxed. (C) Sequence entropy of the fusion peptide and HRA can be fit to a single sinusoidal function with period of  $3.47 \pm 0.02$  residues/turn ( $r = 0.51$ ). (D) Sequence entropy of the fusion peptide alone can be fit to a single sinusoidal function with a period of  $3.51 \pm 0.08$  residues/turn ( $r = 0.59$ ).

Table 2.1 Conservation of small residues (Ala, Gly) within four virus families.

Each family uses a type I fusion protein with an N-terminal fusion peptide. Small residues are shown in bold. Note that even distantly related viruses often conserve the position of small residues.

**Orthomyxoviridae:**

|                |  |
|----------------|--|
| Influenza A H1 | GLF <b>GAIAG</b> FI <b>EGG</b> WT <b>GM</b> IDGWY <b>G</b>                                 |
| Influenza A H3 | GLF <b>GAIAG</b> FI <b>ENG</b> W <b>EG</b> MIDGWY <b>G</b>                                 |
| Influenza A H5 | GLF <b>GAIAG</b> FI <b>EGG</b> W <b>Q</b> GMVDGWY <b>G</b>                                 |
| Influenza B    | <b>GFFGAIAG</b> F <b>LEGG</b> W <b>EGMIAG</b> WH <b>G</b>                                  |
| Influenza C    | IF <b>G</b> IDD <b>LII</b> <b>G</b> VLFV <b>A</b> IV <b>ETG</b> IG <b>GG</b> YLL <b>GS</b> |

**Retroviridae:**

|                         |   |
|-------------------------|---|
| HIV-1, Group M, Clade A | <b>A</b> IG <b>MGA</b> FF <b>LGFLGA</b> AGSTM <b>GAA</b> SITLTV <b>QA</b> |
| HIV-1, Group M, Clade B | <b>AVG</b> IG <b>ALFLGFLGA</b> AGSTM <b>GAA</b> SMTLTV <b>QA</b>          |
| HIV-1, Group M, Clade C | <b>AVG</b> IG <b>AVFLGFLGA</b> AGSTM <b>GAA</b> SITLTV <b>QV</b>          |
| HIV-1, Group O          | <b>AVG</b> LG <b>MLFLGVLSA</b> AGSTM <b>GAAA</b> TT <b>LAVQT</b>          |
| HIV-1, Group N          | <b>AAFGLGALFLGFLGA</b> AGSTM <b>GAA</b> SITLTV <b>QA</b>                  |
| HIV-2                   | <b>G</b> VF <b>VLGFLGFLA</b> T <b>AGS</b> AM <b>GAA</b> SLT <b>LSAQS</b>  |
| SIV, rhesus monkey      | <b>G</b> VF <b>VLGFLGFLA</b> T <b>AGS</b> AM <b>GAA</b> SLT <b>LLTAQS</b> |
| SIV, chimpanzee         | <b>AAFGLGALFLGFLGA</b> AGSTM <b>GAAA</b> V <b>T</b> LTV <b>QA</b>         |

Human T-cell leukemia virus **AVPVAVWLVSALAMGAGVAGGITGSMSLASG**

**Paramyxoviridae:**

|                              |                                   |
|------------------------------|-----------------------------------|
| Human parainfluenza virus 1  | <b>FFGAVIGTIALGVATAAQITAGIALA</b> |
| Human parainfluenza virus 3  | <b>FFGGVIGTIALGVATSAQITAAVALV</b> |
| Simian parainfluenza virus 5 | <b>FAGVVIGLAALGVATAAQVTAVALV</b>  |
| Measles                      | <b>FAGVVLAGAALGVATAAQITAGIAL</b>  |
| Sendai virus                 | <b>FFGAVIGTIALGVATSAQITAGIALA</b> |
| Nipah virus                  | <b>LAGVIMAGVAIGIATAAQITAGVALY</b> |
| Newcastle virus              | <b>FIGAIIGSVALGVATAAQITAASALI</b> |
| Respiratory syncytial virus  | <b>FLGFLLGVGSAIASGVAVS</b>        |

**Arenaviridae:**

|                              |                               |
|------------------------------|-------------------------------|
| Lassa fever                  | <b>GTFTWTLSDSEGGKDTGGYCLT</b> |
| Lymphocytic choriomeningitis | <b>GTFTWTLSDSSGVENPGGYCLT</b> |
| Junin arenavirus             | <b>AFFSWSLTDSSGKDTGGYCL</b>   |
| Machupo virus                | <b>AFFSWSLTDSSGKDMPGGYCL</b>  |
| Guanarito virus              | <b>AFFSWSLSDPKGNDMPGGYCL</b>  |
| Sabia virus                  | <b>GIFSWTITDAVGNDMPGGYCL</b>  |

## 2.3 Results

### Association of Fusion and C-term-TM Peptides in Detergent Micelles

Analytical ultracentrifugation (AUC) of the FP from PIV5 F protein in phospholipid micelles reveals cooperatively assembly into hexamers (Figure 2.2). The FP was dissolved in dodecylphosphocholine (DPC) micelles, and the density of the solution was adjusted to precisely match that of the DPC detergent<sup>82</sup> so that only the protein component contributes to the sedimentation equilibrium. Three samples prepared at differing peptide-to-detergent ratios were each centrifuged at four to five rotor speeds, respectively, for the wild-type and mutant Q120A. The data were then globally analyzed to extract the number of peptides per oligomer as well as the free energy of association<sup>82, 139, 141</sup>. The data conform very well to tightly associating and fully cooperative monomer-hexamer equilibrium (Figure 2.3). The addition of lower-order intermediate states failed to improve the quality of the fit, indicating that the association was highly cooperative and specific for the formation of hexamers relative to other possible association states.

The mutant Q120A also forms hexamers (Figure 2.2C, 2D, 3B), but its association is weaker than that of the wild-type peptide by 13.4 kcal/mol of hexamer, or 2.2 kcal/mol of monomer. Glutamine (Gln) is well known to stabilize the association of TM helices<sup>61</sup> and the magnitude of the effect is similar. Thus, it is likely that the Gln helps stabilize TM helix association within the structure, although this residue is not

absolutely essential for forming the 6HB. Q120 is strongly conserved in related viruses (Table 2.1) and is a promising target for future studies using reverse-engineered viruses.

While the C-term-TM domain has been shown to associate in the full-length protein <sup>138</sup>, the C-term-TM peptide alone does not associate in DPC micelles (Figure 2.4A). However, when unlabelled wild-type FP is introduced at a 1:1 ratio, the C-term-TM-peptide strongly associates (Figure 2.4B), perhaps adopting a structure relevant for the postfusion state. Analysis of the sedimentation curves indicates that the TM peptide self-associates with the FP at least 20-fold less tightly than the corresponding heteromeric interaction with the C-term-TM peptide.

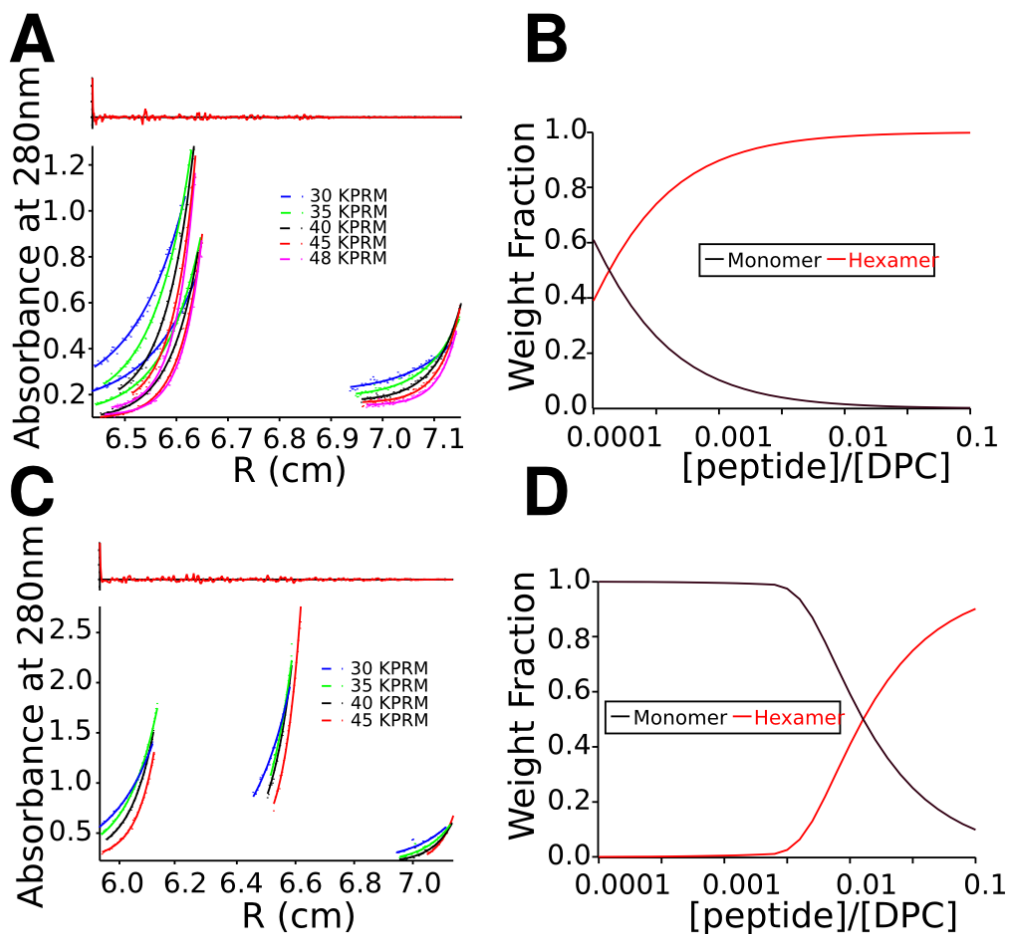


Figure 2.2 Analytical ultracentrifugation of fusion peptide wild-type (A, B) and mutant Q120A (C, D). Single species fitting of wild-type (A) and mutant Q120A (C) PIV5 fusion peptide suggests both associate as hexamers. The top of each panel shows the residuals of single species fitting. The species weight fraction is shown for wild-type (B) and mutant Q120A (D), indicating that hexamer is the dominant species composition for the wild-type, while oligomerization of the mutant Q120A requires significantly higher mole fractions.

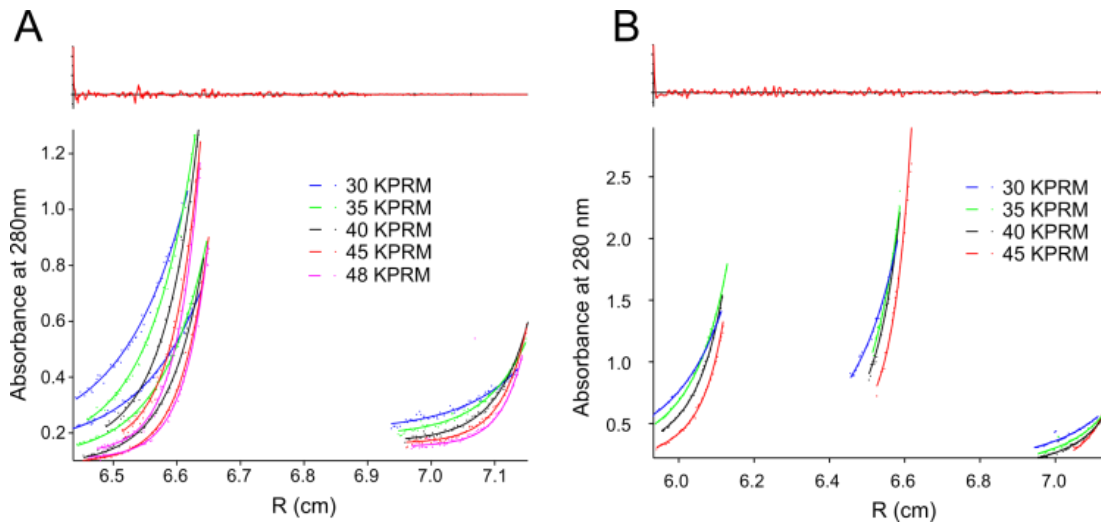


Figure 2.3 Sedimentation equilibrium analysis of the wild-type (A) and mutant Q120A (B) fusion peptides in DPC micelles. The absorbance was measured at 280 nm. The data was fitted as a monomer-hexamer equilibrium, resulting in a  $pK_{\text{dissociation}}$  of 20.1 for the wild-type and a  $pK_{\text{dissociation}}$  of 10.2 for the mutant Q120A. The top of each panel shows the residuals of the fit. In this data analysis the concentration of peptide is the mole ratio of peptide/detergent and therefore is unitless.



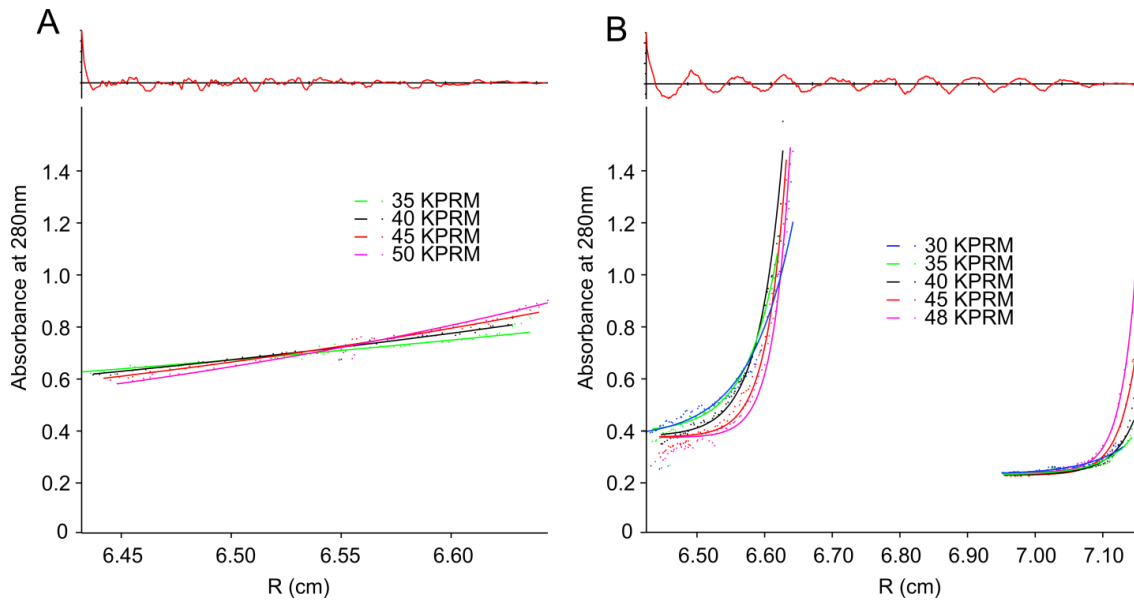


Figure 2.4 Analytical ultracentrifugation of C-term-TM peptide alone (A) and C-term-TM peptide-fusion peptide mixture (B). No significant curvature has been observed in AUC data of TM peptide alone (A) suggesting that C-term-TM peptide does not self-associate under these conditions. Single species fitting for a 1:1 mixture of C-term-TM peptide and fusion peptide (using the original sequence without a Trp label) suggests a strong interaction between these peptides.

## **Fusion Peptides Adopt a TM Orientation in Lipid Bilayers**

The secondary structure and orientation of the wild-type and mutant FP in micelles and deuterium oxide (D<sub>2</sub>O) hydrated bilayers were evaluated using circular dichroism (CD) and attenuated total reflection IR spectroscopy (ATR-IR), respectively. The CD spectra of both peptides in DPC micelles are typical of alpha-helix (Figure 2.5) indicating that the association observed by AUC corresponded to the formation of helical bundles. The IR spectra in the amide I region of the FPs shows a single, sharp peak at 1656 cm<sup>-1</sup>, indicative of a dehydrated helical conformation<sup>142</sup> in bilayers (Figure 2.6 A, B). The dichroic ratio for parallel versus perpendicularly polarized light was 3.2 and 3.6 for the wild-type and mutant, respectively. These values correspond to an orientation of approximately 29° and 22° relative to the membrane normal<sup>143, 144</sup>, assuming the entire peptide is fully helical and the bilayers are well ordered. Deviation from helical geometry or disorder of the bilayer would result in somewhat lower dichroic ratios. In this case, the true angles would be even closer to parallel to the bilayer normal. Thus, both peptides have a strong preference to adopt a TM orientation relative to other possibilities in which the helix was either randomly oriented or oriented parallel to the membrane surface.

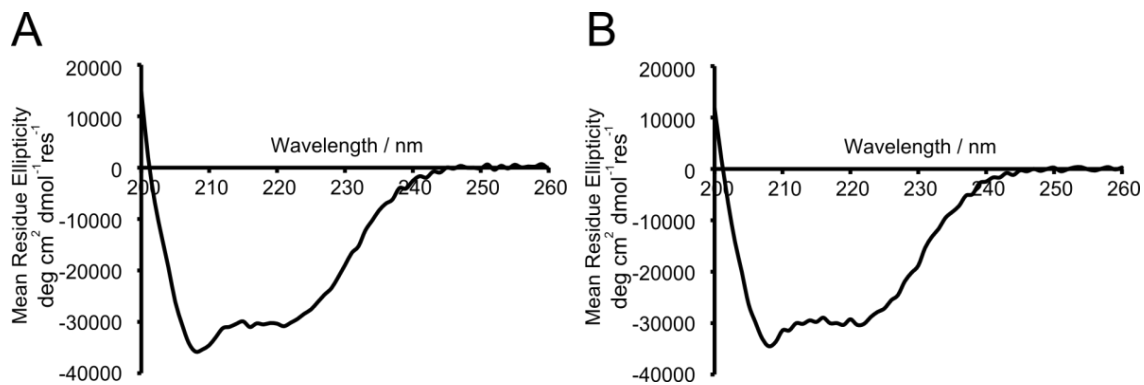


Figure 2.5 CD spectra of fusion peptide wild-type (A) and mutant Q120A (B). The spectra show that both wild-type and Q120A are predominantly  $\alpha$ -helical at a peptide:DPC ratio of 1/200.

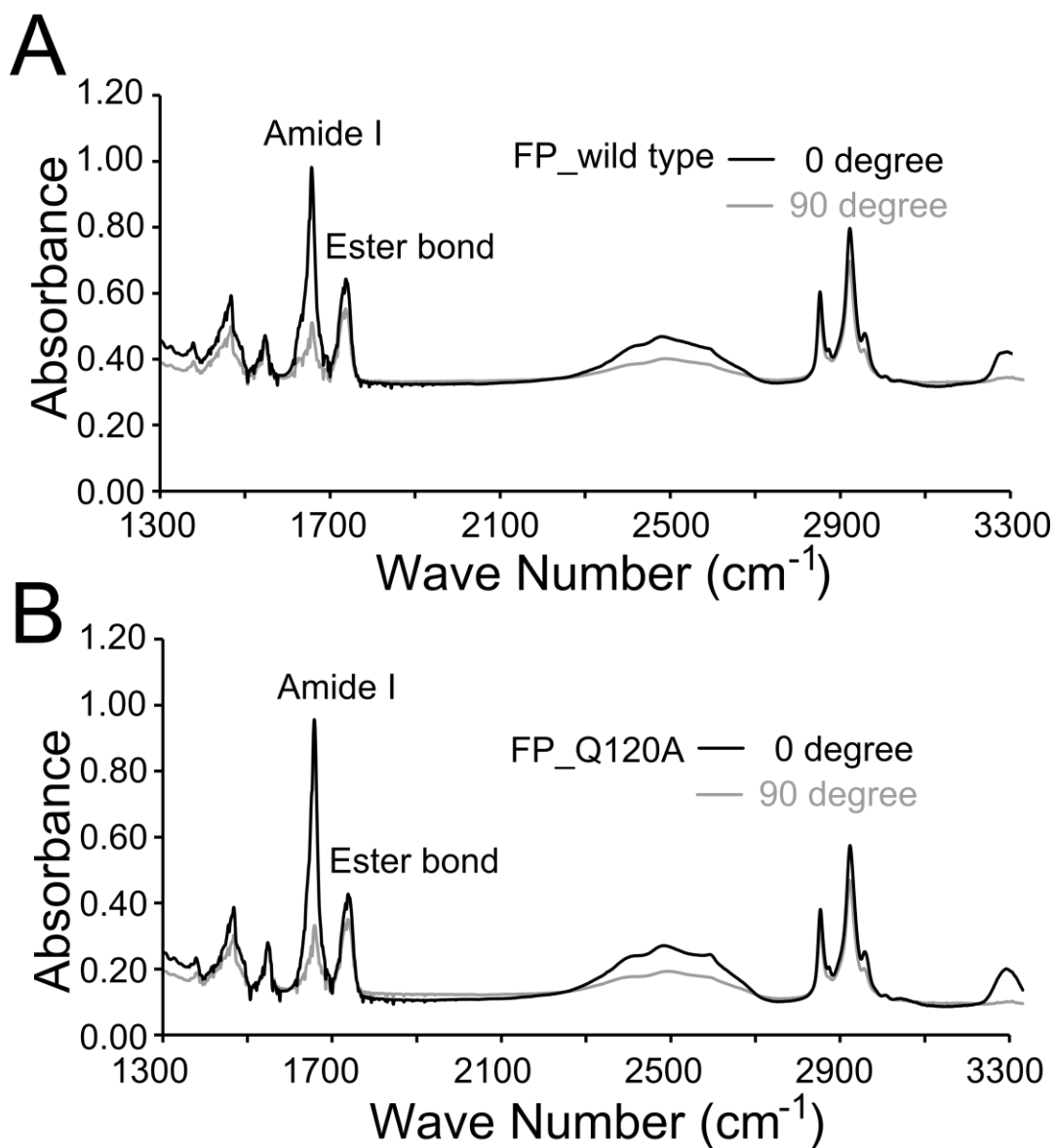


Figure 2.6 Attenuated total reflection IR spectroscopy (ATR-IR) of fusion peptide wild-type (A) and mutant Q120A (B) in phospholipid (POPC) bilayers. The sharp peak at  $1656 \text{ cm}^{-1}$  is indicative of alpha helical secondary structure. The TM orientation is demonstrated by the much greater intensity of the  $1656 \text{ cm}^{-1}$  amide I bond for parallel ( $0^\circ$ ) versus perpendicular ( $90^\circ$ ) polarized incident light (relative to the membrane normal).

## Computational Modeling of the 6HB

To model the FP 6HB, the possible structural space was systematically sampled and scored using a protocol akin to the conformational search of Brunger, Arkin and coworkers<sup>145</sup>. The strong heptad repeat (Figure 2.1C, D) is indicative of a left-handed helical bundle. A left-handed bundle also would be consistent with a continuous helical structure beginning in the soluble heptad repeat A (HRA) and continuing directly to fusion peptide, as the conservation pattern suggests (Figure 2.1C). Moreover, the nature of viral fusion, with asymmetric insertion of peptides into the target membrane, suggests that the FPs comprising the 6HB should adopt a parallel orientation. Symmetric, parallel coiled-coils can be described by a limited number of variables<sup>146</sup>. Three of these:  $\alpha$ -helical phase ( $\varphi$ ), pitch angle ( $\alpha$ ), and superhelical radius ( $R$ ) were allowed to vary and were sampled systematically in search of optimal coiled-coil structures. For each structure, optimal rotamers were selected and the structure then minimized. Each structure was scored using the CHARMM energy function in an implicit membrane environment<sup>147, 148</sup> to select candidate models.

The energy of a particular 6HB conformation depends primarily on the phase,  $\varphi$ . Multiple energy minima are observed as the helices are rotated (Figure 2.7A). Five left-handed structures were selected, corresponding to the lowest energy basins, for further refinement using molecular dynamics (MD) simulations in an explicit fully hydrated lipid bilayer. The lowest predicted energy for an antiparallel orientation 6HB was selected as a

negative control. It was less stable in MD simulations than the low-energy parallel models and was not further pursued (Figure 2.7, 2.8).

MD simulations on the five parallel 6HB structures (labeled according to their phases,  $\phi = 40^\circ, 43^\circ, 88^\circ, 196^\circ,$  and  $300^\circ$ , respectively) show that the orientation of the Gln side chain is crucial for 6HB stability. Two closely related structures,  $\phi = 40^\circ$  and  $43^\circ$  ( $C_\alpha$  RMSD = 1.6 Å), place the Gln in a **d** position within the coiled-coil while the phase  $88^\circ$  structure places the Gln in a **a** position. The remaining two structures have Gln facing the lipid (phase  $196^\circ$  and  $300^\circ$ ), are much less stable than the interior-facing ones (Figure 2.7B, 8), and rapidly depart from their initial structures (as measured by  $C_\alpha$  RMSD), while structures with an interior Gln are stable near the initial structure for 50 ns of MD simulation.

Of the interior facing Gln structures, the phase  $40^\circ$  and  $43^\circ$  models are most stable during the MD simulation and best maintain a symmetric coiled-coil structure (Figure 2.8). These models form a highly stable hydrogen bond network in the interior of the 6HB coiled-coil (Figure 2.9A), consistent with the important role Gln plays in oligomerization (Figure 2.2). The periodically conserved small residues of the FP are found at the helix interface in this model. Of note is the penetration of water into the core of the 6HB from the viral side of the membrane (Figure 2.9B). It is possible that the formation of the FP 6HB structure may reduce the barrier to fusion by initiating

formation of a nucleus for expansion into the later, much larger fusion pore. In the less stable models (phases 88°, 196°, and 300°) the water distribution is not stable due to either a less favorable arrangement of the Gln side chains (88°) or to their location outside the pore (196° and 300°) (Figure 2.10).

The most stable 6HB ( $\phi=40^\circ$ ) structure was then used to compute an FTIR dichroism ratio following the method of Arkin and coworkers<sup>149, 150</sup> where the individual residue dipoles are combined. The computed dichroism ratio is  $2.95 \pm 0.07$  (mean and standard deviation over the MD simulation), in good agreement with the experimental value (Figure 2.6).

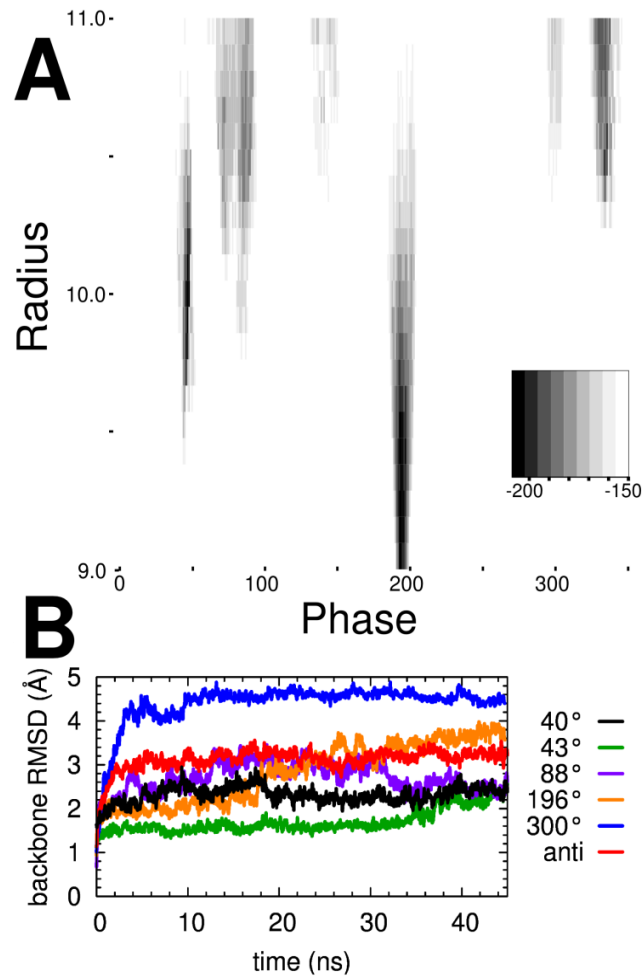


Figure 2.7 Computational prediction of the hexameric helical bundle. (A) Heat map shows predicted energies of coiled-coil models with different radii and phases. Regions in black are predicted to be more energetically favorable. (B) Backbone atom RMSD of five coiled-coil models selected from (A) and one antiparallel model generated using an analogous procedure. Parallel models are labeled according to their phase angle.



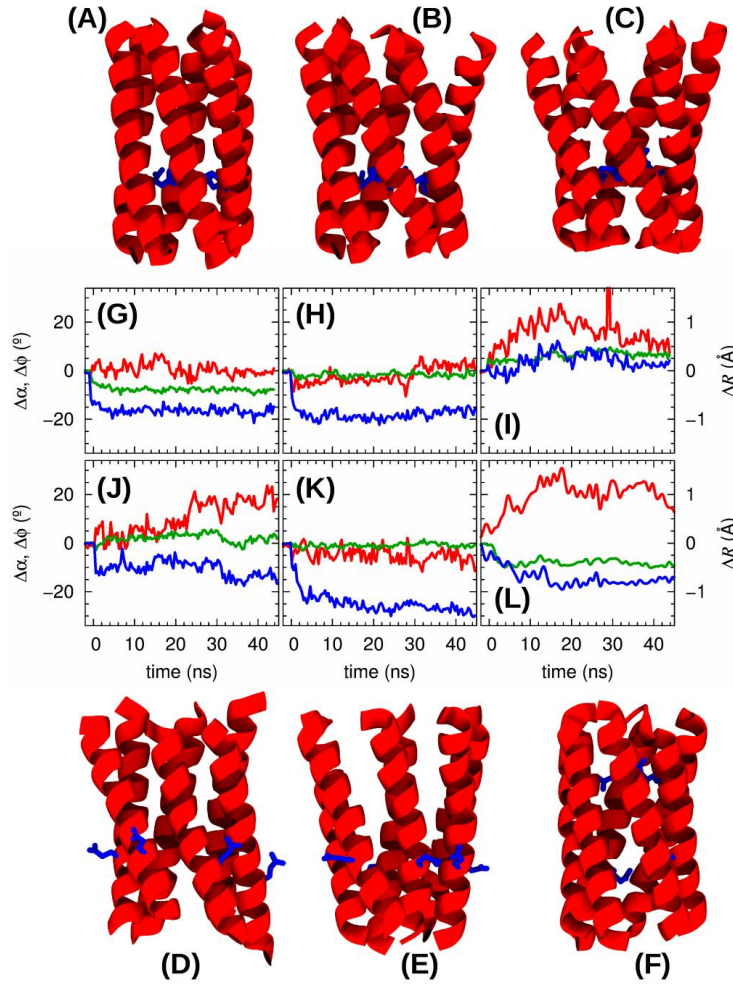


Figure 2.8 Final structures of the MD simulations of different hexamer models and coiled coil parameters as a function of time. (A-F) Final backbone structures in cartoon representations of five parallel models of the hexamers corresponding to the phase angles,  $\varphi$ , of  $40^\circ$  (A),  $43^\circ$  (B),  $88^\circ$  (C),  $192^\circ$  (D),  $300^\circ$  (E), and of the antiparallel model (F). The heavy atoms of the Gln120 side chains are drawn in blue. (G-L) Coiled coil parameters as a function of time of the six models, in the same order as in (A-F); the deviations from the initial values of the pitch angle  $\alpha$  and of the phase angle  $\varphi$  ( $\Delta\alpha$  and  $\Delta\varphi$ ) are plotted in green and red, respectively; the deviation of the radius ( $\Delta r$ ) is plotted in blue.

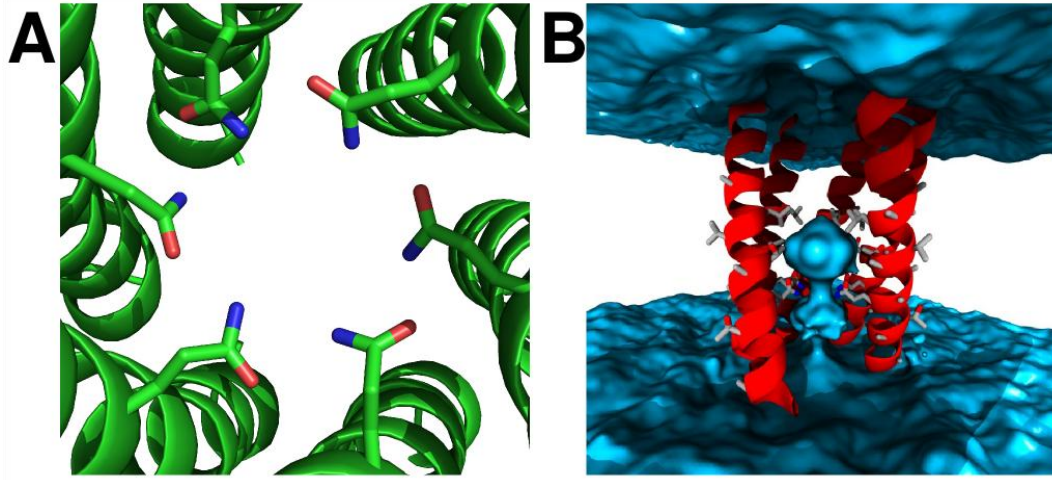


Figure 2.9 Computational model of the PIV5 F fusion peptide hexameric bundle. (A) The Q120 residues form hydrogen bonds with one another as well as waters on the interior. (B) Side view shows the bundle oriented with the N-terminal end (which presumably faces the cellular interior) up. Water is shown in blue. Not shown for clarity are the phospholipids as well as one helix closest to the viewer.

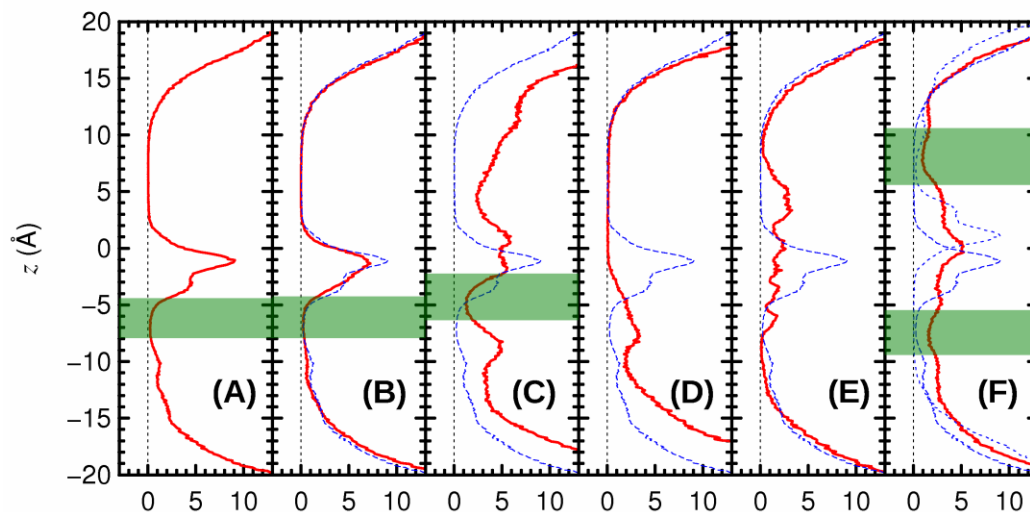


Figure 2.10 Density of oxygen atoms from water molecules through the pore of the hexamer bundle. From left to right, data from the following models are plotted:  $\varphi = 40^\circ$ ,  $43^\circ$ ,  $88^\circ$ ,  $192^\circ$ ,  $300^\circ$ , and the antiparallel model. The density is expressed in number of atoms per  $\text{\AA}$ , and plotted as a red line along the horizontal axis. For comparison, the density profile from the  $\varphi = 40^\circ$  model is also shown as a dashed blue line in the other models (duplicated for the antiparallel model, which is symmetric with respect to the membrane plane). For those models featuring pore-lining Gln120 side chains, the region occupied by their nitrogen and oxygen atoms is highlighted in green.

## 2.4 Conclusion and Discussion

*Comparison of the properties of the FP from PIV5 with other systems:*

Here we provide experimental evidence that the FP from the PIV5 F protein is able to adopt a TM helical conformation when incorporated into lipid bilayers, and that it associates with the C-term-TM helix. Similarly, a FRET assay<sup>151</sup> suggested the C-term-TM domain of influenza HA interacts with its FP, although the orientation of the peptide in the complex was not determined. These findings extend the structural and mechanistic similarity between the PIV5 fusogenic F protein and SNARE proteins to include not only their water-soluble coiled-coil domains, but also their membrane-interactive domains. Recent biochemical and structural studies on SNARE proteins<sup>131</sup> suggest a zippering motion of the water-soluble coiled-coils that continues into the TM domains promoting a heteromeric interaction between the two TM helices to provide part of the driving force for bilayer fusion.

The conformation and TM orientation of the FP from the PIV5 F protein is clearly defined by IR dichroism (Figure 2.6), which showed an average helical tilt of 20° to 30° relative to the membrane normal, and also ruled out the possibility of significant amounts of  $\beta$ -structure. The situation is less clear for other FPs, which often are found to adopt more “oblique-oriented” or “tilted helical conformation”<sup>152</sup>, in which the helix is oriented at 30 to 70° relative to the bilayer normal, either spanning the bilayer or penetrating a single leaflet, depending on the length of the synthetic peptide investigated

<sup>96-98, 109, 116-118</sup>. For example, the N-terminal peptide of gp41 has been reported to adopt a TM <sup>99</sup>, tilted <sup>112-114</sup>, and beta <sup>103, 110-112, 115</sup> conformation in various membrane mimetics. Synthetic versions of the FP from influenza virus HA2 span approximately half of the bilayer width, but as a bent helix <sup>96</sup> or helical hairpin <sup>102</sup> in micelles. However, the hydrophobic region of the FP in the intact virus spans residues up to Arg25, while NMR studies have been conducted with peptides spanning between 20 to 23 residues, with an artificial oligo-Lys tail added to enhance water-solubility. The dynamics and conformational properties of the 20- versus 23-residue peptide differ significantly <sup>102</sup>, as expected for a finely tuned system with multiple low-lying energy wells that are progressively populated during fusion. These distinct structural states, and their sensitivity to small changes in sequence and environment, may be both functionally relevant and reflect the energetic fine-tuning of the landscape and the dynamic nature of fusion.

Within a family, the FPs of viral fusion proteins have highly conserved sequence motifs, such as heptad repeats of small residues, that are similar to those important for association of other oligomeric TM helical bundles <sup>75, 128, 129</sup> (Table 2.1), suggesting that TM helix-helix association might be relevant to fusion. In this regard, it is interesting to compare the avidity of homo- and hetero-oligomer formation for the C-term-TM and FP of the PIV5 F protein. Isolated FPs homo-oligomerize strongly and specifically to a 6HB. The C-term-TM peptide also engages in helix-helix interactions, which have been

experimentally demonstrated using disulfide crosslinking of the full-length protein<sup>138</sup>. However, the present study shows that the C-term-TM helix homo-oligomerizes more weakly than the FP in the absence of the trimeric ectodomain, but associates tightly with the FP (Figure 2.4). The hierarchy of association strengths mimics the assembly process of the ectodomains, in which the weakly associated parallel C-terminal coiled-coil trimer (contiguous with the C-term-TM helix) dissociates and zippers up along the N-terminal coiled-coil (contiguous with the FP) to form the final antiparallel bundle<sup>85-88</sup>. The C-term-TM and FP may likewise zipper as an antiparallel bundle in forming the postfusion state.

*A provisional model for membrane fusion by class I proteins: lipid-centric and “pin-prick” mechanisms.*

In the absence of fusion proteins, the process of bilayer fusion is a physical process with multiple high-energy intermediates<sup>153, 154</sup> corresponding to: 1) diffusion of the membranes together; 2) dehydration of the bilayers as the headgroups of opposing bilayers come into still closer proximity; 3) formation of a lipidic stalk, 4) hemifusion, 5) pore formation and expansion. Viral fusion proteins and SNARE proteins utilize essentially irreversible, energetically favorable conformational transitions to lower the activation energy for membrane fusion<sup>131, 132, 153, 154</sup>. Thus, they are active participants that shape the energy landscape. There are multiple classes of fusogenic proteins, and there is significant variation in the number of fusion proteins per particle, suggesting additional biological requirements, presence of accessory proteins, or lipid compositions

<sup>86</sup>. Here we consider how the class I fusogenic proteins might orchestrate energetic landscape-shaping mechanisms. The present observations provide molecular detail to two limiting hypothetical models, representing extremes in a continuum of kinetic pathways that depend on the protein and experimental variables.

In a “lipid-centric” model of viral fusion, the proteins hold the bilayers in close proximity to promote the progression through lipidic intermediates of fusion (Figure 2.11A). The FP and C-term-TM domains are hypothesized to remain outside of the point of membrane apposition, which is instead made up exclusively of lipids. The ability of the FP to embed deeply into the membrane and engage in favorable C-term-TM to FP interactions provides a mechanism for forcing the two bilayers into close proximity within a very small area, as the coiled-coil domains of pre-associated proteins zipper through the water-soluble regions and extend into the membrane. Favorable FP to C-term-TM interaction provides a continuously downhill process for the protein component, facilitating bilayer-bilayer apposition. Moreover, for systems in which many fusion proteins are required for fusion, the association of the FP in target membranes might bring sufficient fusion proteins near the protein-free zone.

A second model of fusion envisions that fusion proceeds via an initial contact between the TM domains in the two bilayers.. The central point of protein contact can be thought of as a “pin-prick” that expands into a fusion pore. This model is in contrast to

those that propose a gap junction-like pore <sup>107</sup> as only the initial contact is mediated solely by protein domains. The 6HB is hypothesized to be at the center of the contact region between the two membranes. Subsequent pore formation involves the initial protein contact expanding with recruitment of additional lipids with their headgroups facing the growing fusion pore (note the incursion of phospholipids in Figure 2.11B).

To probe the hypothesized mechanism further, we built models of two bilayers in the process of fusion, and asked how the previously defined structural intermediates of the water-soluble and membrane domains of the protein might map onto likely lipid intermediates, lowering the activation energy of the process. Figure 2.11 shows how a hexamer of the PIV5 FPs might serve as a pin-prick to nucleate a pore at three critical points: the extended prehairpin intermediate (left), membrane apposition (center), and postfusion (right). As the conformational change progresses the TM bundles formed by the C-term-TM and FP helices first dock, then coalesce into heteromeric bundles. The initial zone of inter-membrane contact involves favorable protein-protein interactions rather than energetically unfavorable dehydration of the bilayer headgroups, and the fusion of these two helical bundles provides a low-energy pathway to direct fusion of the bilayers, which remain associated with the TM bundles throughout the process.

The latter mechanism provides a rationale for the multiple conformational forms and strong intra-family conservation in the FP sequences, which must associate with



graded affinities in a homomeric as well as heteromeric fashion. It also explains how the addition of various shaped lipids can either promote or inhibit fusion. As the protein conformational change proceeds, the C-term-TM and FP become more tilted (relative to the normal of the initial bilayer). The driving force for tilting includes the zippering of the coiled-coil and favorable heteromeric TM interactions. The recent structure by Rees and coworkers of the MscL MS channel <sup>130</sup> illustrates how mechanical forces from external domains and lipid-specific effects result in changes in helical tilt and channel radius. Changes in the membrane lateral surface pressure profile cause helices comprising the MscL channel to slide relative to each other, increasing their tilt and opening the channel like a diaphragm. In a similar manner, mechanical forces from conformational changes as well as lateral surface pressure effects associated with the lipid composition would couple to the energetics of protein-mediated bilayer fusion.

The highly conserved small glycine and alanine residues, which are found in both MS channels and class I viral FP (Figure 2.1, Table 2.1), are ideally suited for helix sliding because they present relatively smooth interfaces <sup>74, 128</sup>. The channel formed by MscL also expands with these conformational changes, both opening the channel and increasing the surface area available for protein-protein interaction. The hypothetical tilting of the FP and C-term-TM domains would increase the number of residues in contact with the hydrophobic region of the bilayer beyond the length of 20 residues typically seen for TM helices. This longer membrane-suitable region is observed for the

PIV5 C-term-TM and contributes to fusion<sup>138</sup>. It is also observed in other viruses such as influenza<sup>132</sup> and HIV, where shortening the length of C-term-TM helices can halt the fusion process<sup>155, 156</sup>.

Figure 2.11A compares the lipid-only and pin-prick mechanisms; in both cases protein-protein interactions between membrane-embedded helices bring the two bilayers into intimate contact. After the bilayers are brought close together, different proteins might take different pathways to achieve fusion. The zone of adhesion can widen to create a hemifusion intermediate, particularly for situations in which one of the two helices does not fully span the bilayer (Figure 2.11A). Alternatively, the protein might act as a pin-prick to nucleate the fusion pore (Figure 2.11B). The requirements for tight and specific interactions between the membrane-embedded helices will also vary depending on the specifics of the mechanism.

Overall, it seems likely that a continuum of mechanisms is needed, with protein-rich and lipid-rich patches in the fusion pore for many proteins. This in turn will allow for failures leading to lipid mixing-arrested hemifusion when the fusogenic peptides are mutated. The present work favors a protein-centric but not a protein-only fusion mechanism. The pin-prick mechanism should face a less difficult pathway for inter-bilayer interaction to initiate the pore. In addition, it bridges SNARE-like and virus-like mechanisms while explaining why different angles of insertion have been observed for

various fusion peptides. Thus, this mechanism provides a general framework for understanding protein-mediated membrane fusion.

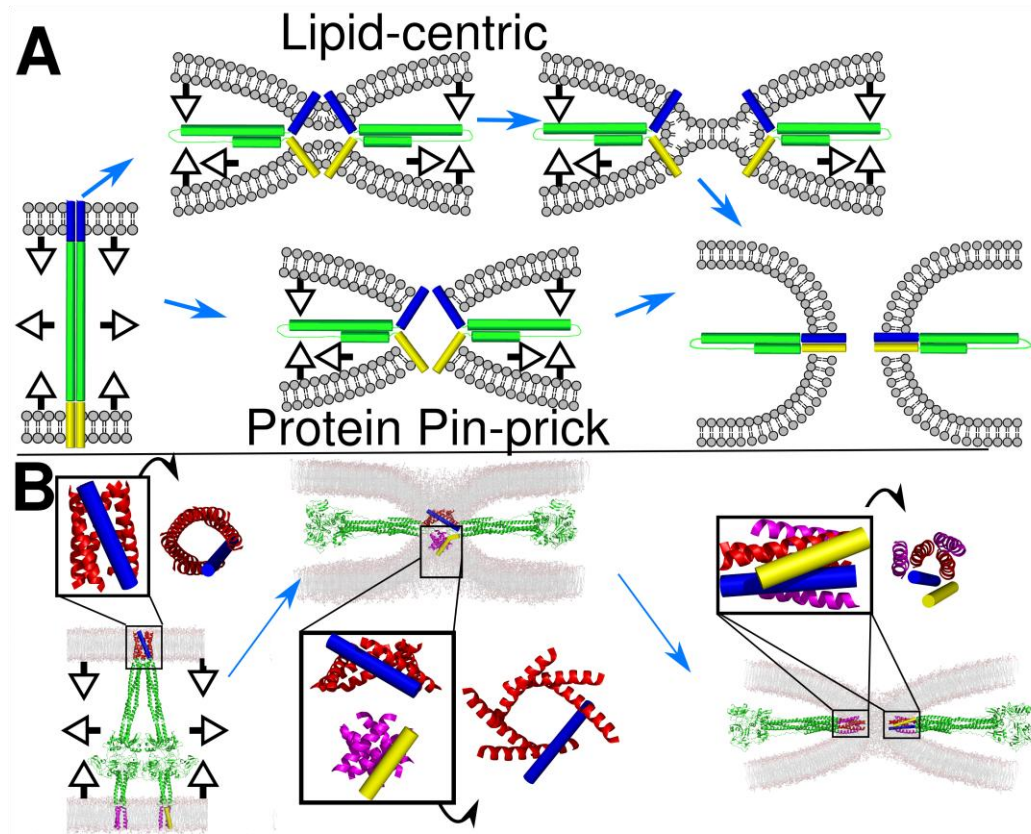


Figure 2.11 Provisional model of PIV5 fusion. (A) Schematic diagrams of the limiting extremes of lipid-centric and pin-prick fusion. (B) Shown is a model of the conformational change of the F protein (FP) hexamer (6HB) from a pre-hairpin, extended intermediate (left) to a point of membrane apposition (middle) and finally to the postfusion state (right). Proposed conformations of the FP 6HB are shown in the insets along with 90° rotations. Note the increased tilt of the FP moving from the extended

intermediate to the point of membrane apposition as well as the recruitment of lipid headgroups to the nascent pore. FPs are shown in red and blue, C-term-TMs are shown in magenta and yellow. The C-term-TM in the middle image contains two trimeric structures<sup>138</sup>.

## 2.5 Method

### Peptide Synthesis, cleavage and purification

The wild-type sequence for PIV5 fusion peptide (FP) is FAGVVIGLAALGVATAAQVTAAVALVKANE. The only glutamine, residue 120 (using the numbering of the fusion protein before cleavage), was mutated into alanine for the mutant Q120A. One tryptophan was added to the C-terminus of both the wild-type FP and the mutant Q120A using a flexible (Gly)<sub>3</sub> linker in order to provide absorbance at 280 nm for concentration measurements. The sequence for C-terminal membrane segment TM (C-term-TM) peptide is VLSIIAIALGSLGLILILLSSVVDWK, and contains a tryptophan for concentration measurements.

Peptide synthesis, cleavage and purification were conducted as previously described<sup>139</sup>. 0.1 mmole scale synthesis was manually conducted on RINK amide resin (Novabiochem) by N-9-fluorenylmethyloxycarbonyl (Fmoc) amino acids (using a four-fold molar excess) in a microwave synthesizer (CEM Discover). The peptide was cleaved using a trifluoroacetic acid (TFA):water:1,2-ethanedithiol ratio of 95:2.5:2.5. Peptide purification was run on a semi reverse phase HPLC (Vydac, C4 column, 250mm x 10mm i.d.) at 60° C in a gradient between solvent A (water with 0.1% TFA) and solvent B' (isopropanol: acetonitrile:water in a ratio of 6:3:1 with 0.1% TFA). The identity and purity of the peptide were confirmed by MALDI-TOF mass spectrometry

(Voyager model DE RP; PerSeptive Biosystems) and analytical reverse phase HPLC (Vydac C4 column).

### **Sedimentation equilibrium of analytical ultracentrifugation (AUC)**

Sedimentation equilibrium experiments and data analysis were performed as described previously<sup>82, 139, 141</sup>. Wild-type and mutant Q120A FPs was mixed with dodecylphosphocholine (DPC) in 2,2,2-trifluoroethanol (TFE). The cocktail was dried in a glass vial, lyophilized over night, and rehydrated with buffer in order to reach the DPC concentration of 8mM. This pH-7.3 buffer contained 100 mM HEPES, 100 mM NaCl and 37% D<sub>2</sub>O in order to match the density of the detergent.

Three groups of samples were prepared as peptide:DPC molar ratios of 1:50, 1:100 and 1:200. The experiments were conducted at 25°C using a Beckman XL-I analytical ultracentrifuge at 30, 35, 40, and 45 kRPM. In addition, some samples used additional experiments at 48 kRPM. Data obtained were globally fitted to a nonlinear least-squares curve by IGOR Pro (Wavemetrics) as previously described<sup>157</sup>.

Two AUC experiments were carried out to identify the interaction between the FP and the C-term-TM peptide. First, the C-term-TM peptide alone was prepared as a peptide:DPC molar ratio of 1:100. No significant curvature has been observed in AUC

data despite running the sample at multiple speeds (30, 35, 40, 45 and 50 kRPM). The C-Term-TM peptide and the FP were mixed in a 1:1 molar ratio at a total peptide:DPC molar ratio of 1:50 (Figure 2.4B left side) and 1:100 (Figure 2.4B right side). The experiments were conducted under the same conditions as those for the FP alone. The FP used here was synthesized using the original sequence, without Trp labeled, thus the FP itself did not have absorbance at 280nm. The significant curvature from the mixture suggests the strong interaction between C-term-TM peptide and FP. The concentration of the C-term-TM peptide was identified based on the absorbance at 280nm. The concentration of the FP was identified using a micro-balance and dissolved in TFE. Thus the determined concentration for the FP here might have 5% error from absorbing water from atmosphere, and presumably contributed the fitting error in the left side of the curves.

To estimate a lower limit of the strength of C-term-TM association with FP, the avidity of C-term-TM homooligomerization with and without FP were compared. For C-term-TM alone, the association is weak. The midpoint occurs in the range of 1:3-1:5 peptide:detergent. The midpoint is similar when fit as a dimer, trimer, or hexamer. In the presence of FP, the midpoint of heteromeric association occurs at approximately 1:100 peptide:detergent. Because the heteromeric interaction appears to involve an equal number of FP and C-term-TM peptides, for this calculation the molecular weight, extinction coefficient, and partial specific volume values for the FP and C-term-TM were

averaged and fit to a monomer-12mer equilibrium. The ratio of these self-association values provides a conservative estimate of C-term-TM association of at least 20 times stronger association in the presence of FP than in isolation.

### **Circular Dichroism (CD)**

CD spectra were collected with a Jasco J-810 spectropolarimeter and a 0.1 cm quartz cell using a 1nm step at 25°C. Peptide at 12.5  $\mu\text{M}$  was incorporated into 2.5 mM detergent DPC using the method described above and rehydrated into aqueous buffer containing 10 mM disodium phosphate of pH 7.4. The CD spectrum of each peptide was obtained by subtracting the spectrum of DPC alone and averaging over three scans.

### **Attenuated total reflection IR spectroscopy (ATR-IR)**

ATR-IR experiment and data analysis were conducted as previously described<sup>143</sup>,<sup>144</sup>. 140  $\mu\text{L}$  of 300  $\mu\text{M}$  wild-type or Q120A mutant FP was mixed with 25.5  $\mu\text{L}$  of 32.9 mM 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (*POPC*), in order to make a peptide:POPC ratio 1:20. The cocktail was loaded on the surface of ATR Ge crystal evenly and dried by air. The film was rehydrated by  $\text{D}_2\text{O}$ -saturated air overnight in closed environment of  $\text{D}_2\text{O}$  bath. All infrared spectra were measured in a Nicolet Magna-IR 860 spectrometer using 1  $\text{cm}^{-1}$  resolution. During data acquisition high-purity  $\text{N}_2$  gas continuously purged the spectrometer and  $\text{D}_2\text{O}$  hydrated high-purity  $\text{N}_2$  gas was



continuously purged upon the sample film to eliminate the spectral effects of water in air. The polarized mirror was adjusted to 0° and 90° in order to create incident light oriented parallel and perpendicular to the lipid normal, respectively. Each spectrum of a peptide was subtracted by the spectrum of the crystal alone at 0° and 90°, respectively. A total of 64 scans were averaged and Fourier-transformed to both wild-type and mutant. The dichroic ratio of 1656 cm<sup>-1</sup> amide I bond absorption is computed for parallel (0°) versus perpendicular (90°) polarized incident light relative to the membrane normal. The dichroic ratio was then applied to equations in <sup>143</sup>, in order to calculate the peptide orientation relative to the membrane normal.

### **Sequence conservation**

Because buried positions are more conserved than solvent or lipid exposed positions, an  $\alpha$ -helix will, in general, show a sinusoidal conservation pattern with approximately 3.6 residue periodicity <sup>158</sup>. If the helix is bent as part of a coiled coil, seven residues occur over every two turns of the helix, giving an average of 3.5 residue periodicity. A beta sheet, however, would be expected to show approximately 2 residues per period, as the residues alternate sides of the strand.

To determine the sequence conservation of the FP, sequences from the NCBI non-redundant database (February 9, 2009) were selected if the sequence matched the PIV5 f-

peptide sequence (FAGVVIGLAALGVATAAQVTAVAL) to an E-value of 1 or less<sup>159</sup>. The program Cd-hit<sup>160</sup> was used to remove sequences with higher than 90% sequence identity. The sequence entropy of each sequence position is calculated as: Entropy =  $-\sum_{i=1,20} \{ f_i \ln f_i \}$ , where  $f_i$  is the fraction of residue  $i$ . The entropy,  $x$ , was fit to a sine wave according to the formula:  $y = a \cdot \sin[2\pi(x+b)/c] + d$ , where the periodicity is found in variable  $c$ .

### **Creation of coiled coil models**

All-atom protein backbones of the TM coiled-coil FP hexameric bundle (6HB) were created using the Crick parameterization<sup>146</sup> according to a previously published method<sup>161</sup>. Three parameters were allowed to vary: alpha helical phase ( $\phi$ ), pitch angle ( $\alpha$ ), and superhelical radius ( $R$ ). The other coiled coil parameters were held fixed.  $\phi$  was varied from 0° to 359° in 1 deg steps;  $\alpha$  from 5° to 20° in 1 deg steps; and  $R$  from 9.0 to 11.0 Å in 0.1 Å steps. These ranges were chosen based on the values observed in other coiled coils and such that no backbone clashes would occur in the complex.

For each backbone 6HB model, side chains were placed using the program scap<sup>162</sup> and hydrogens placed by reduce<sup>163</sup>. Each structure was subject to a constrained minimization in CHARMM22<sup>147</sup> of 50 steps to decrease, but not remove, the penalty for a clash in a given structure and to not move significantly from the initial coiled coil

parameters. The energy of the structure was then calculated using CHARMM22 and IMM1 implicit solvation<sup>148</sup> with a non-bonding interaction cutoff of 9.0 Å. The selected minimized models had the following ( $\phi$ ,  $\alpha$ ,  $R$ ) parameters before simulation: (40°, 10.3 Å, 22.1°), (43°, 10.4 Å, 15.9°), (88°, 10.4 Å, 16.0°), (196°, 9.7 Å, 13.6°), and (300°, 10.7 Å, 15.6°). The parameters of the antiparallel model were: (24°, 10.2 Å, 9.0°) with a z-offset of 0.5 Å.

### **Molecular dynamics simulations**

Models of the FP hexameric bundle (6HB) were embedded in a lipid bilayer composed of 140 POPC molecules (80x80 Å in size), capped on each side by a water layer of 18 Å thickness (6,500 water molecules in total). Periodic boundary conditions were applied. 11 Na<sup>+</sup> and 11 Cl<sup>-</sup> ions were distributed in the water region, corresponding to a salt concentration of about 150 mM. The six peptides and the ions were described by the CHARMM27 force field<sup>164</sup>, the water molecules by the TIP3P force field<sup>165</sup>, and the lipid molecules by the united-atom force field recently developed by Hénin *et al*, which provides nearly identical physicochemical properties to the CHARMM27 lipid<sup>166</sup>. The van der Waals interactions were truncated at 12 Å, and a grid resolution of 0.75 Å was used to treat the electrostatic interactions with the Particle-Mesh-Ewald (PME) scheme<sup>167</sup>.

Each 6HB system was simulated by molecular dynamics (MD), using a time step of 2 fs. A Langevin thermostat<sup>168</sup> was applied to maintain a temperature at 310 K (~37 C), and a Langevin-piston barostat<sup>169</sup> to keep a pressure of 1 bar along the bilayer normal. In the two directions parallel to the membrane, instead, a constant surface tension of 20 dyn/cm<sup>2</sup> was enforced. The NAMD program<sup>170</sup> was used to perform all the MD calculations presented here.

Because the interactions between the peptides and the surrounding lipids are of major importance to this study, and due to the fact that the starting HB structures were modeled within an implicit membrane, a rather long equilibration phase was performed. For each starting structure, a restraint of 10 kcal/mol/Å was applied to the peptide heavy atoms for the first 2 ns, and on the backbone heavy atoms only for the following 7 ns. The time evolution of the system was monitored during the following 50 ns of MD simulation without restraints.

The density of water oxygen atoms in Figure 2.8 was computed by counting all atoms within a radius of 12 Å from the central axis of the bundle. We used this criterion to account for the wide aperture of certain 6HB models ( $\varphi = 88^\circ, 300^\circ$ ) and the oscillations of the bundle with respect to the membrane normal, while not including at the same time a detectable number of water molecules in the lipid interstitial regions: water density profiles computed with different radii do not differ significantly between  $z = -15$

Å and 15 Å. The water density isosurface in Figure 2.9B shows instead the distribution of all water molecules of the system in the HB ( $\varphi = 40^\circ$ ) simulation.

### **FTIR Dichroism calculations**

Following the method of Arkin and coworkers<sup>150</sup>, we calculated the expected dichroism ratio by summing the contributions to carbonyl groups to each polarization given the backbone dipoles vectors in the molecular dynamics trajectory. For these calculations 5% disorder was assumed. Because non-helical termini and the first three helical residues are expected to exchange with D<sub>2</sub>O, only residues 108 through 126 were used for the calculation. Including the full FP gives very similar results.

### **Modeling of the fusion process**

To model the fusion process, both the PIV5 F protein and the membrane needed to be modeled. A 200 Å by 200 Å POPC lipid bilayer was created using VMD<sup>171</sup>. For early stages of the fusion model, the bilayer was perturbed using a cosine function with an increasing amplitude centered at the region that would become the point of membrane apposition. Later stages incorporated increasing fractions of the conformation of a catenoid that makes up the final, postfusion conformation of the model. The diameter of this pore is estimated to be 20 Å based on experimental results for influenza hemagglutinin<sup>105, 172</sup>. For the catenoid, the midpoint of the membrane followed the

mathematical formula while lipids were rotated to be perpendicular to the surface. Lipid density was maintained between the initial flat membrane surface and the catenoid structure.

To model the F protein, first the available prefusion and postfusion crystal structures were combined to create a model of the prehairpin intermediate that bridges the viral and target cell membranes. Transmembrane domains were modeled either as coiled coils extended into the membrane or using the existing model of the prefusion TM<sup>138</sup>. Loops were modeled using loopy<sup>173</sup>, the side chains were repacked using Rosetta<sup>174</sup>, and the structure minimized using the CHARMM22 energy function<sup>147</sup>. The structure of the postfusion state comes from the closely related hPIV3; the sequence of PIV5 was threaded onto the structure using Rosetta and minimized using the CHARMM energy function.

For fusion intermediates, the membrane associated domains were first modeled. The hexameric model of the fusion protein was taken as the initial state for the prehairpin intermediate. The helices were then tilted in 10° increments and slowly moved outward from the center to mimic the conformational change proposed for the MscL channel<sup>130</sup>. In parallel, two TM domain trimers<sup>138</sup> were initially at a distance due to the conformation of the prehairpin intermediate, then brought close together and tilted relative to each other as may occur under the tension caused by zippering of the soluble coiled coil domains.

The soluble portions of the F protein trimers were tilted to connect to the FPs at each step. Loops between the FP and HRA were connected using loopy<sup>173</sup> and Pymol sculpting. The conformation of the exterior coiled coil, HRB, was then modified to connect to the TM domain using an loopy in-house loop modeling program and the BBQ backbone modeling program<sup>175</sup>.

## **2.6 Contributions**

Yao Zhang performed all the experiments and related analysis in this chapter, Jason E. Donald, Giacomo Fiorin, Vincenzo Carnevale, and David R. Slocower performed the computational modeling and molecular dynamic simulations.

This chapter has been published in Proceedings of the National Academy of Sciences, 2011:108(10):3958-63.

## Chapter 3: Experimental and computational evaluation of forces directing the association of transmembrane helices

### 3.1 Abstract

The forces that define the interactions of transmembrane helices have been evaluated using a model membrane-soluble peptide (MS1), whose packing is modeled on the two-stranded coiled-coil from GCN4. The thermodynamic stability of water-soluble coiled-coils depends on the side chain at the buried **a** position of the repeat, favoring large hydrophobic residues over small side chains. Here we show that just the opposite is true for the membrane-soluble peptide. Analytical ultracentrifugation and equilibrium disulfide interchange show that the stability of MS1 is greatest when Gly is at each **a** position of the heptad repeat (MS1-Gly), followed by Ala > Val > Ile. Moreover, MS1-Gly has a strong tendency to form antiparallel dimers, MS1-Ala forms a mixture of parallel and antiparallel dimers, while MS1-Val and MS1-Ile have a preference to form parallel dimers. Calculations based on exhaustive conformational searching and rotamer optimization were in excellent agreement with experiments, in terms of the overall stability of the structures and the preference for parallel vs antiparallel packing. The MS1-Gly helices are able to achieve more favorable and uniform packing in an antiparallel dimer, while MS1-Val and MS1-Ile have more favorable van der Waals interactions in a parallel dimer. Finally, the electrostatic component arising from the partial charges of the backbones become significant in the antiparallel MS1-Gly and MS1-Ala conformations, due to close packing of the helices. Thus, van der Waals



interactions and electrostatic interactions contribute to the stability and orientational preferences of the dimers.

### 3.2 Introduction

While our understanding of the features stabilizing the structures of water-soluble proteins has reached an advanced state, a parallel understanding of membrane protein folding is only beginning to emerge.<sup>59</sup> Previous work on water-soluble proteins suggests that the burial of hydrophobic residues plays a crucial role for folding.<sup>176, 177</sup> Similarly, the replacement of large apolar sidechains with smaller residues in the interior of membrane proteins results in introduction of cavities with a concomitant loss in thermodynamic stability.<sup>178</sup> From this perspective, the packing of large apolar sidechains can stabilize the folded structure of membrane protein. By contrast, statistical,<sup>74, 75, 179-181</sup> computational<sup>77, 182, 183</sup> and experimental<sup>77, 182, 184</sup> studies have demonstrated that small side chains, such as Ser, Ala and Gly, occur frequently at the helix-helix interface of membrane proteins, suggesting that the appropriate packing of these residues might provide an even stronger driving force for transmembrane (TM) helix association. We therefore compared the effects of packing large vs. small apolar sidechains, using a simple transmembrane TM helical dimer.

Coiled-coils, such as the leucine zipper from GCN4,<sup>176, 177</sup> have a repetitive 7-residue repeat providing a conceptually simple system for studying sidechain packing. By convention, the residues at the **a** and **d** positions of the heptad pack in the core of a coiled coil. The stability of water-soluble coiled-coils scales with the size and hydrophobicity of the sidechains at the **a** position increasing over the series Gly < Ala < Val < Ile.<sup>177, 185</sup>

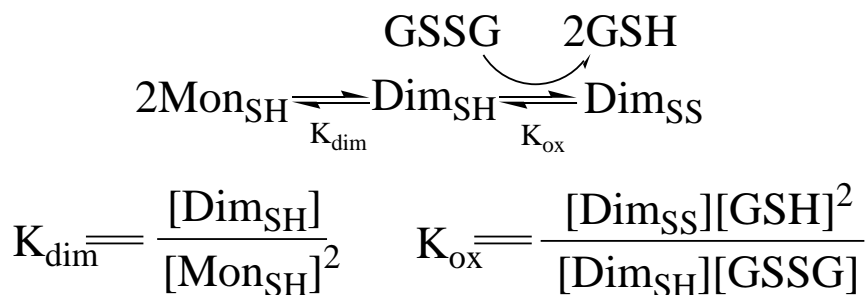
Interestingly here we show just the opposite rank for MS1, a membrane-soluble version of a leucine zipper.

### 3.3 Results and Discussions

MS1 is a membrane-soluble derivative of GCN4-P1, rendered lipophilic by converting its exposed polar sidechains to apolar residues, while maintaining the core residues constant.<sup>60</sup> We synthesized a series of MS1 variants in which each of the four **a** positions was varied to Gly, Ala, Val and Ile (Figure 3.1). Each of these peptides is predominantly helical in dodecylphosphocholine (DPC) micelle as determined by circular dichroism over the entire range of peptide/DPC ratios studied here (Figure 3.2). Their assembly was first examined in DPC micelles by analytical ultracentrifugation (AUC) under conditions where the density of the solution is adjusted by addition of D<sub>2</sub>O to eliminate the mass contribution of DPC. The degree of association of membrane peptides in micelles depends on the concentration of peptide in the micelle phase as reflected in the peptide/detergent ratio. Over all experimentally accessible peptide/detergent ratios, MS1-Gly was fully dimeric, MS1-Ala adopted a monomer-dimer equilibrium, and the most hydrophobic peptides, MS1-Val and MS1-Ile, were predominantly monomeric (Table 3.1, Figure 3.3). This ranking is precisely the opposite of that found in water-soluble structures.

To explore the association strength of the MS1 variants we employed the method of equilibrium thiol/disulfide exchange,<sup>141</sup> which is well-suited for examining weak interactions. The N-terminus was modified with a flexible three-glycine linker followed by a cysteine (Figure 3.1C). After peptides were incorporated into detergent micelles, redox buffer was added to bring the system to the following equilibrium (scheme 1).

*Scheme 1.*



The two steps in scheme 1 are linked but depend differently on the peptide concentration: the dimerization step ( $K_{\text{dim}}$ ) is a reversible bimolecular association reaction that depends on the reduced monomer concentration ( $\text{Mon}_{\text{SH}}$ ); the subsequent oxidation step ( $K_{\text{ox}}$ ) is also reversible, but independent of the concentration of the peptide and dependent on the ratio of oxidized (GSSG) to reduced glutathione (GSH). Using this function we fit curves to obtain the parameters  $K_{\text{dim}}$  and  $K_{\text{ox}}$  for each MS1 variant (Table 3.2; Figure 3.4, 3.5A). Comparison of  $\Delta G_{\text{dim}}$  for each variant suggests that the amino acid in the **a** position aids the association of membrane helices in increasing order of: Gly > Ala > Val > Ile, in good agreement with the AUC data. Clearly, the association of the helices increases as the size of the core positions side chain decreases.

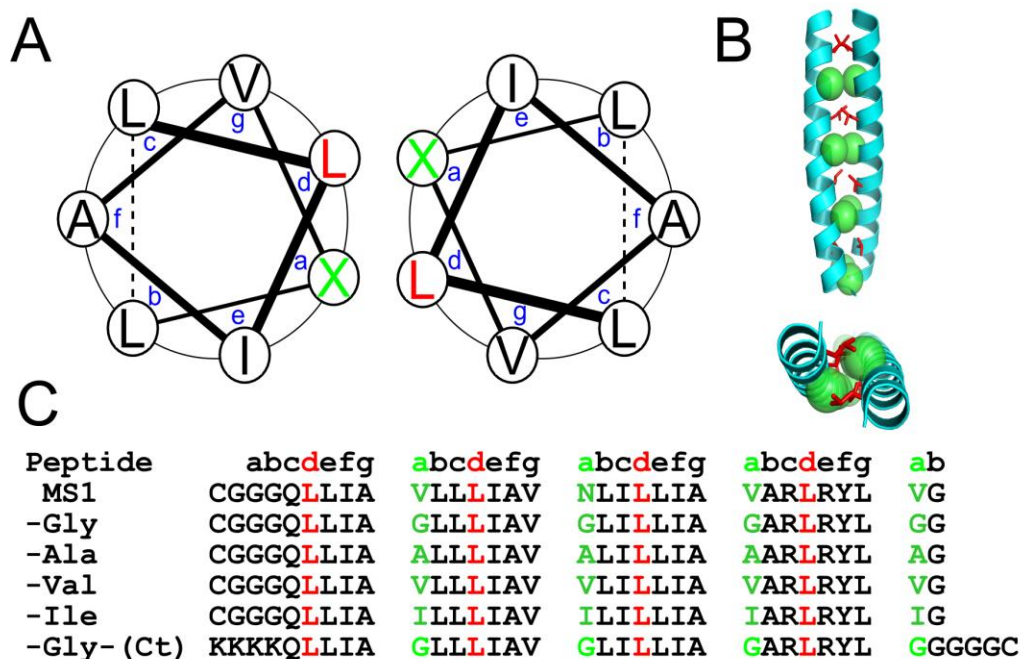


Figure 3.1 Helical wheel (A), computational model showing side and top view (B), and sequence of MS1 variants (C). MS1, -Gly, -Ala, -Val, and -Ile are N-terminally Cys-modified. -Gly-(Ct) is C-terminally Cys-modified. The variable **a** positions are shown in green, and the Leu at **d** in red. All peptides had an N-terminal acetyl group.

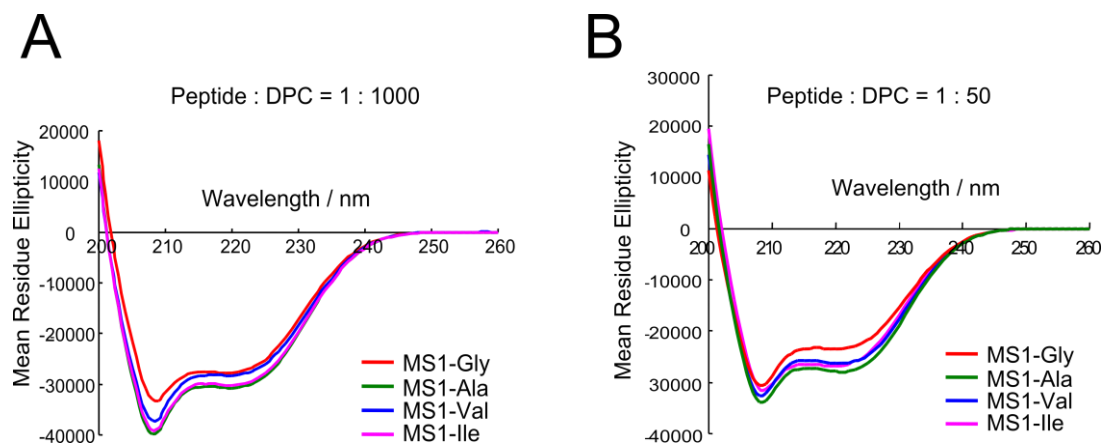


Figure 3.2 CD Spectra of MS1 variants (units are  $\text{deg cm}^2 \text{dmol}^{-1} \text{res}^{-1}$ ). The spectra show that all peptides are predominantly  $\alpha$ -helical at a peptide/detergent ratio of 1/1000 or 1/50 (matching the concentration range for the thiol-disulfide exchange equilibria). The magnitude of the ellipticity at 222 nm is the same within the experimental error, which derives primarily from the concentration determination used to compute the mean residue ellipticity. Buffer conditions and other methods are given below.

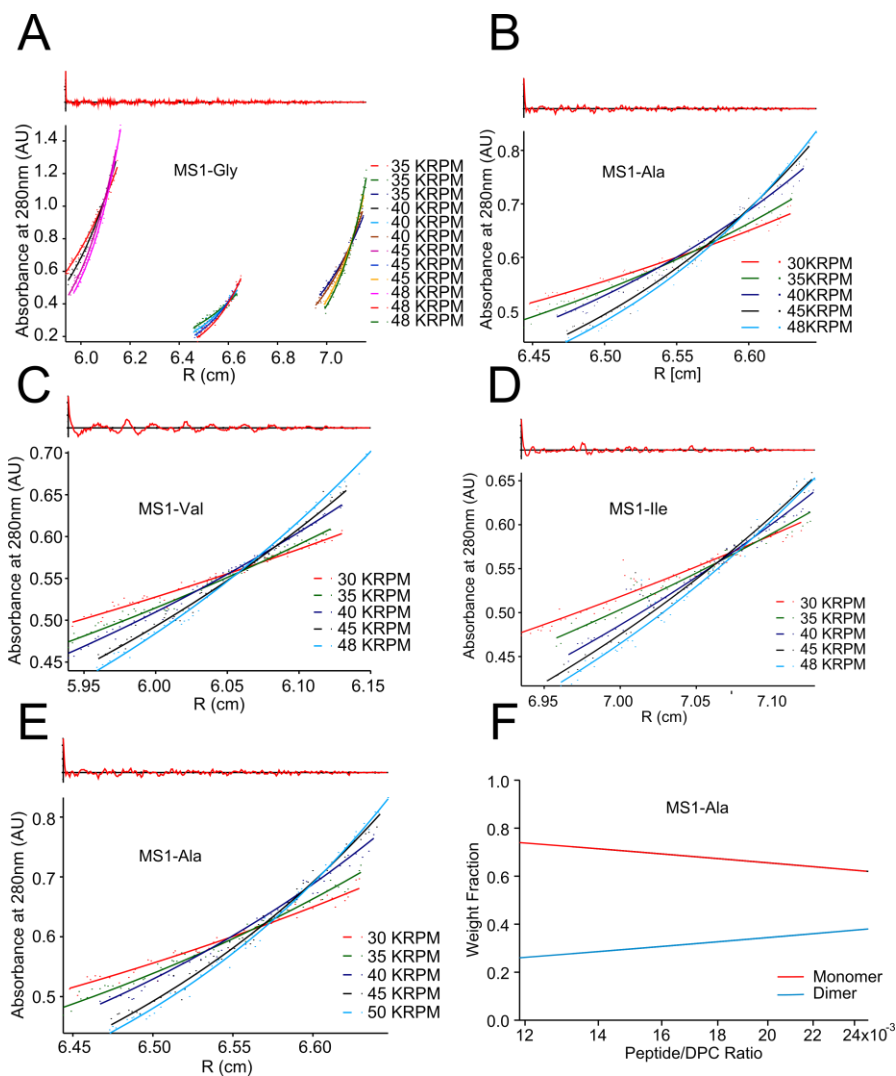


Figure 3.3 Sedimentation equilibrium analysis of the reduced MS1 variants in DPC micelles: (A) MS1-Gly, (B) MS1-Ala, (C), MS1-Val, (D) and MS1-Ile. The absorbance was measured at 280nm. The top panels shows the residuals of single species fitting to data at respectively 30, 35, 40, 45, 48 KRPM. (E) MS1-Ala is fitted with monomer-dimer equilibrium, resulting in  $pK_{dim}$  equal to -1.3. In this analysis the concentration of peptide is mole ratio of peptide/detergent, and hence is unitless. (F) Monomer/dimer species



distribution of MSI-Ala as a function of peptide/DPC ratio based on fitting in (E). Buffer conditions and other methods are given below.

Table 3.1 Association states of MS1 variants determined by AUC

|         | Observed MW | Monomer MW | Ratio <sup>a</sup> |
|---------|-------------|------------|--------------------|
| MS1-Gly | 6300±600    | 3214.4     | 2.0±0.2            |
| MS1-Ala | 5600±200    | 3270.5     | 1.7±0.1            |
| MS1-Val | 3200±300    | 3359.3     | 1.0±0.1            |
| MS1-Ile | 3100±500    | 3438.8     | 1.0±0.2            |

<sup>a</sup> Ratio = Observed MW / monomer MW

Table 3.2 pK<sub>dim</sub> and pK<sub>ox</sub> obtained from analysis of data in Figure 3.2A

|      | pK <sub>dim</sub> | pK <sub>ox</sub> | ΔG <sub>dim</sub> (Kcal/mol dimer) <sub>-</sub> |
|------|-------------------|------------------|---|
| MS1  | -2.6              | 1.5              | -3.6  |
| -Gly | -3.0              | 2.6              | -4.1  |
| -Ala | -1.6              | 1.2              | -2.2  |
| -Val | 0.9               | -0.8             | 1.2   |
| -Ile | 1.9               | -2.0             | 2.6   |

The error is estimated to be approximately 10% based on the error in experimental concentrations of the reduced and oxidized peptides

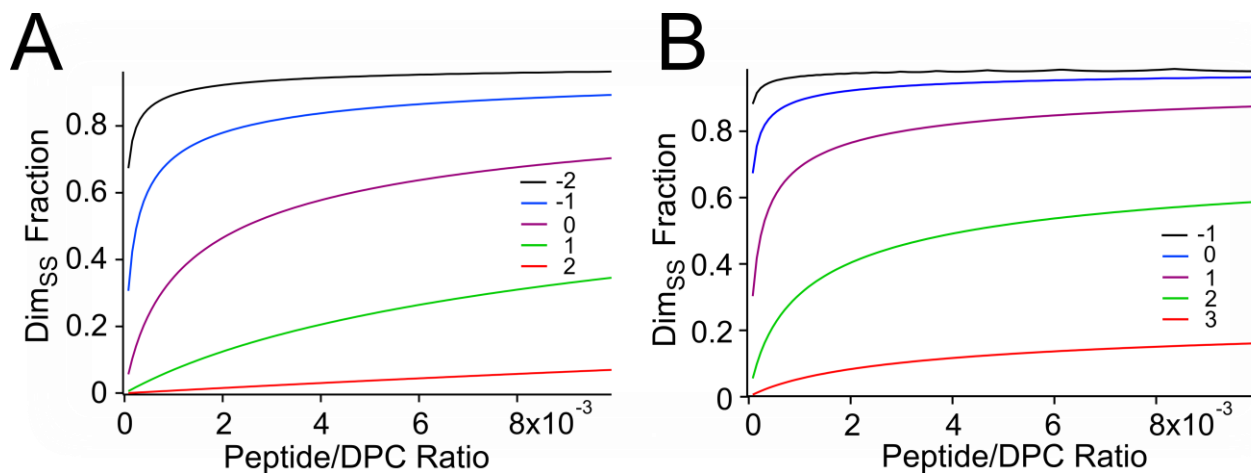


Figure 3.4 Simulation of disulfide exchange equilibrium related to various dimerization affinities ( $K_{dim}$ ) and redox ratios ( $K_{ox}$ ). The figure shows theoretical fractions of crosslinked dimer as a function of peptide/DPC ratio at (A) varying  $pK_{dim}$  ( $pK_{ox}=0$ ), and at (B) varying  $pK_{ox}$  ( $pK_{dim}=0$ ). Note that the shapes of the curves vary depending on both parameters. Thus, the figures demonstrates that thiol disulfide equilibrium method is sensitive to both the free energy of association as well as and redox potential for a given peptide.

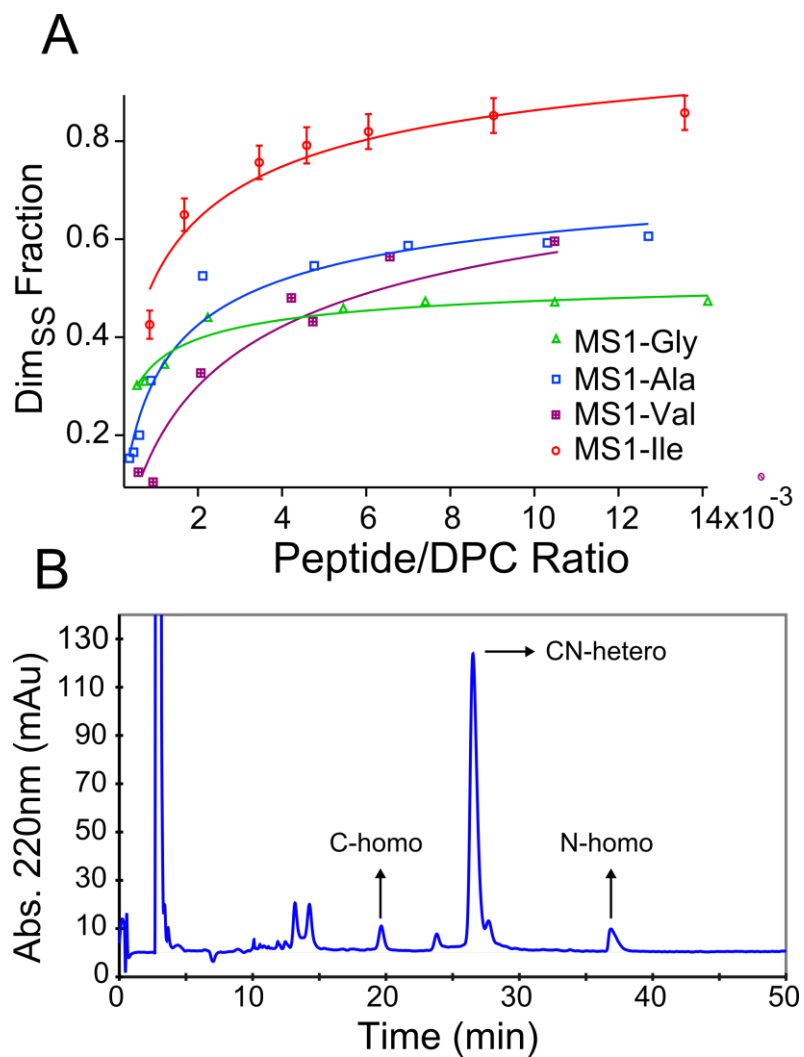
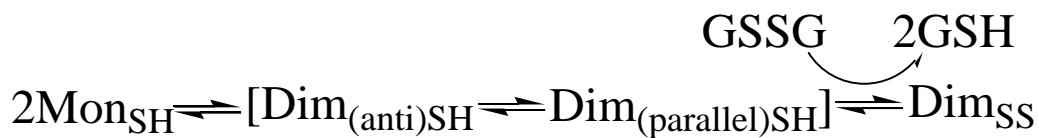


Figure 3.5 (A) Analysis of the fraction of crosslinked dimer as a function of peptide/DPC ratios for each MS1/variant. The theoretical curve describes the least-squares fit to scheme 1 (Table 3.2). The standard errors in the experimental points are similar for each peptide and are indicated for MS1-Ile (Others are not so shown for clarity). (B) HPLC chromatogram after redox equilibration of the C-terminal and N-terminal Cys-modified MS1-Gly mixture indicates that MS1-Gly prefers an anti-parallel orientation. The other peaks are glutathione adducts.

Interestingly, as the size of the side chain at the **a** position decreases, the ease of disulfide formation (reflected in  $K_{ox}$ ) becomes less favorable. This was surprising, given that the peptides have a Gly3 linker between the helical ends and the Cys. As long as the helices pack in a parallel manner, the flexibility of the linker should easily accommodate any subtle differences in helix-packing (two Cys-Gly3 linkers could extend up to about 20 Å, while interhelical distances in dimers vary by only a couple Å). Thus, we considered the possibility that MS1-Gly prefers to assume an antiparallel orientation. In this case, association of  $Mon_{SH}$  would remain favorable, but the oxidation step would require unfavorable intramolecular rearrangement of the antiparallel dimer,  $Dim_{(anti)SH}$ , to the parallel  $Dim_{(parallel)SH}$  to allow disulfide formation, because of the need to shift the equilibrium from one favoring antiparallel to parallel dimers upon disulfide formation (Scheme 2).

**Scheme 2.**



To test this hypothesis, we synthesized a C-terminally Cys-modified peptide (Figure 3.1C, -Gly-(Ct)), which was mixed in equal amounts with N-terminally Cys-labeled MS1-Gly under reversible redox conditions. If MS1-Gly has the same preference to form parallel dimers as antiparallel dimers, then a ratio of 1:1:2 (N-terminal

homodimer : C-terminal homodimer : heterodimer) is expected. However, the experimental ratio is 1:1:14, indicating that MS1-Gly strongly prefers to form anti-parallel dimers (Figure 3.5B).

To probe further the orientation of MS1-Gly and the other variants, the peptides with N-terminal Cys residues were individually air-oxidized to force an N-terminal crosslink (Figure 3.6, 3.7). Under these conditions peptides with a strong tendency to form antiparallel dimers might be expected to oligomerize as shown in Figure 3.3C. To avoid the precipitation of polymers during centrifugation we performed the experiment on samples that were approximately (75±5)% oxidized. The ratio between the computed molecular weight from a single-species fit and the computed monomeric molecular weight roughly reflects the degree of oligomerization (Table 3.3). The computed ratio for MS1-Gly is 6.9, supporting the expectation that MS1-Gly prefers an anti-parallel orientation. MS1-Ala has a ratio of about 3, in agreement with the conclusion that this peptide prefers to form weak, antiparallel dimers. The ratio for MS1-Val and MS1-Ile is less than two, again consistent with the suggestion that they form even weaker parallel dimers. Thus, as the side chains in the core positions become smaller, the helices prefer to form antiparallel orientation.

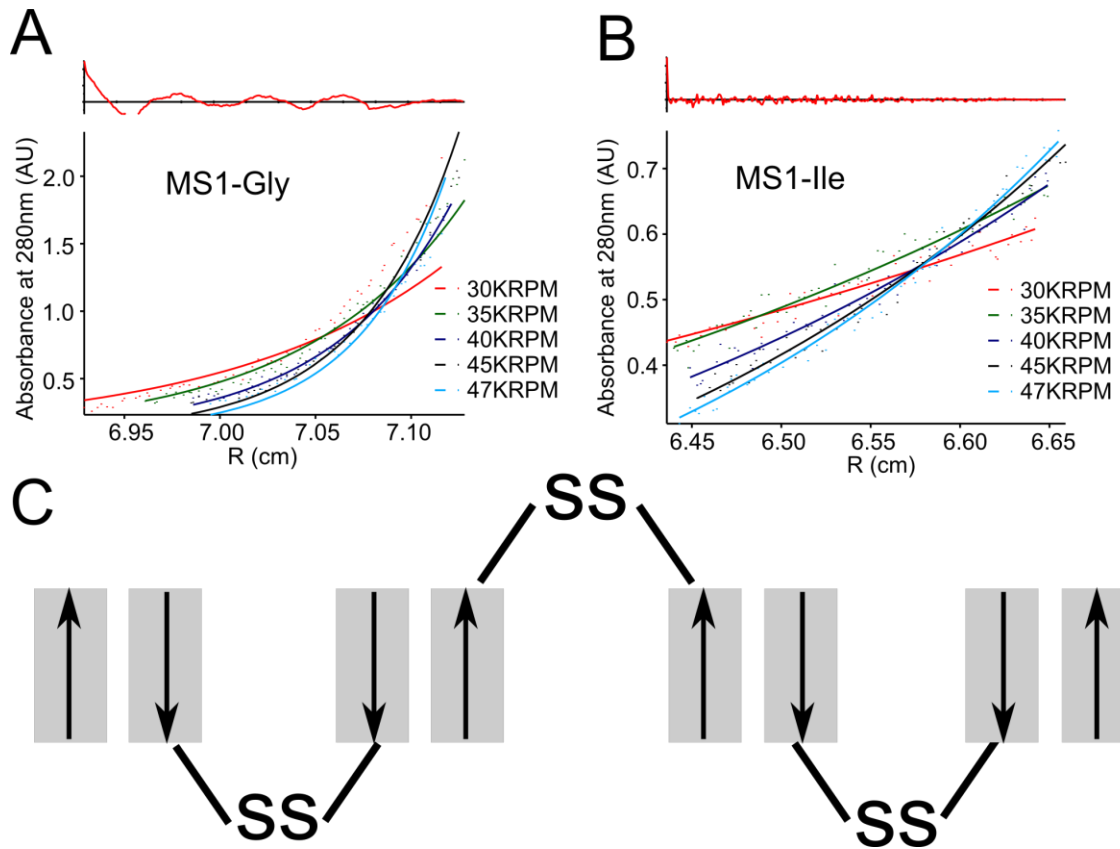


Figure 3.6 Analytical ultracentrifugation of approximately 80% disulfide-bonded MS1-Gly (A) and MS1-Ile (B). The greater degree of curvature in panel A vs. B is indicative of greater oligomerization. (C) Oligomerization of MS1-Gly and MS1-Ala via formation of antiparallel dimers.

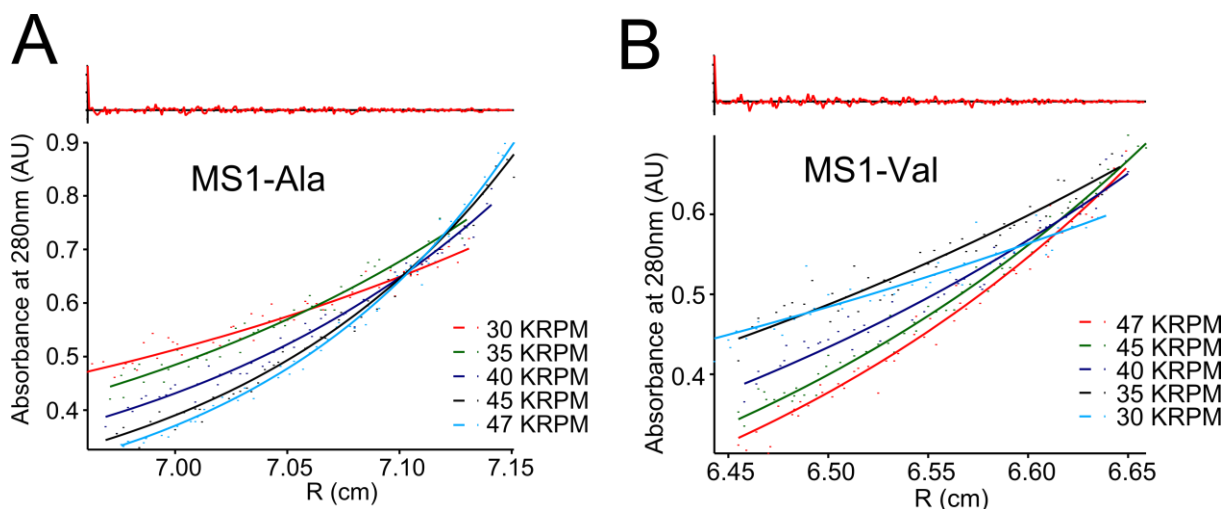


Figure 3.7 Sedimentation equilibrium analysis of the MS1 variants in DPC micelles: (a) MS1-Ala and (b) MS1-Val. Absorbance is measured at 280nm. Top panels shows residuals of single species fitting to data at respectively 30, 35, 40, 45, 47 KRPM. Peptides were first air-oxidized in buffer. The percentage of air oxidized dimer was determined by analytical HPLC to be (75±5)%. Buffer conditions and other methods are given below.

Table 3.3 Degree of association of air-oxidized MS1 variants in DPC micelles determined by analytical ultracentrifugation

|         | Observed MW | Monomer MW | Ratio <sup>a</sup> |
|---------|-------------|------------|--------------------|
| MS1-Gly | 22200±800   | 3214.4     | 6.9±0.1            |
| MS1-Ala | 9200±1200   | 3270.5     | 2.8±0.4            |
| MS1-Val | 5000±1700   | 3359.3     | 1.5±0.5            |
| MS1-Ile | 4300±1300   | 3438.9     | 1.3±0.4            |

<sup>a</sup> Ratio defined as in Table 3.1 (Main text). The errors are large for low observed MW, due to the small curvature in the radial distribution curves, adding uncertainty in the fitted baseline parameter.

To investigate the energetic and structural mechanisms behind these observations we built parallel and antiparallel computational models for the MS1 variants. While long simulations in bilayers would be essential to fully evaluate the relative energetic contributions from helix-helix, helix-lipid and lipid-lipid components, successful models<sup>186</sup> and designs<sup>82</sup> of transmembrane proteins have been achieved by probing helix-helix packing interactions alone, using a much simpler gas phase potential energy function. For each sequence, the conformational space available to parallel and antiparallel two-stranded coiled coils (Figure 3.8A) was globally searched using a molecular mechanics force field to compute the difference in energy between the homodimer versus the isolated monomers as described in the supplement. The resulting energy landscapes (Figure 3.8B) have global minima corresponding to structures in which the variable **a** position projects towards the core of the structure as in Figure 3.1B.



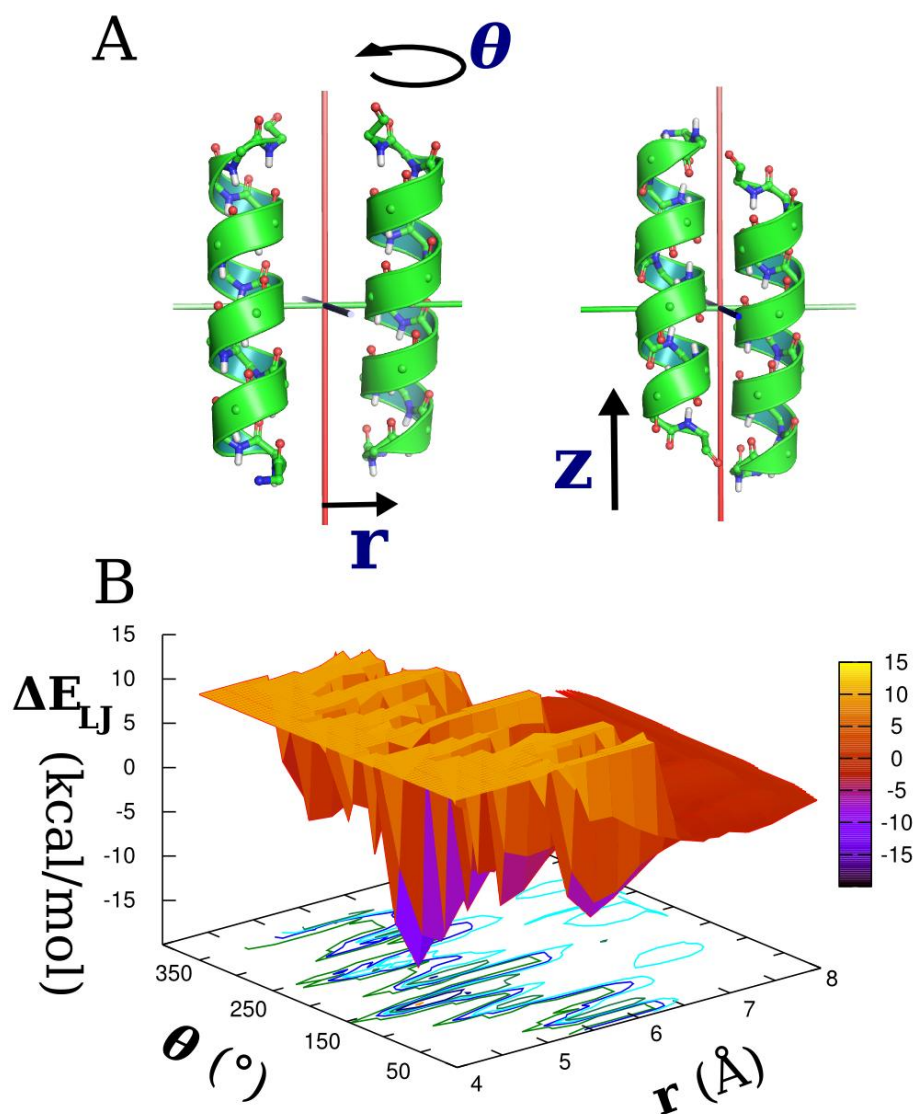


Figure 3.8 (A) Sampling Crick parameters for parallel ( $r$ ,  $\theta$ ) and anti-parallel ( $r$ ,  $\theta$ ,  $z$ -translation) dimers. (B) Energy landscape showing the difference in computed Leonard-Jones energy  $E_{LJ}$  (for the dimer versus two monomers) of MS1-Gly in a parallel orientation. The minimum in the surface has a helical phase ( $\theta$ ) of  $154^\circ$ , allowing packing of the Gly residues at the helix/helix interface as in Figure 3.1B.

The calculations are in remarkable agreement with experiment, given the stark simplicity of the calculations. The global minimum energy conformations (GMEC) for MS1-Gly and MS1-Ala correspond to antiparallel structures, which also allow the closest approach of the helices (Table 3.4). By contrast, the GMEC conformations for MS1-Val and MS1-Ile, correspond to parallel structures. To gain insight into the interactions responsible for these structures, some of the energetic components were investigated, specifically the change in Leonard-Jones energy ( $\Delta E_{LJ}$ , approximating the van der Waals component) and the electrostatic term associated with interactions between the partial charges of the main chain atoms at the interface ( $\Delta E_{bb}$ ). The values of  $\Delta E_{LJ}$  for the GMEC structures correlate with the experimental ranking ( $\Delta G_{dim}$ ), both in terms of overall energetics of association as well as the preference for parallel versus antiparallel structures (Table 3.5). Moreover, although the magnitude of  $\Delta E_{bb}$  depends on the electrostatic treatment employed in these calculations, there is a clear trend towards greater stabilization of the antiparallel structure as the residues at the **a** position (and hence the inter-helical separation) become smaller.

Table 3.4 Energetic contributions for minima in energy landscape <sup>a</sup>

| Peptide | Parallel        |                 | Anti-Parallel   |                 | $\Delta(\text{orientation})^b$ |                        |
|---------|-----------------|-----------------|-----------------|-----------------|--------------------------------|------------------------|
|         | $\Delta E_{LJ}$ | $\Delta E_{bb}$ | $\Delta E_{LJ}$ | $\Delta E_{bb}$ | $\Delta \Delta E_{LJ}$         | $\Delta \Delta E_{bb}$ |
| Gly     | -23.5           | 2.8             | <b>-38.0</b>    | <b>-1.4</b>     | -14.6                          | -4.2                   |
| Ala     | -31.8           | 1.8             | <b>-36.1</b>    | <b>-0.9</b>     | -4.3                           | -2.7                   |
| Val     | <b>-35.2</b>    | <b>0.8</b>      | -26.2           | -1.0            | 9.1                            | -1.8                   |
| Ile     | <b>-32.3</b>    | <b>0.4</b>      | -28.8           | -0.1            | 3.5                            | -0.5                   |

<sup>a</sup> Units are kcal/mol based on the CHARMM Force Field. Values in bold give the values of  $\Delta E_{LJ}$  and  $\Delta E_{bb}$  associated with the GMECs. <sup>b</sup>  $\Delta(\text{orientation})$  represents the energetic difference between the global minimum for the anti-parallel vs. parallel orientations.

Table 3.5. Crick parameters at the potential energy minimum

| Peptide | Parallel |                    | Anti-Parallel |                    |                              |
|---------|----------|--------------------|---------------|--------------------|------------------------------|
|         | R(Å)     | $\theta(^{\circ})$ | R(Å)          | $\theta(^{\circ})$ | $Z_{\text{trans}}(\text{Å})$ |
| Gly     | 4.3      | 154                | 3.6           | 130                | -2.1                         |
| Ala     | 4.4      | 152                | 4.3           | 135                | -1.4                         |
| Val     | 4.6      | 154                | 4.5           | 145                | -1.8                         |
| Ile     | 4.9      | 146                | 4.8           | 150                | -1.2                         |

See Figure 3.8A (main text) for description of parameters

### 3.4 Conclusion and Discussion

These studies together with other studies of MS1 variants<sup>77</sup> show that small residues at TM helix-helix interfaces allow helices to come into close contact, concomitantly increasing their van der Waals interactions.<sup>80, 180</sup> Thus, they are in agreement with previous studies highlighting the importance of van der Waals interaction,<sup>178</sup> while also demonstrating the important role that small residues can play in allowing particularly efficient packing to occur. Finally, we show that close interhelical distances associated with packing of small sidechains can additionally facilitate interhelical electrostatic interactions between the partial charges of backbone atoms.

### **3.5 Methods**

#### **Peptide synthesis and purification**

Peptides were synthesized as C-terminal carboxamides at a 0.1 mmole scale on RINK amide resin (Novabiochem) by N-9-fluorenylmethyloxycarbonyl (Fmoc) amino acids (four-fold molar excess) on a Symphony peptide synthesizer. Standard coupling conditions are shown in ref.<sup>141</sup>

Side chain deprotection and simultaneous cleavage from the resin was performed using a mixture of trifluoroacetic acid (TFA)/water/triisopropylsilane/1, 2-ethanedithiol (92.5/2.5/2.5/2.5) and precipitated with cold diethyl ether. The peptides were purified by reverse phase HPLC using a preparative (Vydac, C4 column, 250mm x 9.4mm i.d.) in a gradient between solvent A (water with 0.1% TFA) and solvent B (isopropanol /acetonitrile 2:1 with 0.1% TFA). The purity of the peptides was confirmed by analytical reverse phase HPLC (Vydac C4 column) and MALDI-TOF mass spectrometry (Voyager model DE RP; PerSeptive Biosystems).

#### **Preparation of peptide detergent micelles**

MS1/variants and detergent dodecylphosphocholine (DPC) were dissolved in 2,2,2-trifluoroethanol (TFE) and mixed in a glass vial. The TFE was removed under a stream of nitrogen and subsequently in vacuum. The samples were re-hydrated with

100mM Tris-HCl (pH 8.6), 100 mM KCl and 1 mM ethylenediaminetetraacetic acid (EDTA) unless noted otherwise.

### **Circular dichroism**

CD spectra were acquired with a Jasco J-810 spectropolarimeter and a 0.1 cm quartz cell at 25°C. 50µM MS1/variants were incorporated in DPC at detergent/peptide molar ratios of 50 and 1000. CD measurements were carried out in aqueous buffer containing 2.5 mM Tris-HCl (pH 8.6), 2.5 mM KCl and 25 µM EDTA. The CD spectra were averaged over three scans using a 1 nm step. Baseline obtained by DPC in buffer was subtracted from all peptide spectra.

### **Thiol-disulfide exchange equilibria with glutathione redox buffer**

Thiol-disulfide exchange was conducted as described in reference.<sup>141</sup> 25 µM MS1/variants were incorporated into DPC at different DPC/peptide ratios from 50 to 1000. The samples were incubated in a reversible redox condition of a glutathione buffer containing 0.45 mM oxidized (GSSG) and 1.05 mM reduced (GSH) glutathione for 4 hours to reach equilibrium, followed by quenching with HCl at final concentration of 0.12 M. The mixtures were separated by analytical RP-HPLC and each peak was identified by MALDI. The area of the peak was converted into mass using standard curves.

The data in Figure 3.2A are presented as a plot of cross-linked dimer fraction as a function of peptide/DPC ratio, fitted by IGOR Pro (Wavemetrics) using functions derived from equilibrium scheme 1 (see main text), resulting in minus log of the association constant ( $pK_{\text{dim}}$  and  $pK_{\text{ox}}$ ) for all MS1/variants.

The experimental observables are the concentrations of the reduced and oxidized species. The HPLC assay does not distinguish between  $\text{Mon}_{\text{SH}}$  and  $\text{Dim}_{\text{SS}}$  so the total concentration of reduced species,  $[\text{Pep}_{\text{SH}}]$ , is given by the sum of  $[\text{Mon}_{\text{SH}}]$  and  $2[\text{Dim}_{\text{SH}}]$ . Glutathione adducts are also observed, but are generally much lower in concentration, and can be neglected because the equilibria of interest required to compute  $K_{\text{dim}}$  and  $K_{\text{ox}}$  involves only the concentration of the reduced species ( $\text{Mon}_{\text{SH}}$  plus  $\text{Dim}_{\text{SH}}$ ) as well as  $\text{Dim}_{\text{SS}}$  at equilibrium.

$$P_T = [\text{Pep}_{\text{SH}}] + 2[\text{Dim}_{\text{SS}}] \quad \text{eqn: S1a}$$

$$P_T = [\text{Mon}_{\text{SH}}] + 2[\text{Dim}_{\text{SH}}] + 2[\text{Dim}_{\text{SS}}] \quad \text{eqn: S1b}$$

$$P_T = [\text{Mon}_{\text{SH}}] + [\text{Mon}_{\text{SH}}]^2 K_{\text{dim}} + [\text{Mon}_{\text{SH}}]^2 K_{\text{dim}} K_{\text{ox}} [\text{Glut}_{\text{SS}}]/[\text{Glut}_{\text{SH}}]^2 \quad \text{eqn: S1c}$$

Equation S1c was solved numerically for  $[\text{Mon}_{\text{SH}}]$  as a function of  $[P_T]$ ,  $K_{\text{dim}}$ ,  $K_{\text{ox}}$ , and  $[\text{Glut}_{\text{SS}}]/[\text{Glut}_{\text{SH}}]^2$  using the root-finding algorithm in IGOR Pro.  $[P_T]$  and  $[\text{Glut}_{\text{SS}}]/[\text{Glut}_{\text{SH}}]^2$  are experimentally determined quantities, leaving only  $K_{\text{dim}}$  and  $K_{\text{ox}}$  as dependent variables. The data are then expressed as a plot of fraction of the peptide in the

disulfide form ( $\text{frac} = 2[\text{Dim}_{\text{SS}}]/\text{P}_{\text{T}}$ ), and the values of the dependent variables obtained by non-linear least squares fitting to the equation:

$$\text{frac} = \{[\text{MonSH}]^2 K_{\text{dim}} \cdot K_{\text{ox}} \cdot [\text{Glut}_{\text{SS}}]/[\text{Glut}_{\text{SH}}]^2\} / \text{P}_{\text{T}} + \text{C}$$

in which C is a constant, which was found to be close to zero for all peptides.

### **Disulfide formation between MS1-Gly and MS1-Gly-(Ct)**

Also, in order to measure the orientation of Gly-variant dimer, 12.5 $\mu\text{M}$  MS1-Gly and 12.5 $\mu\text{M}$  MS1-Gly-(Ct) (see Figure 3.1 in main text) were incorporated into DPC in a ratio of 1:100. Disulfide cross-linking experiments were applied for 24 hours with GSSG/GSH ratio of 1:4 and total concentration of 1.5 mM, followed by analytical HPLC.

### **Analytical ultracentrifugation**

Sedimentation equilibrium experiment and data analysis were described as in refs.<sup>82, 141</sup> The experiments were conducted at 25 °C using a Beckman XL-I analytical ultracentrifuge at respectively 30, 35, 40, 45, 48 KRPM. 200  $\mu\text{M}$  of each MS1 variant was incorporated into 20 mM DPC, in a buffer containing 100mM Tris-HCl (pH 8.6), 100 mM KCl, 1 mM EDTA. Two groups of sample preparation were employed for different purposes.



In order to measure the oligomerization of reduced MS1 variants, 1mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was added to buffer to keep the peptides reduced and buffer containing 48% D<sub>2</sub>O is used to density match the detergent. Additional multiple conditions of MS1-Gly were run at 333 μM peptide in 20 mM DPC and 100 μM peptide in 20 mM DPC in order to make sure MS1-Gly adopts a fully dimer conformation.

In order to measure the orientation of MS1 variants, samples were air-oxidized overnight before sedimentation. Buffer containing 48% D<sub>2</sub>O is used to density match the detergent. The oxidation percentage is quantitatively measured (70±5)% by analytical HPLC after sedimentation.

Data obtained were globally fitted by nonlinear least-squares curve by IGOR Pro (Wavemetrics) as previously described.<sup>157</sup> Peptide partial specific volumes and the molecular mass calculated for 48% D<sub>2</sub>O exchange were calculated using previously described methods.<sup>187</sup> The solvent density (1.0621 g/ml) and aqueous solution molar extinction coefficients at 280 nm were calculated using program Sednterp. These coefficients were multiplied by the molar detergent concentration to provide molar ratio concentration units. All these values were kept constant during global fitting.

## Computational modeling.

The original MS1 peptide was designed using GCN4's backbone as a template, which is a parallel coiled coil.<sup>60</sup> In this work, we model the MS1 variants using a coiled coil description of the backbone. To generate the coiled coil backbones, we used Crick's equations.<sup>188</sup> For the parallel coils, the super-helical radius and the  $\alpha$ -helical phase were varied to generate the model backbones. For anti-parallel coils, the super-helical radius, the  $\alpha$ -helical phase and the z-Translation were varied to generate the model backbones. For coiled coils of this size, some of the degrees of freedom can be held at constant values.<sup>161</sup> The pitch was set to 190 Å. The rise per residue was set to 1.51 Å. The  $\alpha$ -helical radius was set to 2.25 Å. The  $\alpha$ -helical frequency was set to 102° (360° / 3.5 residues per turn). For computational efficiency, we modeled the MS1 sequence starting at the residue prior to the first variable **a** position and including a total of 16 residues (Figure 3.1). This is the most hydrophobic region, which should correspond to the transmembrane portion of MS1.

| Degree of Freedom                  | Sampled Ranges            |                             |
|------------------------------------|---------------------------|-----------------------------|
|                                    | Parallel                  | Antiparallel                |
| super helix radius (r)             | 2.5 – 8 Å; steps of 0.1 Å | 2.5 – 7.5 Å; steps of 0.1 Å |
| $\alpha$ -helix phase ( $\theta$ ) | -180° – 180°; steps of 5° | -180° – 180°; steps of 5°   |
| Z-translation (z)                  | N/A                       | -3 – 3 Å; steps of 0.1 Å    |

To complete the models of each MS1 variant, the side-chains needed to be built on the backbone. Each side chain, except for alanine and glycine, was modeled using 30 conformations from an energy-ranked rotamer library.<sup>189</sup> The energy of each side-chain

conformation with the backbone and each pair of side-chain conformations was computed using an implementation of the CHARMM force field,<sup>190</sup> including IMM1 membrane solvation.<sup>148</sup> To select the proper rotamer at each position, the CHARMM energy table was processed by an Integer Linear Programming (ILP) optimization algorithm.<sup>191</sup> Our C++ implementation of the algorithm uses the Gnu Linear Programming Toolkit (GLPK)<sup>192</sup> to solve for the global minimum energy configuration (GMEC) of side-chains. Lastly, each model was minimized using CHARMM. The minimization was run for 1000 steps, using the Adopted Basis Newton-Raphson algorithm and restrained by placing harmonic forces on the backbone C $\alpha$  atoms.

In order to evaluate the association energy, the energy of an isolated, ideal and independently repacked helix was used as a reference, unbound state. The single helix for each MS1-variant sequence was modeled using the exact same protocol as stated above.

$$E_{association} = E_{complex} - 2E_{unbound}$$

To understand further the role of electrostatics in membrane proteins, we computed the contribution of the interfacial residues backbone atoms (the residues that are varied in this study) in both the winning parallel and anti-parallel models. Because we are interested only in the rank-ordering of the electrostatic contributions, the dielectric was set to 1. All MS1-variant sequences preferred their anti-parallel model for electrostatics

(based on the minimum-energy conformations of the parallel and antiparallel conformations). However with larger residues at the interface the difference had a marginal preference (Figure 3.9).

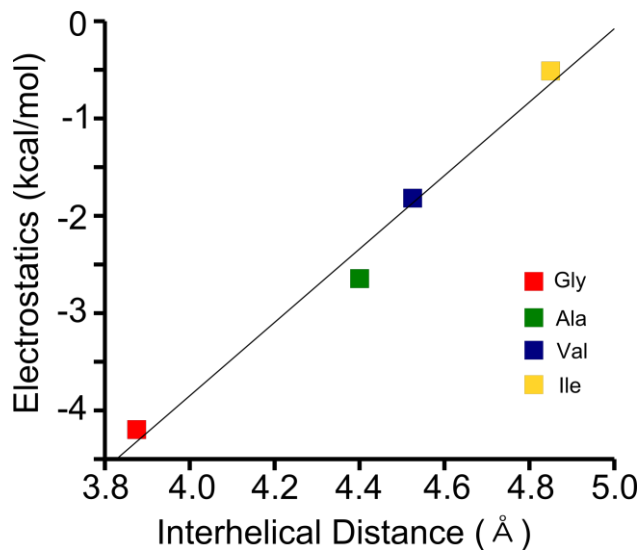


Figure 3.9 Electrostatics contribute to the orientation preference of MS1 variant models. The difference between electrostatic energy in the parallel and anti-parallel models correlates with inter-helical distance (average of parallel and anti-parallel distances). The glycine model shows the largest electrostatic energy difference between parallel and anti-parallel orientations.

### **3.6 Contributions**

Yao Zhang performed all the experiments and related analysis in this chapter, Daniel W. Kulp performed the computational modeling.

This chapter has been published in *Journal of the American Chemical Society*, 2009:131(32):11341-3.

## **Chapter 4: pH-switchable peptide model for membrane insertion and pore formation**

### **4.1 Abstract**

As therapeutic potentials for macromolecules, like peptides and proteins, are increasingly characterized, there is a need to develop the means to deliver them in to the cytosolic environment (where these agents carry out the expected functions) across the hydrophobic barrier of membrane. Efforts to develop a variety of intracellular drug delivery systems as viral vector, lipoplexes, nanoparticles and amphiphilic peptides have been made, but various challenges as delivery efficiency, toxicity and controllability remain to be overcome. Here we have designed and characterized a series of pH-switchable pore formation (PSPF) peptides as a potential delivery agent. Successful low pH-triggered, PSPF-mediated release specific for ATP and miRNA from red blood cells was characterized. Furthermore, various biophysical studies (Trp fluorescence, CD, SEC, AUC and ATR-FTIR) show that the decreased pH destabilizes the PSPF stability in aqueous systems while promoting their membrane insertion. Together, these results suggest a model that reduced pH drives PSPF to insert into membrane, leading to target-specific escape through pore formation.

## 4.2 Introduction

The ability to introduce targeted substances into a cell's interior would greatly enhance our ability to interface with cellular processes.<sup>160, 193</sup> Though for a small class of molecules cellular uptake can be spontaneous, the general task, known as the delivery problem, is largely unsolved.<sup>193, 194</sup> This is because biological membranes serve as effective barriers that prevent most substances from freely flowing into and out of cells and between organelles. To allow flux of desired target, organisms depend on membrane-inserted protein channels and transporters. Thus a potential solution to the delivery problem is via engineering of custom channels or transporters.

In nature, a common feature of these carrier proteins is their controllability. A channel or transporter responsible for the flux of an important molecule can generally be activated or inactivated by the cell as needed.<sup>195</sup> For example, channel-forming toxin peptides, found in each of the three domains of life, generally become active after a proteolytic cleavage event.<sup>196</sup> This is also a desirable feature in engineered carrier proteins as controlled delivery could lead to targeted delivery in pharmaceutical applications.

Here we aim to design peptides that bind to biological membranes and form pores only at low pH, but are minimally interactive at high pH. The pores can serve as channels for transport of appropriately-sized target, while the pH switch provides a convenient

manner in which to control the activity. Further, because of the lower pH environment in the endosome, the uptake of such peptides by endocytosis could allow endosomal escape of material present in the extracellular environment into the cell.

To realize this pH-switchable behavior, we considered three thermodynamic states in our design process (Figure 4.1). At high pH, the peptide should be “stored” in a water-soluble form that does not interact with the membrane. A good way to encode this is to assure the formation of a stable water-soluble helical bundle at high pH. Lowering of pH should destabilize this state, allowing peptide monomers to interact with the membrane. Here we consider either a surface-adsorbed form, in which helical monomers are engaged with the membrane surface, or a fully inserted state capable of forming a channel. Because insertion and channel formation are thermodynamically linked, the relative stability of the inserted versus surface-adsorbed states will have a concentration dependence, with higher peptide concentrations favoring insertion and channel formation.

To minimize membrane association at high pH, the water-soluble bundle should be very stable and its exterior should interact more favorably with water than the membrane at these conditions. The most hydrophobic and potentially membrane-interacting region of the peptide is buried in the core in this state. At low pH, both of these factors ideally need to be reversed - the stability of the water-soluble bundle should decrease, producing a population of dissociated monomers poised to interact with the



membrane, while the hydrophobicity of the peptide (and thus its preference to interact with the membrane) should increase.

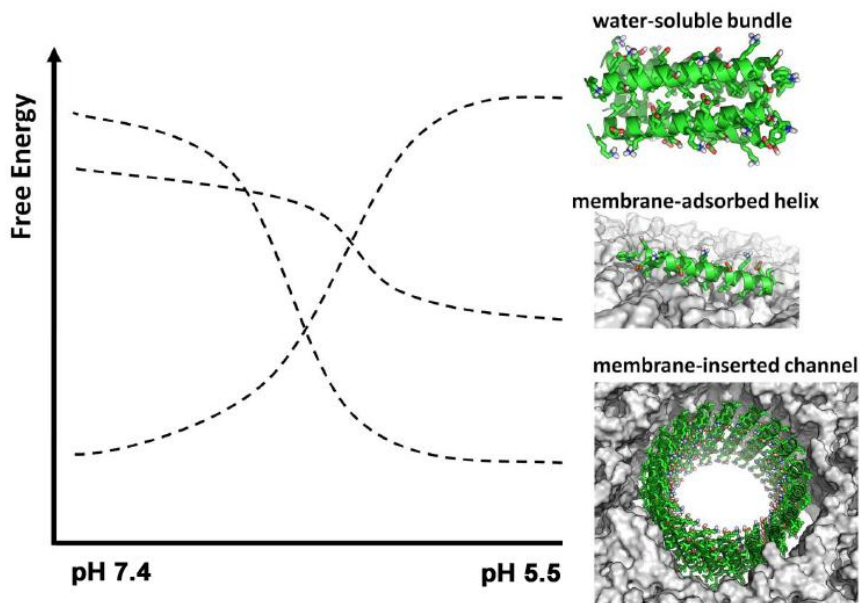


Figure 4.1 The desired free energy diagram of the designed peptide as a function of pH. Lowering pH should destabilize water-soluble bundle state and stabilize first membrane-associated monomeric state and then, in a concentration-dependent manner, the membrane-inserted channel state.

We achieve this pH modulation of stability and hydrophobicity by including amino acids in the peptide sequence whose charge state and hydrophobicity are pH-

dependent, such as Asp, Glu and His, and considering the stability of the water-soluble coiled coil-like bundle. In addition, we must also consider the specific inter-residue interactions of the membrane-inserted pore in selecting the design sequence, as we are interested in stabilizing a specific pore-forming state at low pH, rather than simply ensuring membrane insertion. For example, peptides that simply insert into membranes or those that insert and form indiscriminately large pores or even cause lysis are abundant in nature,<sup>197</sup> but would constitute unsuccessful endpoints of our design efforts either because of lack of pore formation or potential toxicity. Thus, the overall design procedure combined the use of pH-switchable residues with the consideration of inter-residue contacts and stabilities of both the water-soluble as well as membrane-inserted pore states.

Here, we present the design and characterization of a series of pH-switchable pore formation (PSPF) peptides. Several of the designed peptides associate with the membrane in a pH dependent manner. The most promising peptides also show features of pH dependent pore formation.

### 4.3 Results

#### Rational selection of amino-acid choices

The design goal was to create a water-soluble peptide that associates into a stable coiled-coil bundle at high-to-neutral pH, while preferring a membrane-inserted channel state at low pH. This means that upon pH decrease, the nonpolar residues facing inward in the soluble bundle, should invert and face the lipid phase in the membrane-inserted channel (Figure 4.2). Since canonical coiled coils have only seven environmentally distinct positions, referred to as the heptad and designated with letters **a** through **g** (Figure 4.2A), we focused on choosing the appropriate amino-acids for each of these seven sites. Furthermore, each site in our design must play two roles - stabilizing the water-soluble, “hydrophobic-inside” state at high pH and the membrane channel, “hydrophobic-outside” state at low pH. To impart stability on the water-soluble bundle, we chose to adhere to the canonical Leu-zipper coiled-coil motif, meaning that coiled-coil positions **a** and **d** were set to Leu. These same residues face the lipid phase in the membrane channel state, and Leu residues are ideal for this task as well (Figure 4.2B). The solvent-exposed **b**, **c**, and **f** positions in the water-soluble bundle should be polar to impart solubility and fold specificity, and these can also be used to modulate bundle stability through their innate helix propensities. In the membrane-channel state, these positions are also water-facing, as they point into the center of the channel, so their polar nature is appropriate here as well. However, unlike in the water-soluble state, **b** and **c** positions are also located at the inter-helical interface of the channel. Thus, the importance of these positions goes beyond their physico-chemical character and includes potential interactions stabilizing specific

interfacial conformations of channel helices. The inter-helical geometry in the channel state is important as it ultimately defines the shape and even size of the entire channel.

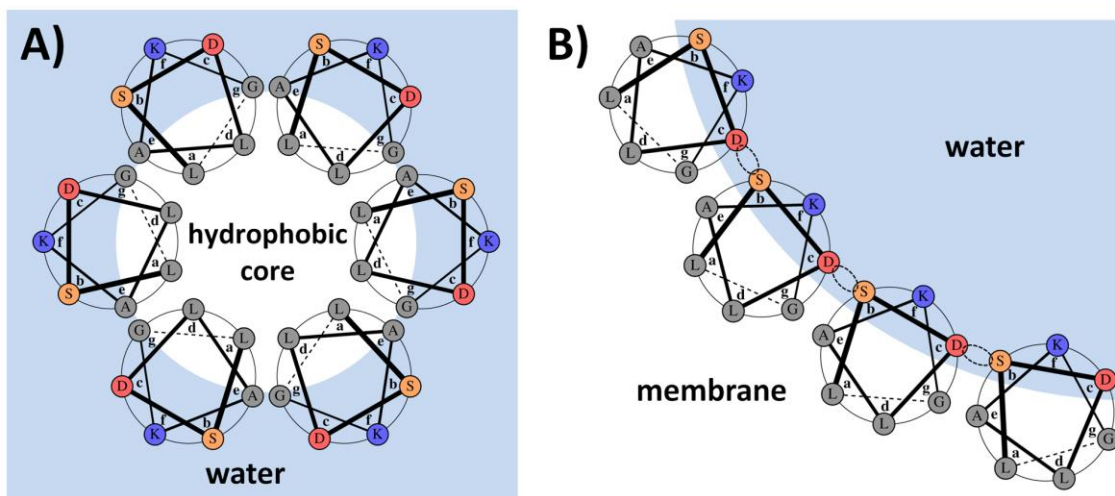


Figure 4.2 The design concept illustrated using one of the designed sequences (PSPF-DKG). Hydrophobic residues are either lining the bundle the core in the water-soluble state (A), or are facing the lipid membrane in the membrane channel state (B). Dotted circles illustrate potential hydrogen bonding in the channel state. Heptad positions in both panels are labeled according to the water-soluble state. The amino acid choices at each position are shown in Table 4.1.

At the **f** position, we chose to consider Lys or Gln - amino acids with favorable helix propensities<sup>198</sup> that are common at this position in coiled coils. At position **b** we considered Ser because of its polar nature and relatively high helix propensity, as well as

due to its high preponderance in closely-packing helix-helix interfaces in TM proteins.<sup>75</sup> The **c** position was chosen as the pH-sensing switch. We considered amino acids Glu and Asp at this position as their protonation state is dependent on pH, causing them to be more protonated, less charged and thus more hydrophobic at lower pH. Although the pK<sub>a</sub> of the carboxylic side-chains groups of Glu and Asp in water are around 4.0, somewhat lower than the typical endosomal pH or ~5.5, significant shifting in protonated populations would still be expected relative to neutral pH, and the collective effect of having multiple closely-spaced acidic groups on one face of a helix will likely increase the effective pK<sub>a</sub> of the side-chains. An additional significance of Glu and Asp residues is their potential ability to participate in inter-helical hydrogen bonding (see Figure 4.2B), thus further dialing in a specific, closely packed inter-helical geometry in the membrane-channel state (see also structural modeling below). As a way of testing the importance of the pH switch residue, we also considered the amino acid His at the **c** position. The side-chain of His titrates at pH ~6.1, but it is more charged at acidic pH than at neutral pH. Because of this reversed pH sensitivity compared to Asp and Glu, His provides a convenient point of reference.

Positions **e** and **g** are located along the helix-helix interface in both the water-soluble and the membrane-channel states. Because the primary driver of the water-soluble bundle stability is the canonical leucine-zipper motif, we opted to choose small

hydrophobic residues at **e** and **g** with the primary purpose of stabilizing a closely-packed TM helical interface<sup>75</sup>.

Table 4.1 Amino-acid choices considered in design

| Position in water | Function in water, high pH           | Position in membrane | Function in membrane, low pH   | Amino acid Choice |
|-------------------|--------------------------------------|----------------------|--|-------------------|
| a                 | Helical bundle hydrophobic core      | b                    | Membrane-facing.   | Leu               |
| b                 | Solvent-exposed, imparts solubility. | d                    | Small residue for helical interface, potential inter-helical hydrogen bonding.   | Ser               |
| c                 | Solvent-exposed, imparts solubility. | e                    | Trigger residue, changes protonation state/ hydrophobicity at low pH. Potential inter-helical hydrogen bonding.                | Asp, Glu, His     |
| d                 | Helical bundle hydrophobic core.     | f                    | Membrane-facing.   | Leu               |
| e                 | Modulation of helical propensity.    | g                    | Small residue for helical interface.   | Ala               |
| f                 | Solvent-exposed, imparts solubility. | a                    | Solvent-exposed in channel state (inner channel lining). Imparts folds specificity by encoding helical orientation preference. | Lys, Gln          |
| g                 | Modulation of helical propensity.    | b                    | Small residue for helical interface.   | Ala, Gly          |

Based on the criteria list above, a group of sequences has been generated for the pH-switchable pore formation (PSPF) peptide (Table 4.2).

Table 4.2 The sequence of PSPF peptides

| Peptide         | Sequence          |                   |                   |                   |
|-----------------|-------------------|-------------------|-------------------|-------------------|
| Heptad membrane | <b>cdefgab</b>    | <b>cdefgab</b>    | <b>cdefgab</b>    | <b>cdefgab</b>    |
| PSPF-DQA        | WSD <u>DLAQA</u>  | LS <u>DLAQA</u>   | LS <u>DLAQA</u>   | LS <u>DLAQA</u>   |
| PSPF-DQG        | WS <u>DLAQG</u>   | LS <u>DLAQG</u>   | LS <u>DLAQG</u>   | LS <u>DLAQG</u>   |
| PSPF-DKA        | WSD <u>DLAKA</u>  | LS <u>DLAKA</u>   | LS <u>DLAKA</u>   | LS <u>DLAKA</u>   |
| PSPF-DKG        | WSD <u>DLAKG</u>  | LS <u>DLAKG</u>   | LS <u>DLAKG</u>   | LS <u>DLAKG</u>   |
| PSPF-EQA        | WSE <u>ELAQ</u> A | LSE <u>ELAQ</u> A | LSE <u>ELAQ</u> A | LSE <u>ELAQ</u> A |
| PSPF-EQG        | WSE <u>ELAQG</u>  | LSE <u>ELAQG</u>  | LSE <u>ELAQG</u>  | LSE <u>ELAQG</u>  |
| PSPF-EKA        | WSE <u>ELAKA</u>  | LSE <u>ELAKA</u>  | LSE <u>ELAKA</u>  | LSE <u>ELAKA</u>  |
| PSPF-EKG        | WSE <u>ELAKG</u>  | LSE <u>ELAKG</u>  | LSE <u>ELAKG</u>  | LSE <u>ELAKG</u>  |
| PSPF-HQA        | WS <u>HLAQA</u>   | LS <u>HLAQA</u>   | LS <u>HLAQA</u>   | LS <u>HLAQA</u>   |
| PSPF-HQG        | WS <u>HLAQG</u>   | LS <u>HLAQG</u>   | LS <u>HLAQG</u>   | LS <u>HLAQG</u>   |
| PSPF-HKA        | WS <u>HLAKA</u>   | LS <u>HLAKA</u>   | LS <u>HLAKA</u>   | LS <u>HLAKA</u>   |
| PSPF-HKG        | WS <u>HLAKG</u>   | LS <u>HLAKG</u>   | LS <u>HLAKG</u>   | LS <u>HLAKG</u>   |
| Heptad Water    | <b>abcdefg</b>    | <b>abcdefg</b>    | <b>abcdefg</b>    | <b>abcdefg</b>    |

In the column of peptide names, the heptad position for the membrane-soluble form is labeled after the one-letter amino acid code at top, and the heptad position for the water-soluble form is label after the one-letter amino acid code at bottom.

## Cellular release assays

Red blood cell (RBC) lysis assays have been used to screen the functional efficacy of the peptides upon delivery (Table 4.3). The release of ATP, miRNA and hemoglobin has been studied at both pH 7.5 and 5.4. The peptide is designed to selectively deliver the nucleotides or ribonucleic acid, with sizes similar to ATP and miRNA, across the membrane only at pH5.5. The desirable peptide should also negate membrane disruption, as assessed by leakage of proteins such as hemoglobin at both pHs. Therefore the peptides were first screened for hemolytic activity at both pH 7.5 and pH 5.4. None of the twelve peptides had hemolytic activity at either pHs. When screening for ATP and miRNA release at 5 $\mu$ M, PSPF-DQA, PSPF-DKG, and PSPF-EKG showed relatively high release percentage for ATP (more than 20%) and miRNA (more than 10%) at pH 5.4, and also low release percentage at pH7.5 for both ATP and miRNA (less than 10%). Among the top three peptides screened out of RBC assays, PSPF-EKG has been further characterized to reveal the mechanism of action.



Table 4.3 RBC Lysis assay of PSPF peptides

| Peptide  | RBC Lysis Assay (% calculated compared to triton-x-100) |       |                   |       |                     |       |
|----------|---|-------|-------------------|-------|---------------------|-------|
|          | Hemoglobin Release                                      |       | %ATP at 5 $\mu$ M |       | %miRNA at 5 $\mu$ M |       |
|          | pH7.5   | pH5.4 | pH7.5             | pH5.4 | pH7.5               | pH5.4 |
| PSPF-DQA | none  | none  | 3.81              | 17.91 | 0.81                | 18.61 |
| PSPF-DQG | none  | none  | 3.21              | 8.61  | 0.06                | 0.16  |
| PSPF-DKA | none  | none  | 4.64              | 5.79  | 4.13                | 0.79  |
| PSPF-DKG | none  | none  | 7.54              | 24.1  | 5.54                | 7.47  |
| PSPF-EQA | none  | none  | 3.69              | 3.61  | 0.46                | 0.02  |
| PSPF-EQG | none  | none  | 3.36              | 9.52  | 0.15                | 2.64  |
| PSPF-EKA | none  | none  | 2.02              | 3.66  | 0.51                | 0.21  |
| PSPF-EKG | none  | none  | 3.38              | 27.3  | 0.14                | 12.54 |
| PSPF-HQA | none  | none  | 6.17              | 11.72 | 0.02                | 1.43  |
| PSPF-HQG | none  | none  | 5.69              | 10.55 | 0.2                 | 0.64  |
| PSPF-HKA | none  | none  | 0.93              | 5.7   | 0.43                | 0.02  |
| PSPF-HKG | none  | none  | 32.1              | 39.22 | 72.28               | 0.44  |

## **Peptide engagement with the lipid bilayer by tryptophan fluorescence**

To detect the engagement of peptides with lipid vesicles, tryptophan (Trp) fluorescence was measured for PSPF-DQA, DKG and EKG. The extent of environmental change around the N-terminal Trp was determined by the observed shift and changes in intensity of the fluorescence signal. Blue shifts (Table 4.3), correspond to a more hydrophobic environment, such as that which would occur to the Trp upon membrane interaction or insertion. The majority of the PSPF- peptides studied showed minimal blue shifting at pH 7.4 and larger shifts at pH 5.5 (Table 4.4). PSPF-DQA shows small detectable shift at pH 5.5 (-1 nm), whereas PSPF-DKG and EKG show blue shifts of approximately 3 nm (350 to 347 nm) each at pH 5.5. PSPF-HKG also showed a significant shift from 351 to 341 nm at pH 5.5.

Despite different experimental conditions, Trp fluorescence shifts among all the peptides correlate strongly with ATP release at pH 5.5, with  $R^2$  of 0.74 if linear regression is applied (Figure 4.3). At pH 5.5, a larger shift in Trp fluorescence (likely due to insertion into the membrane of Trp) corresponds to greater release of ATP (likely from membrane insertion and pore formation). This suggests that the peptides are acting in a similar manner in both experimental assays and consistent with pH-sensitive insertion and pore formation.

Table 4.4 Trp fluorescence of PSPF- series peptides with various amounts of lipid vesicles.

| Peptide<br>PSPF- | pH7.4                 |                   |  |                          | pH5.5                 |                   |                                 |                         |
|------------------|-----------------------|-------------------|--|--------------------------|-----------------------|-------------------|---------------------------------|-------------------------|
|                  | $\lambda_{\max}$ (nm) |                   | $\Delta \lambda_{\max}$<br>(nm) <sup>#</sup> | % Intensity<br>Increase* | $\lambda_{\max}$ (nm) |                   | $\Delta \lambda_{\max}$<br>(nm) | % Intensity<br>Increase |
|                  | 0 $\mu$ M Lipid       | 200 $\mu$ M Lipid |  |                          | 0 $\mu$ M Lipid       | 200 $\mu$ M Lipid |                                 |                         |
| DQA              | 352                   | 351               | -1   | 32                       | 348                   | 347               | -1                              | 38                      |
| DQG              | 354                   | 353               | -1   | 18                       | 350                   | 358               | -2                              | 33                      |
| DKA              | 354                   | 353               | -1   | 18                       | 349                   | 358               | -1                              | 38                      |
| DKG              | 355                   | 351               | -4   | 36                       | 350                   | 347               | -3                              | 38                      |
| EQA              | 352                   | 352               | 0  | 6                        | 349                   | 348               | -1                              | 21                      |
| EQG              | 355                   | 354               | -1   | 15                       | 349                   | 346               | -3                              | 42                      |
| EKA              | N/A                   | N/A               | N/A  | N/A                      | N/A                   | N/A               | N/A                             | N/A                     |
| EKG              | 354                   | 352               | -2   | 28                       | 350                   | 347               | -3                              | 52                      |
| HQA              | N/A                   | N/A               | N/A  | N/A                      | N/A                   | N/A               | N/A                             | N/A                     |
| HQG              | N/A                   | N/A               | N/A  | N/A                      | N/A                   | N/A               | N/A                             | N/A                     |
| HKA              | N/A                   | N/A               | N/A  | N/A                      | 349                   | 347               | -2                              | 34                      |
| HKG              | N/A                   | N/A               | N/A  | N/A                      | 351                   | 341               | -10                             | 72                      |

#  $\Delta \lambda_{\max} = \lambda_{\max}$  at 200  $\mu$ M Lipid -  $\lambda_{\max}$  at 0  $\mu$ M Lipid

\*% Intensity Increase = (Intensity at 200  $\mu$ M Lipid - Intensity at 0  $\mu$ M Lipid) / Intensity at 0  $\mu$ M Lipid

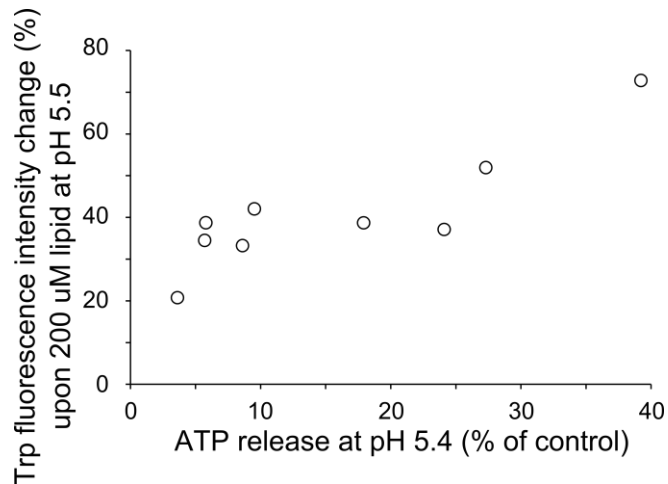


Figure 4.3 Correlation between ATP-release by PSPF peptides and the degree of lipid engagement as assessed by the fractional change of Trp-fluorescence signal upon addition of 200  $\mu$ M lipid vesicles.

## **The association properties of PSPF peptides in an aqueous system**

### *Size exclusion chromatography*

The association state of the designed peptide PSPF-EKG was initially investigated by size exclusion chromatography (SEC)<sup>46</sup> using a Superdex 75 column (GE Healthcare) eluted at pH 7.4 (150 mM NaCl, 50 mM Tris) and pH 5.5 (150 mM NaCl, 50 mM MES) respectively. In addition, PSPF- DKG was also investigated to determine the effect of substituting Asp for Glu on the stability of the water-soluble bundle at each pH. To determine the approximate oligomerization states, four standards were used, shown by blue eluting peaks in Figure 4.4: blue dextran (2,000,000 g/mol), carbonic anhydrase (29,000 g/mol), cytochrome C (12,400 g/mol) and aprotinin (6,500g/mol).

PSPF-EKG eluted with an apparent molecular weight 6.5-fold higher than the calculated molecular weight at pH 7.4 and 5.2-fold at pH 5.5 (Figure 4.4, Table 4.5), both as a single species. Noticeably PSPF-EKG presented a peak with significantly lower intensity and a broad trailing feature when eluting at pH 5.5, indicating that the decreased pH has increased the propensity to interact with column, which may act as a mimic of the membrane phase (Figure 4.4A). Dissociation during elution might also contribute to the peak shape, indicative of a lower stability of the water-soluble helical bundle. Similarly, PSPF-DKG eluted with an apparent molecular weight 6.0-fold higher than the calculated molecular weight at pH 7.4 as a single species and nearly failed to elute at pH5.5 (Figure 4.4A), indicating the lower pH drove the peptide to interact with the column. Furthermore,

the salt concentration has been increased to 2M and the shoulder of elution peak for PSPF-EKG still exists at pH 5.5 (Figure 4.4C, D). Also, Asp at the putative **a** position makes the PSPF-DKG more sensitive to the pH decrease than Glu in PSPF-EKG, in term of driving the peptide's preference away from the aqueous phase (Figure 4.4A).

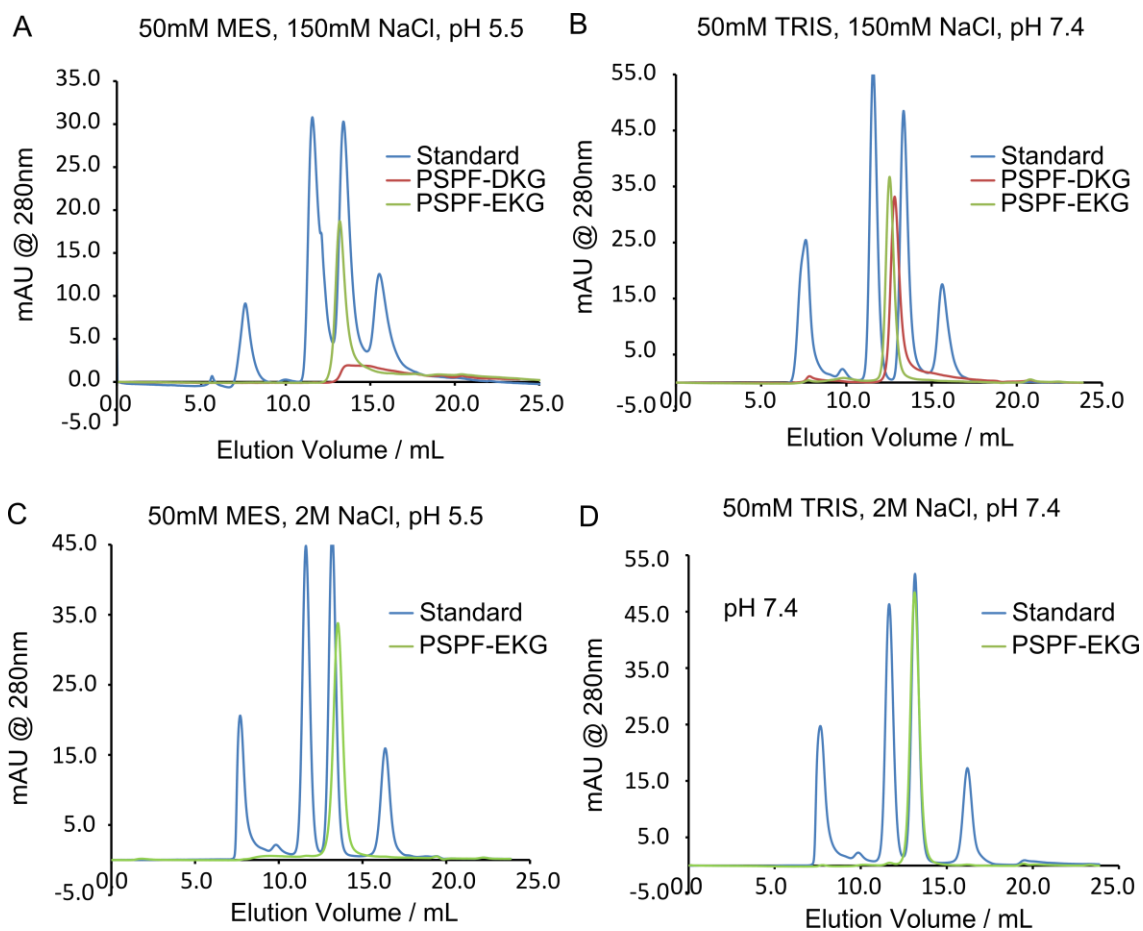


Figure 4.4 Size exclusion chromatography of PSPF-EKG and PSPF-DKG at each pH. Both PSPF-EKG and PSPF-DKG elute as a single species corresponding to the oligomerization of hexamer at pH 7.4 (B). PSPF-EKG elutes as a single-species peak with a significant shoulder at pH5.4 and the major peak corresponds to a formation of

hexamer. PSPF-DKG almost fails to elute at PH 5.5 (A). The salt concentration has been increased to 2M and the shoulder of elution peak still exists at pH 5.5 (C, D).

Table 4.5 Apparent molecular weight and calculated oligomerization state based on size exclusion chromatography for PSPF- EKG and PSPF- DKG at both PHs

|                        | PSPF-EKG |         | PSPF-DKG |        |
|------------------------|----------|---------|----------|--------|
|                        | pH 7.4   | pH 5.5  | pH 7.4   | pH 5.5 |
| Apparent MW            | 19,000   | 15,000# | 17,000   | N/A    |
| Oligomerization State* | 6.6      | 5.2     | 6.0      | N/A    |

\* Oligomerization State = Apparent MW/ Monomer MW

# Major peak

### ***Sedimentation equilibrium of analytical ultracentrifugation***

Analytical ultracentrifugation (AUC) sedimentation equilibrium<sup>46, 53</sup> was applied to further investigate the association state and affinity of the water-soluble bundles of both PSPF- EKG and PSPF- DKG. The peptides were studied at 100  $\mu$ M peptide concentration and pH 7.4 (150 mM NaCl, 50 mM Tris) or pH 5.5 (150 mM NaCl, 50 mM MES). The parameters were globally fit to data collected over multiple rotor speeds (35, 40, 45, 50 KRPM). Fitting the curve to a single MW species suggested an apparent molecular weights for PSPF- EKG of 18,000 $\pm$ 30 at pH 7.4 (Figure 4.5A) and 16,000 $\pm$ 30 at pH 5.5 (Figure 4.5B). This agrees well with the data from size exclusion chromatography and points to a hexameric association state at both pHs for PSPF-EKG. The data can be further fit to a monomer-hexamer equilibrium, resulting in an association

energy  $\Delta G$  of -6.3 kcal/mol monomer at pH 7.4 and -5.6 kcal/mol monomer at pH 5.5 (Table 4.2). Also, as shown in the plot of species weight fraction, the concentration of peptide required to associate at pH 7.4 was lower than at pH 5.5 (Figure 4.5B, D). Together it suggests that decreased pH destabilized the helix bundle of PSPF-EKG.

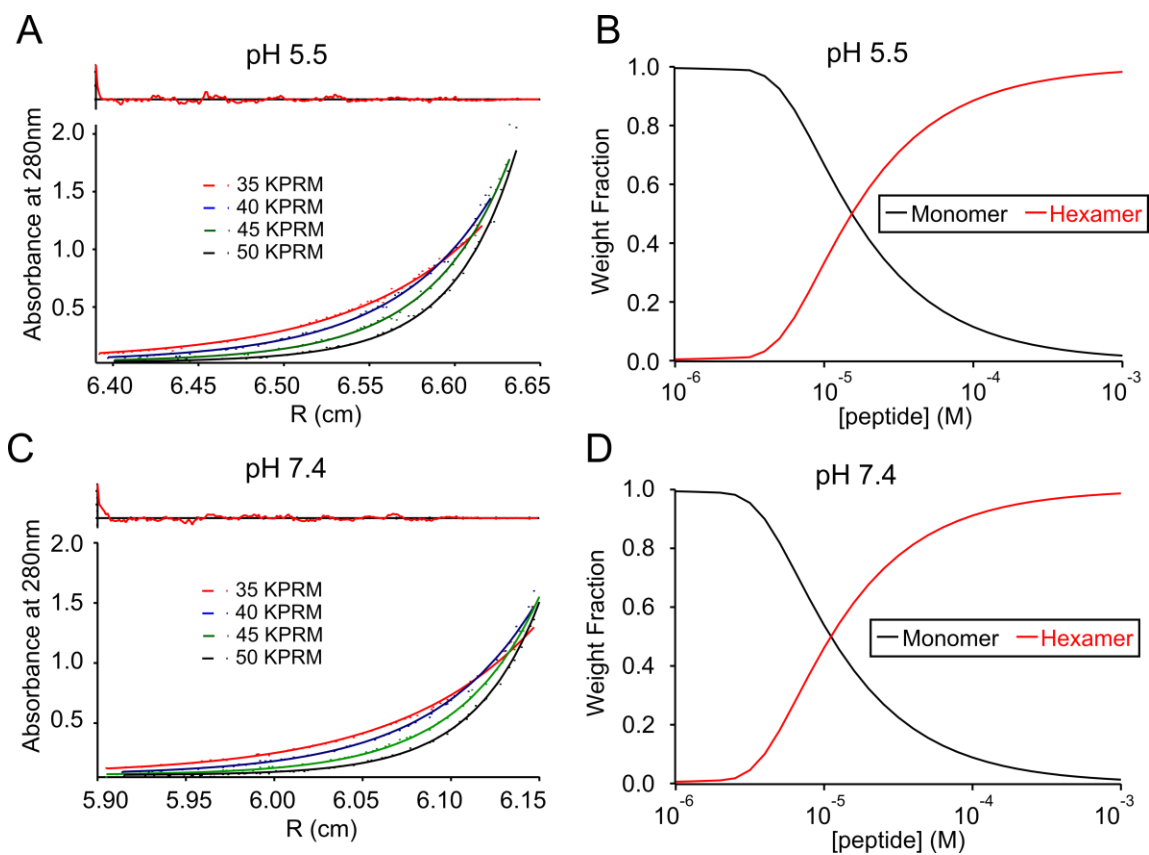


Figure 4.5 AUC sedimentation equilibrium of PSPF-EKG at pH 5.5 (A) and 7.4 (C). Single species fitting of PSPF-EKG suggests a hexameric association state at both pH 7.4 (A) and pH 5.5 (C). For each peptide and pH condition, the top plot shows the single species fitting with residuals above while the below plot shows the species weight fraction. Then the data has been fit with a monomer-hexamer equilibrium model at both

pHs. The dissociation state and dissociation energy is shown in Table 4.6. The weight fraction distributions have also been plot for pH 5.5 (B) and pH 7.4 (D).

For PSPF-DKG, a global fit resulted in a single-species apparent molecular weight of  $17,000 \pm 30$  at pH 7.4 (Figure 4.6A), which was 6.0-fold higher than the calculated molecular weight and again agrees well with size exclusion chromatography. The equilibrium of PSPF-DKG has also been fit into the equilibrium of monomer-hexamer with association energy  $\Delta G$  of  $-6.5$  kcal/mol monomer (Table 4.6). The single-species apparent molecular weight for PSPF-DKG at pH 5.5 was  $24,000 \pm 60$  (Figure 4.6B). This could represent a heterogeneous set of association states, taken together with the broad elution peak observed in the size exclusion chromatography.

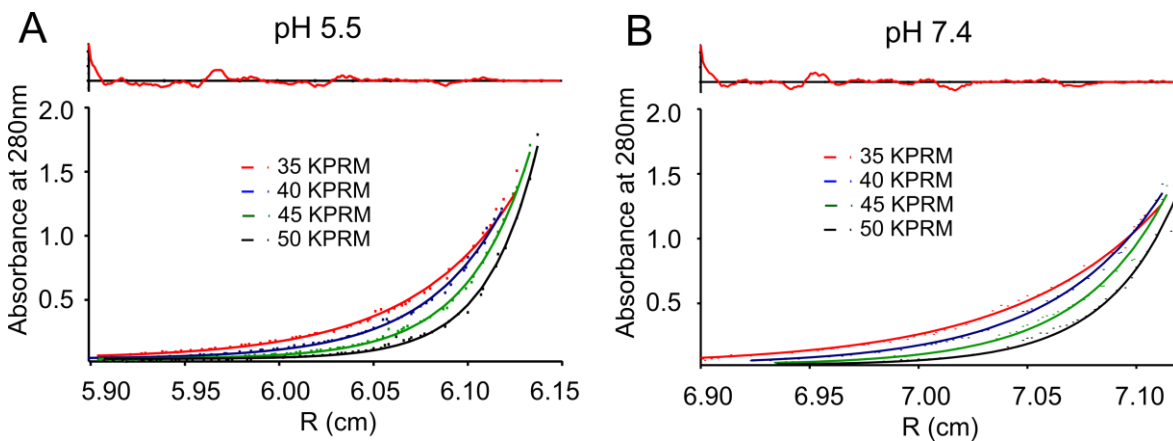


Figure 4.6 AUC sedimentation equilibrium of PSPF-DKG at pH 5.5 and 7.4. Single species fitting of PSPF-DKG suggests it associates as a hexamer at pH 7.4 and reaches an



apparent molecular weight of approximately 24,000 at pH 5.5. For each peptide and pH condition, the top plot shows the single species fitting with residuals above while the below plot shows the species weight fraction.

Table 4.6 Analytical ultracentrifugation (AUC) sedimentation equilibrium for PSPF-EKG and PSPF- DKG at pH 5.5 and 7.4.

|   | PSPF-EKG  |           | PSPF-DKG  |           |
|---|-----------|-----------|-----------|-----------|
|   | pH 7.4    | pH 5.5    | pH 7.4    | pH 5.5    |
| Apparent MW                                       | 18,000±30 | 16,000±30 | 17,000±30 | 24,000±60 |
| Oligomerization State*                            | 6.2       | 5.5       | 6.0       | N/A       |
| $-\log(K_{\text{dissociation}})$                  | 28.0±0.4  | 24.8±0.1  | 28.7±0.4  | N/A       |
| Association $\Delta G^{\#}$<br>(kCal/mol monomer) | -6.3      | -5.6      | -6.5      | N/A       |

\* Oligomerization State = Apparent MW/ monomer MW

# Association  $\Delta G = 2.303 * RT * \log(K_{\text{dissociation}}) / 6$

### ***Circular dichroism and thermal denaturing***

Circular dichroism (CD) suggests that PSPF-KEG adopts an alpha-helical secondary structure at both pHs (Figure 4.7).

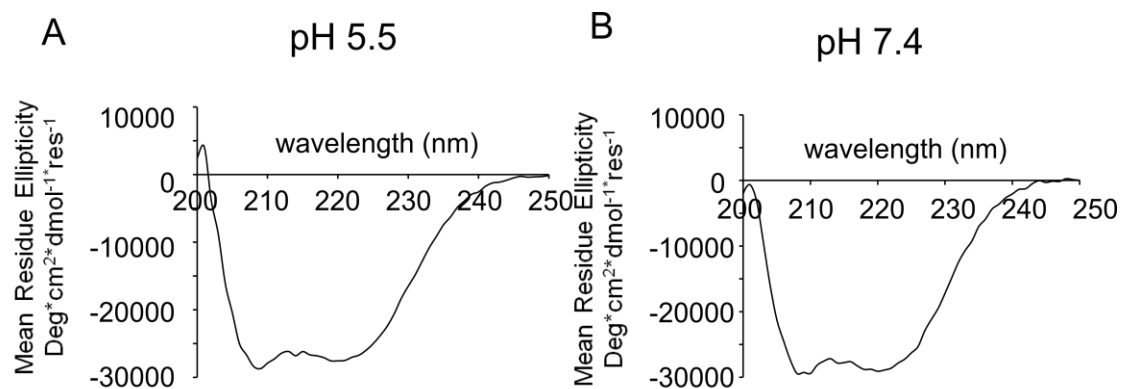


Figure 4.7 Circular dichroism of PSPF-EKG suggests an alpha-helical secondary structure at both pHs.

Furthermore, thermal denaturation by circular dichroism (CD)<sup>199</sup> was used to study the thermal stability of the PSPF- EKG hexamer at multiple concentrations (2 $\mu$ M, 4  $\mu$ M and 20  $\mu$ M), and at both pH 7.4 (Figure 4.8A) and pH 5.5 (Figure 4.8B). For each pH, to the curves were analyzed according to the Gibbs-Helmholtz Equation, using global least squares fitting of  $\Delta H_m$ ,  $T_m$  and baselines.  $T_m$  was chosen as a global parameter defined with a reference concentration of 4  $\mu$ M.  $\Delta C_p$  was also included, but over the range of experimental data examined, this parameter was not well defined.

$$\text{Gibbs-Helmholtz Equation: } \Delta G = \Delta H_m(1 - T/T_m) - \Delta C_p[T_m - T + T \ln(T/T_m)]$$

Here  $\Delta G$  refers to the unfolding energy upon thermal denaturation,  $T$  refers to temperature,  $T_m$  refers to the melting temperature at which  $\Delta G$  equals to zero.  $\Delta H_m$  refers

to the enthalpy at  $T_m$ , and  $\Delta C_p$  refers to the change in the heat capacity over the temperature range.

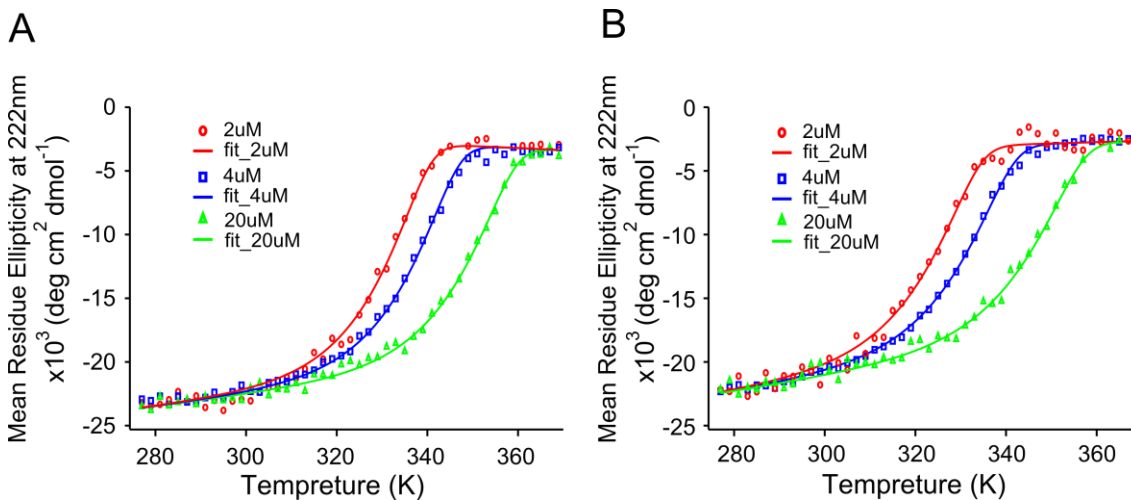


Figure 4.8 Thermal denaturation of PSPF- EKG at pH 7.4 (A) and 5.5 (B). The data are fit to the Gibbs-Helmholtz Equation.

The enthalpy at pH 7.4 is 22.0 kcal/mol monomer and is approximately 12% higher than at pH 5.5 (19.6 kcal/mol monomer) (Table 4.7). The value of enthalpy at both pHs are typical for designed water-soluble helix bundles.<sup>200</sup> The melting temperature  $T_m$  is 339.0 K at pH 7.4 and is 5.6 K higher than at pH 5.5 (333.4K). The concentration of PSPF-EKG required to have 50% of the total amount of peptide remain folded at 300K was calculated to be 0.31 $\mu$ M at pH 5.5, which was approximately double the concentration of peptide required for 50% folding at pH 7.4 (0.14 $\mu$ M). These data suggest that decrease in pH destabilizes the folding of PSPF-EKG.

Table 4.7 Fitting results for CD thermal denaturation of PSPF- EKG at both pH 7.4 and pH 5.5

| pH  | $\Delta H$ (kcal/mol monomer) | $T_m$ (K)       | [ PSPF-EKG ] at 50% fold and 300K |
|-----|-------------------------------|-----------------|-----------------------------------|
| 7.4 | 22.0 $\pm$ 0.1                | 339.0 $\pm$ 0.1 | 0.14 $\mu$ M                      |
| 5.5 | 19.6 $\pm$ 0.1                | 333.4 $\pm$ 0.1 | 0.31 $\mu$ M                      |

## **The structural properties of PSPF- peptides in a membrane micelle system**

### *Sedimentation equilibrium of analytical ultracentrifugation*

AUC sedimentation equilibrium of PSPF-EKG in detergent micelles pointed to a weak oligomerization at both pHs. PSPF-EKG was dissolved in N-tetradecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (C-14 betaine) micelles. The density of the solution was adjusted by D<sub>2</sub>O to precisely match that of the C-14 betaine detergent at both pH 7.4 and pH 5.5 (50 mM sodium phosphate and 150 mM NaCl), so that only the peptide component contributes to the sedimentation equilibrium.<sup>60, 139</sup>

Three samples prepared at different peptide-to-detergent ratios (1:50, 1:100, 1:200) were each centrifuged at four rotor speeds (35, 40, 45, 50 KMRP) at each pH. The data could be fit into a monomer-trimer, monomer-tetramer, and monomer-higher oligomer equilibrium, suggesting that PSPF-EKG weakly associates in detergent micelle. Figure 4.9 showed an example in which a monomer-trimer equilibrium is fit to the data at pH 7.4 (Figure 4.9A) and pH 5.4 (Figure 4.9C), and the weight fraction distribution has been shown in Figure 4.9B and 4.9D.

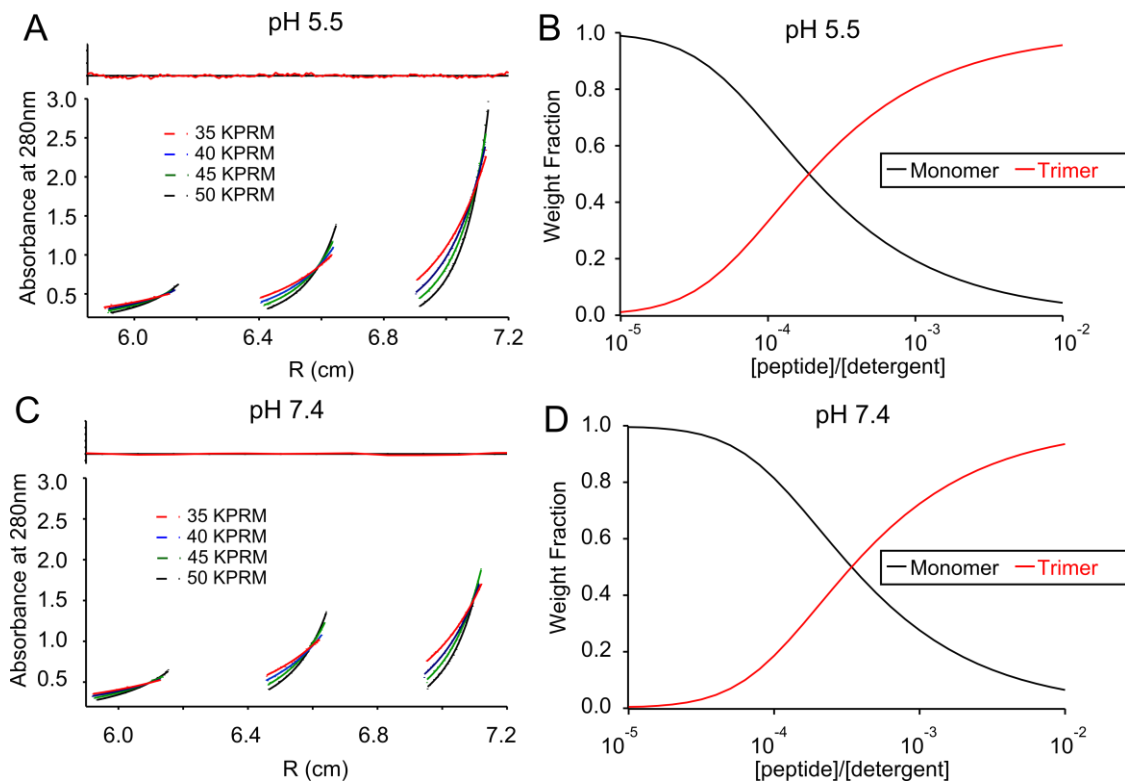


Figure 4.9 The single-species fitting of AUC sedimentation in detergent micelles for PSPF-EKG at pH 7.4 (A) and 5.5 (C). Species weight fraction of PSPF-EKG at pH 7.4 (C) and pH 5.5 (D) as the data were globally fit to a monomer-trimer equilibrium as an example.

## The Orientation of PSPF- EKG in a Lipid Bilayer

### *Attenuated total reflection IR spectroscopy*

The secondary structure and orientation of PSPF-EKG in deuterium oxide (D<sub>2</sub>O) hydrated bilayers were evaluated using attenuated total reflection IR spectroscopy (ATR-IR).<sup>143, 144, 201</sup> The IR spectra in the amide I region of the PSPF-EKG showed a single peak at 1656 cm<sup>-1</sup>, indicative of a dehydrated helical conformation in bilayers (Figure 4.10). The dichroic ratio for parallel versus perpendicularly polarized light was 1.5. This order parameter would correspond to an orientation of approximately 75° relative to the membrane normal, assuming the bilayers are well ordered and the entire peptide is fully helical. The result suggests that the majority of peptide lies parallel to the lipid surface, and rules out the possibility of the peptide being oriented predominantly perpendicular to the bilayer surface. The fact that the computed angle is less than 90° is also consistent with a small amount of peptide adopting a vertically inserted conformation, in equilibrium with the predominant form, although other models could also lead to the observed 75° angle.

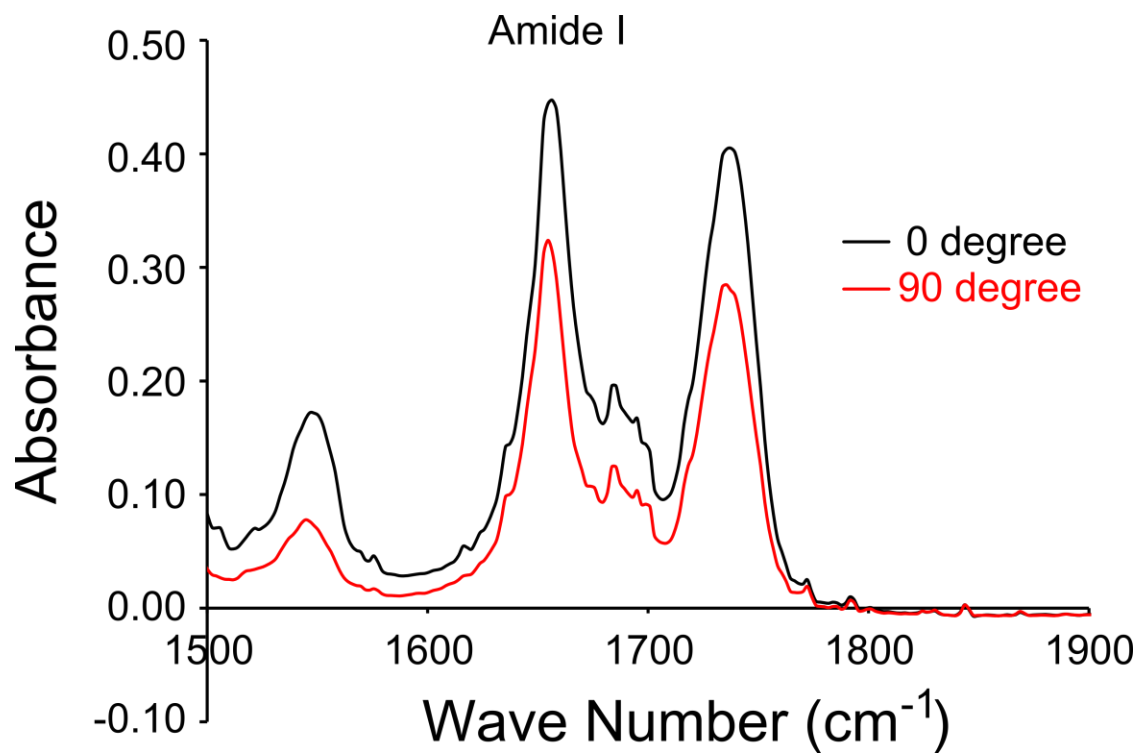


Figure 4.10 ATR-IR of PSPF-EKG in phospholipids (POPC) bilayers. The peak at 1656 cm<sup>-1</sup> is indicative of alpha helical secondary structure. The orientation is demonstrated by the ratio of peak area of the 1656 cm<sup>-1</sup> amide I bond for parallel (0°) versus perpendicular (90°) polarized incident light (relative to the membrane normal).



#### 4.4 Conclusion and Discussion

Therapeutic macromolecules such as peptides and proteins are easily cleared from the bloodstream and require assistance for intracellular delivery in order to reach their intended targets and achieve the desirable therapeutic effects. Decades of research effort have been devoted to develop delivery agents with high efficiency and low toxicity.<sup>202,</sup>

203

Viral vectors are considered to be successful delivery agents and have been extensively studied for gene therapy. Viral vector based gene therapy has demonstrated promising results,<sup>101</sup> but this potential life-saving delivery technique can also be risky. The death of a patient in the Paris trial suggests that viral vectors might also induce undesirable gene insertion and this potential danger is currently uncontrollable.<sup>204, 205</sup>

Most non-viral carriers are synthetic chemical conjugates. Active ingredients (drugs) are usually linked or enclosed into a vehicle and delivered into the cell via endocytosis or membrane fusion, or via a yet to be determined mechanism.<sup>206, 207</sup> These vehicles are typically designed as liposomes/lipoplexes,<sup>208</sup> cationic macromolecules polymer,<sup>209</sup> polypeptide,<sup>210</sup> and protein,<sup>211</sup> amphiphilic polymer/polypeptide,<sup>212</sup> nanoparticles<sup>213-215</sup> and cell penetrated peptides (CPP).<sup>83</sup> Native sequences such as fusogenic peptide from viral fusion protein<sup>216</sup> have also been manipulated as a cargo carrier to cross the barrier of cell membranes. A number of these approaches have also

entered clinical trials, but most of them reached a bottleneck due to high toxicity or lack of manipulability.<sup>207</sup>

Here, we have designed a series of pH-switchable peptides as potential candidates for intracellular (lysosomal) drug delivery. One of the top candidates, PSPF-EKG, has stood out from RBC lysis assay, in terms of highest target molecule delivery efficiency at selective pH (5.4). Lack of hemolytic activity ruled out the possibility of undesirable membrane disruption by PSPF-EKG at both pHs. Also, the nice correlation between ATP release at pH 5.5 and Trp-fluorescence at pH 5.4 upon lipid titration (Figure 4.3), indicates that membrane insertion presumably plays a key role in ATP release.

RBC Lysis data also provide a direct comparison among all the designed peptides. Firstly, we have three options of pH-trigger residues in this peptide series. Asp and Glu residues both presented expected pH-switchable ATP and miRNA release in peptides PSPF-DQA and PSPF-DKG, indicating the carboxyl side chain groups respond efficiently to environmental pH change, though their intrinsic pKa of the unperturbed sidechain is around 4. The third trigger candidate, His, failed to show significant pH preferences in terms of ATP or miRNA release. However PSPF-HKG induced high ATP release percentage at both pHs. Presumably His will induce pore formation in a pH-independent manner. Nevertheless, all the His variants ran into solubility issues in the

further biophysical characterization and thus are not considered as preferred candidates for further pharmaceutical development.

Secondly, Lys and Gln have been chosen in **f** positions in order to provide helix propensities in aqueous system and solvent exposure surface in membrane system. The RBC lysis results did not discriminate between these two residues when comparing the performance of the aspartate and glutamate peptide variants (PSPF-EKG versus PSPF-EQG, PSPF-EKA versus PSPF-EQA).

The third screened parameter is the choice between Ala or Gly for residues packed in the helix interface. This part of the design was in light of previously discovered fact that small residues were preferred in TM helix interaction interface to stabilize the final folded state (TM helix bundle).<sup>75, 139</sup> In the case of PSPF-EKG versus PSPF-EKA, Gly resulted in a much higher pH-switchable ATP and miRNA release. The results agree with the previous conclusion that Gly in TM helical interface drives stronger TM helix association than Ala,<sup>139</sup> presumably because Gly stabilizes the helix interaction via weak C $\alpha$ -H interaction.<sup>217</sup>

A variety of biophysical assays have been applied in order to obtain a comprehensive mechanism of PSPF-EKG's pH switchable pore formation. We first

looked at the structural conformation and folding stability of PSPF-EKG in aqueous solution. CD, AUC and SEC suggest that PSPF-EKG forms a stable helix bundle at both pHs (Figure 4.11A), which is expected due to the designed canonical Leu-zipper coiled-coil motif. AUC and thermal denaturing have been further used to study the folding stability difference between pH 7.4 and pH 5.5. The free energy of helix bundle has increased by 0.7 kcal/mol monomer upon pH decrease. Both  $\Delta C_p$  and  $T_m$  decrease at pH 5.5 versus pH 7.4, suggesting that PSPF-EKG is better packed in higher pH. Also, in SEC PSPF-EKG presented a significant shoulder upon elution at pH 5.5 versus a sharp peak at pH 7.4. The shoulder did not disappear even as the salt concentration in buffer increased from 150 mM to 2 M. The data suggest pH decrease destabilizes the stability of PSPF-EKG in aqueous system (Figure 4.11A, B), thus validating the first consideration of the original design.

We characterized PSPF-EKG in micelles and bilayers. Equilibrium sedimentation AUC suggests PSPF-EKG adopts a monomer-oligomer equilibrium in C14-betaine micelles at both pHs (Figure 4.11C, D), with slightly higher stability at pH 5.5 than at pH 7.4 (Figure 4.9B, D). A unique oligomerization state could not be determined by AUC due to weak association. Furthermore, the orientation of PSPF-EKG has been studied by ATR-FTIR in POPC lipid bilayers. The average dichroic angle is about 75 degrees with respect to the lipid normal, revealing that the majority of peptides are in a membrane-surface-absorbed state and adopt a vertical conformation with respect to the lipid normal.

This state presumably corresponds to the monomer state identified by AUC (Figure 4.11C). Also, some of the peptides adopt a TM orientation, which might reflect a weakly associated oligomeric form (Figure 4.11D). This dynamic equilibrium between vertical monomer in membrane-surface-absorbed state and TM oligomer state, presumably induces membrane pore formation and plays a crucial role in ATP and miRNA release (Figure 4.11D).

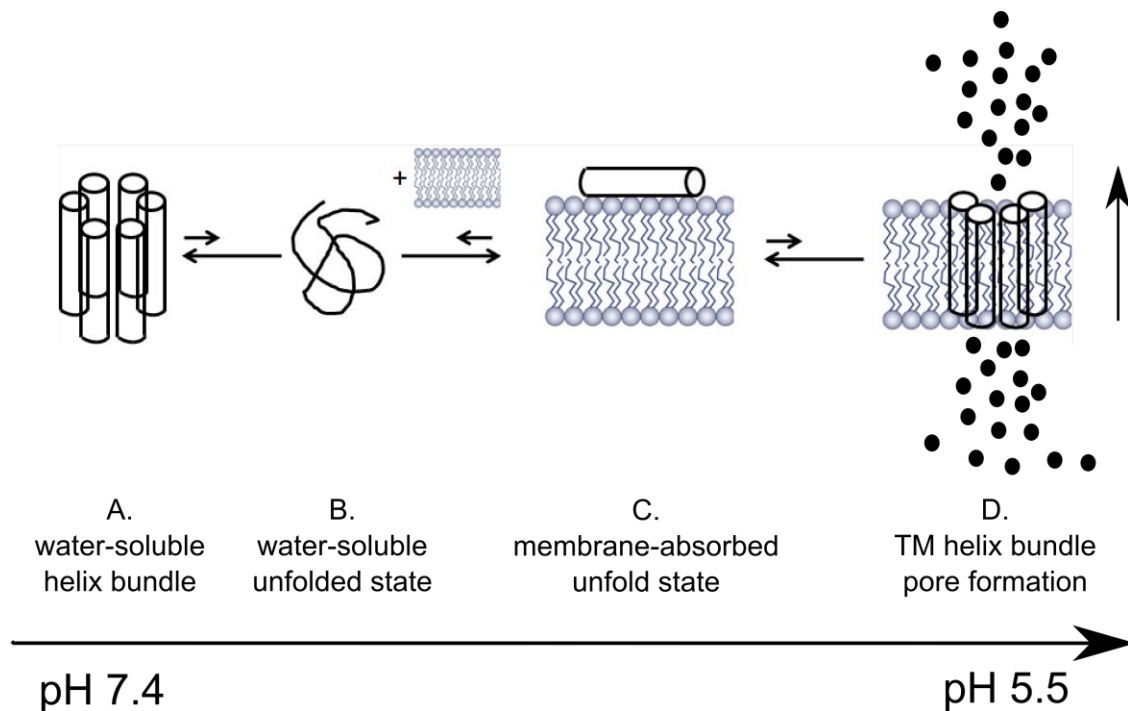


Figure 4.11 Model of PSPF-EKG membrane insertion and pore formation upon pH decrease.

All the biophysical characterizations have corroborated a model for the mechanism of pH-controllable pore formation and ATP/miRNA release (Figure 4.11).

The primary future direction for enhancing the designs is to increase the population of TM peptide relative to the surface-absorbed state. There are two possible avenues that we will pursue towards this aim. Firstly, the inter-peptide interactions in the TM state will be strengthened to favor oligomerization. Secondly, the energetic cost of insertion into the bilayer will be decreased by modifying the hydrophobicity of these peptides.

## 4.5 Methods

### Hemolysis assay

#### *Hemoglobin*

Human Red Blood hemolysis assay was by carried out as described elsewhere<sup>218</sup> with modifications: 5 ml human blood from healthy individuals were transferred into a 50 ml centrifuge tube and either resuspended in 35 ml buffer pH 5.4 (150 mM NaCl, 20 mM MES) or pH 7.5 (150 mM NaCl, 20 mM Hepes). Red Blood Cells (RBCs) were washed 3 times with the appropriate buffer and finally resuspended in a total of 50 ml buffer (pH 5.4 or 7.5). For the assay 175  $\mu$ l of buffer solution (pH 5.4 or 7.5) was added into each well of a clear-bottom 96-well plate followed by 50  $\mu$ l of resuspended RBCs (approx.  $2.5 \times 10^7$  cells) in the appropriate buffer (for RBC transfer wide bore pipet tips were used to avoid cell damage). Test PSPF peptides (New England Peptide<sup>TM</sup>) at the appropriate concentration were diluted in 25  $\mu$ l PBS and then added to the cells. All steps were done with chilled buffers and on ice. The suspension was then mixed 6-8 times by pipetting with wide bore tips, the plate was covered and incubated at 37°C for indicated time. After incubation the cells were centrifuged for 5 min at 500 x g and 150  $\mu$ l of the supernatant was transferred into a new 96-well clear-bottom plate. Absorbance at 541 nm was measured and hemolysis was normalized to RBCs which have been incubated in the presence of 1 % Triton X-100 (100 % hemolysis).

#### *Micro-RNA mir-16*

The release of micro-RNA mir-16 from RBCs was determined using stem-loop PCR as described elsewhere<sup>219</sup>. Briefly, 5  $\mu$ l of supernatant was processed with TaqMan MicroRNA Cells-to-CT Kit (Applied Biosystems) according to manufacturers' protocol and quantitative PCR reaction was performed on an ABI (Applied Biosystems) 7500 Fast Real Time PCR System using standard cycling conditions<sup>220</sup>. The derived Ct values for

mir-16 (Applied Biosystems cat. no.: 4373121) in each experiment were transformed into copy numbers using a linear equation derived from a standard curve which was run in parallel.

### ***ATP***

To quantitatively determine the amount of Adenosine TriPhosphate (ATP) in the supernatant, the ATPLite assay kit (Perkin Elmer; Waltham, MA) was used according to the manufacturers' instructions using 100  $\mu$ l supernatant per reaction point.

### **Tryptophan fluorescence excitation wavelength**

The fluorescence spectra were collected on a Fluorolog spectrofluorometer. The tryptophan fluorescence of each peptide was measured at both pH5.5 and pH7.4 (30m M Phos and 150m M NaCl ), with and without lipid titration.<sup>221</sup> The lipid stock was prepared with 90% POPC and 10% POPG, and the final concentration of lipid after titration is 200 $\mu$ M. The peptide concentration in each measurement was 2 $\mu$ M.

### **CD measurement and thermal denaturation**

CD spectra were collected with a Jasco J-810 spectropolarimeter using a 1-nm step at 4 °C, at both pH5.5 and pH7.4 (30m M Phos and 150m M NaCl ).<sup>139</sup> The PSPF-EKG peptide concentration was 2  $\mu$ M. The CD spectrum was obtained by averaging over three scans.



The helical CD signal at 222nm for 2 $\mu$ M, 4 $\mu$ M and 20 $\mu$ M was monitored as temperature increased from 4 °C to 96 °C at both pH5.5 and pH7.4 (30m M Phos and 150m M NaCl ) , in a 2 °C steps.<sup>46</sup> The parameters from the Gibbs-Helmholtz Equation were fit to the data as shown in Result 3.4.4.3. The fitting was demonstrated in the previous publication.<sup>200</sup>

### **Size exclusion chromatography**

Size exclusion chromatography (SEC) of 100 $\mu$ M PSPF-EKG and 100 $\mu$ M PSPF-DKG were measured by AKTA FPLC machine (GE) using a Superdex 75 column (GE) eluted at pH 7.4 (50 mM Tris, 150 mM NaCl) and pH 5.5 (50 mM MES, 150 mM NaCl) respectively, at 25°C.<sup>46</sup> Four standards were used: blue dextran (2,000,000 g/mol), carbonic anhydrase (29,000 g/mol), cytochrome C (12,400 g/mol) and aprotinin (6,500g/mol). In order to test the effect of salt concentration upon peptide elution, the elutions of PSPF-EKG were also measured at pH 7.4 (50 mM Tris, 2M NaCl) and pH 5.5 (50 mM MES, 2 NaCl), respectively.

### **Sedimentation Equilibrium of Analytical Ultracentrifugation (AUC)**

Sedimentation Equilibrium of Analytical Ultracentrifugation (AUC) of 100 $\mu$ M PSPF-EKG was measured at 25 °C using a Beckman XL-I analytical ultracentrifuge at 35, 40, 45, and 50 kRPM, at both pH 7.4 (50 mM Tris, 150 mM NaCl) and pH 5.5 (50 mM MES, 150 mM NaCl). The data was globally fit to a nonlinear least squares curves by IGOR Pro (Wave-metrics) as previously demonstrated.<sup>82, 139, 141, 157</sup>

The AUC measurement of PSPF-EKG has also been measured in N-tetradecyl-N,N dimethyl-3-ammonio-1-propanesulfonate (C-14 betaine) micelles. 17% D<sub>2</sub>O in buffer was used to precisely match the density of 8mM C-14 betaine micelle at pH 7.4 (50 mM Phos, 150 mM NaCl) and 22% D<sub>2</sub>O was used for pH 5.5 (50 mM Phos, 150 mM NaCl). Three groups of samples were prepared as peptide:DPC molar ratios of 1:50, 1:100, and 1:200 at both pHs. The data with three peptide/detergent ratios and four rotor speeds (35, 40, 45, and 50 kRPM) was globally fit to a nonlinear least squares curves by IGOR Pro (Wave-metrics) as previously demonstrated.<sup>82, 139, 141, 157</sup>

### **Attenuated Total Reflection IR Spectroscopy (ATR-IR)**

ATR-IR of PSPF-EKG was measured by a Nicolet Magna IR 4700 spectrometer using 1 cm<sup>-1</sup> resolution.<sup>143, 144, 201</sup> 5.0<sup>-7</sup> mole PSPF-EKG in trifluoroethanol (TFE) was mixed with 20 fold mole of 1-palmitoyl-2-oleoyl-snglycero-3-phosphocholine (POPC) and dried into a thin film on the surface of ATR Ge crystal evenly by N<sub>2</sub> gas. The film was rehydrated by D<sub>2</sub>O-saturated air overnight in closed environment of D<sub>2</sub>O bath. During data acquisition, the polarized mirror was adjusted to 0° and 90°, creating incident light oriented parallel and perpendicular to the lipid normal respectively. The infrared spectrum of each condition was averaged over 256 scans. The dichroic ratio of 1656 cm<sup>-1</sup> amide I bond absorption is computed for parallel (0°) versus perpendicular (90°) polarized incident light relative to the membrane normal and has been used to calculate the peptide orientation as previously shown.<sup>143</sup>

#### **4.6 Contributions:**

William F DeGrado, Gevorg Grigoryan, Michael Bryan, Jason Donald and Yao Zhang developed the design principles; Rene Bartz, Stephen C. Beck, Nathalie Innocent, David Tellers and Vasant Jadhav from Merck. & Co. performed the cellular release assays; Michael Bryan performed Trp fluorescence, CD scan, and thermal denaturing assays, Yao Zhang performed thermal denaturing data analysis, SEC, AUC, ATR-FTIR and the related data analysis.

## Chapter 5: The design and characterization of a membrane metalloprotein

### 5.1 Introduction

Despite the significant progress in understanding membrane protein folding,<sup>14, 59, 82, 139, 157</sup> the rational design of membrane proteins still remains challenging. Previous efforts have demonstrated successful examples of water-soluble metalloprotein design.<sup>222-225</sup> Also, the design has been further applied to ferrous-ferric redox chemistry as a protein scaffold and exhibited enzymatic functions.<sup>226</sup> Similar metal-binding properties in the membrane phase have long been desired in order to design novel membrane biomaterials and advance our understanding of membrane protein folding.<sup>157</sup>

Here, we present a designed membrane di-iron binding protein MDF (Membrane Duo Ferro). This membrane protein binds to two ferrous cations within a single site and forms a ferrous complex which can then be oxidized into a diferric species upon air oxidation. The design will contribute to our understanding of membrane protein folding, and provides groundwork for designing metalloenzyme that can catalyze the oxidation of small molecules in the membrane phase

## 5.2 Results

The design rational focuses on three major design decisions. Firstly, a His/Glu chelation system was chosen as the metal-binding center, because it is commonly observed in natural and previously designed systems (Figure 5.1A, Table 5.1).<sup>222, 223, 225, 227-229</sup> Secondly, a four-helix bundle, which has been proven successful in natural and previously designed water-soluble metalloproteins,<sup>223, 225, 227, 228, 230</sup> was chosen as the minimal catalysis scaffold for the metal binding site (Figure 5.1B), In contrast to the heterotetramer scaffold which is abundant in designed water-soluble metalloproteins (some were designed as a single chain linked by loops),<sup>224-226</sup> the design of our membrane metalloprotein starts with a homotetramer template in order to simplify the system. Each monomer helix contains both His and Glu residues in order to provide the necessary chelating side-chains. Thirdly, the majority of the backbone sequence was adopted from a previously successfully designed and characterized transmembrane model helix MS1 (Table 5.1),<sup>60</sup> to sidestep any complications with protein aggregation which has been observed in previously computationally designed membrane proteins in house. Based on these criteria the sequence of MDF has been generated as shown in Table 5.1.

Table 5.1 The sequences of DF<sub>tet</sub>B, MS1 and MDF

| Peptide              | g | abcdefg | abcdefg | abcdefg        | abcdefg      |
|----------------------|---|---------|---------|----------------|--------------|
| DF <sub>tet</sub> B* | E | LEELESE | LEKILED | <b>EERHIEW</b> | LEKLEAK LEKL |
| MS1                  |   | BQLLIA  | VLLLIIV | NLILLIA        | VARLRYL VG   |
| MDF                  |   | KKW     | LLLLIAS | <b>ELIHLIL</b> | LALLRYL VG   |

\*DF<sub>tet</sub>B is one chain of a designed water-soluble metal-binding heterotetramer.<sup>226</sup> The chelating residues (Glu and His) have been highlighted in red

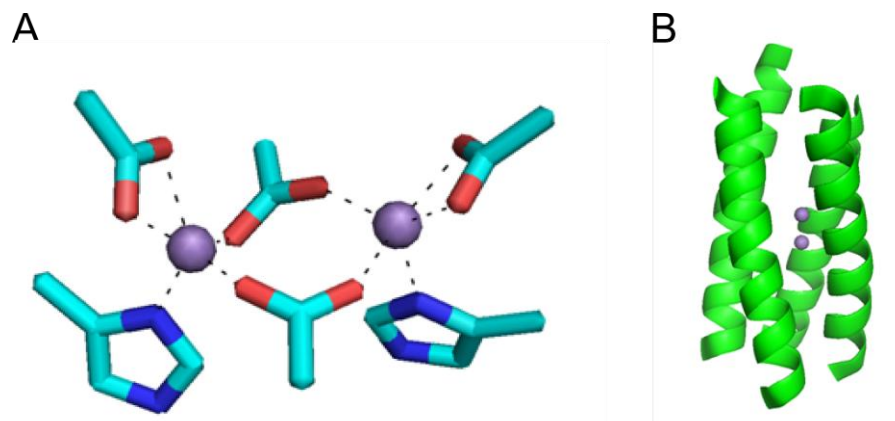


Figure 5.1 Frequently observed His/Glu chelated metal system in both natural and designed proteins<sup>230</sup> (A). Four-helix bundle scaffold for MDF (B).

Designed as a membrane helix, MDF is insoluble in water but can be dissolved and studied in dodecylphosphocholine (DPC) detergent micelles, which mimic the membrane system. A variety of biophysical characterizations have been applied to MDF in order to understand its folding properties. The secondary structure of the apo form studied by circular dichroism (CD), suggests that MDF adopts an alpha-helical secondary structure in micelles at pH7.4 (buffer condition, same buffer in the following assays) (Figure 5.2A). Analytical ultracentrifugation (AUC) sedimentation was also applied to study the association state of MDF at pH7.4. The density of the solution was adjusted by D<sub>2</sub>O to precisely match that of the DPC detergent such that the sedimentation equilibrium was only contributed by MDF peptide component. Fitting the curve to a single molecular weight (MW) species fitting suggests an apparent MW for MDF of 12,500±50 (Figure

5.3A), indicating that a tetramer is formed by the MDF helices. Therefore MDF self-assembles into a membrane four-helix bundle in detergent micelles.

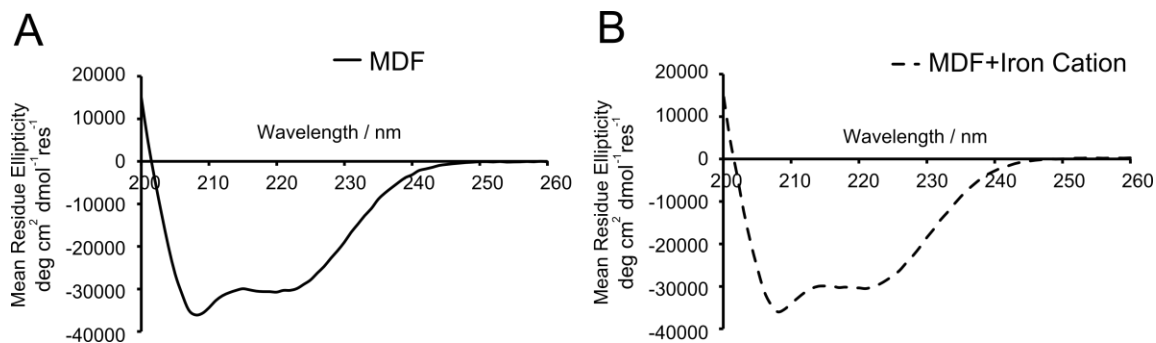


Figure 5.2 CD spectra of MDF (A) and iron-bound MDF (B) in DPC micelles. The spectra show that both apo and bound forms are predominantly  $\alpha$ -helical at a peptide:DPC ratio of 1:25.

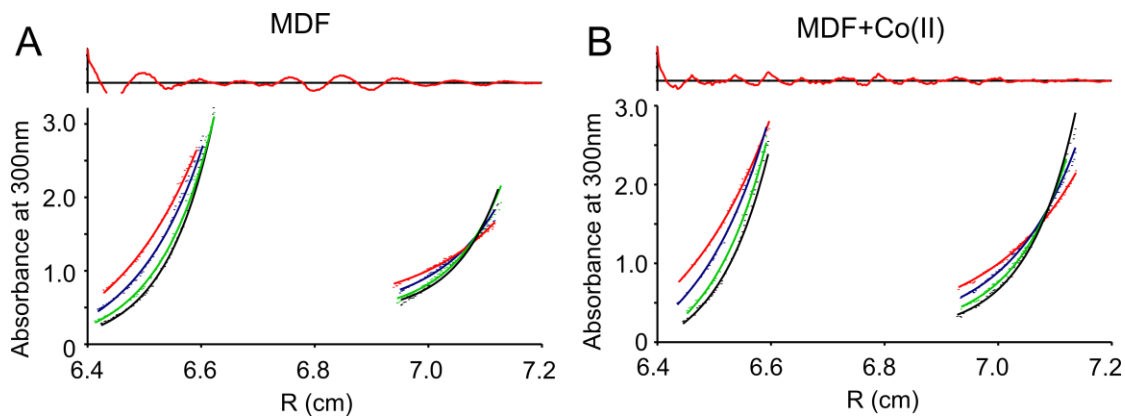


Figure 5.3 Sedimentation equilibrium analysis of MDF (A) and Co(II)-bound MDF (B) in DPC micelles. The absorbance was measured at 300 nm. The single-species fitting suggests a tetramer conformation for both apo and bound forms. The Top of each panel shows the residuals of the fit.

Co(II) is extremely sensitive to its environment, and therefore the UV–visible spectrum can provide useful information on the stoichiometry of metal-binding and the metal-coordination environment.<sup>225</sup> The molar extinction coefficient of Co(II) increases as the coordination number decreases. As Co(II) was titrated into MDF, three absorption bands with  $\lambda_{\text{max}}$  530nm, 550nm and 600nm were observed (Figure 5.4A), presumably due to the cobalt ion's d–d transitions after coordination by the sidechains of MDF. The extinction coefficient at 550nm was calculated to be  $460 \text{ M}^{-1} \text{ cm}^{-1}$  and is indicative of a hexacoordinate Co(II) complex within MDF.<sup>225, 231</sup> Furthermore, a series of concentration of Co(II) was titrated to MDF and the absorbance at 550nm was plot as a function of Co(II)/MDF molar ratios (Figure 5.4B). The data fits nicely with an equilibrium model of two equivalent Co(II)s binding to one MDF tetramer(Scheme 1), resulting in  $-\log(K_d)$  of 14.0. Equilibria with other binding stoichiometries have also been applied, yielding significantly poorer fitting quality.

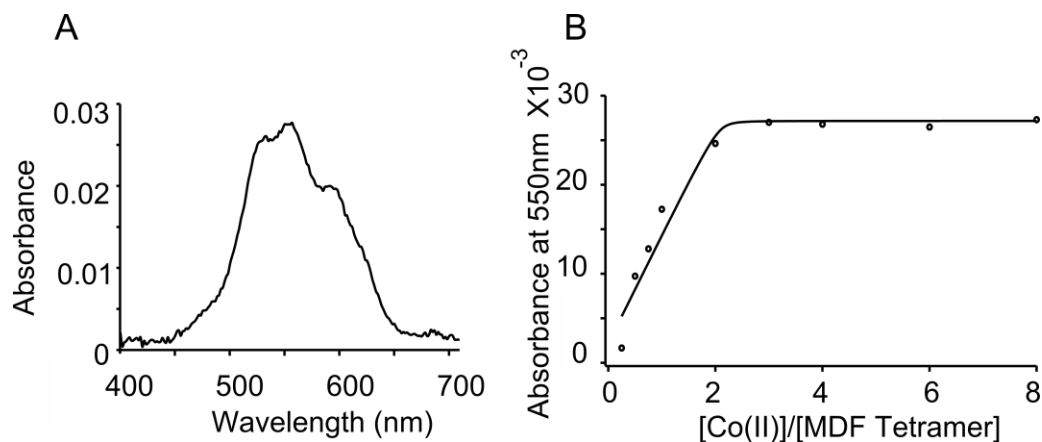
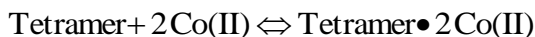


Figure 5.4 (A) Co(II) spectrum of MDF, (B) Titration of MDF with Co(II)



*Scheme 1.*



$$K_d = \frac{[\text{Tetramer} \bullet 2\text{Co(II)}]}{[\text{Tetramer}][\text{Co(II)}]^2}$$

The association state of MDF•2Co(II) complex, where Co(II) and MDF tetramer are mixed at stoichiometric ratio (2:1), has also been measured by AUC sedimentation. The single-species fitting indicates an apparent MW of 11,700±50 (Figure 5.3B). Therefore the metal-bound form remains tetrameric. Moreover, CD spectra of addition of Fe(II) into MDF tetramer at stoichiometry ratio (2:1) suggested an alpha-helical secondary structure (Figure 5.2B), indicating that Fe(II)-binding does not affect the secondary structure of MDF. Therefore the metal-bound form also adopts a four-helix bundle conformation.

Finally Fe(II) has been used as a spectroscopic probe to study the kinetics of metal binding.<sup>224, 226</sup> Since the ferrous form is rapidly oxidized into ferric form in the presence of ambient O<sub>2</sub>, the kinetics of diferric oxo-bridged species formation could be considered as the kinetics of the metal binding step. The time course of the absorbance of MDF tetramer upon the addition of 2 equivalents of Fe (II) was measured by UV-vis spectroscopy. The intensity increase at 320 nm, which results from the ligand-to-metal

charge transfer of the diferric oxo-bridged species, was collected.<sup>224, 226</sup> The curve of the intensity increase at 320 nm as a function of time was fitted with the equilibrium in scheme 2 and resulted in a metal-binding rate of  $1.6\text{s}^{-1}$ ; about 60 fold faster than the rate observed in the water-soluble metalloprotein ( $1.5\text{min}^{-1}$ ).

**Scheme 2.**

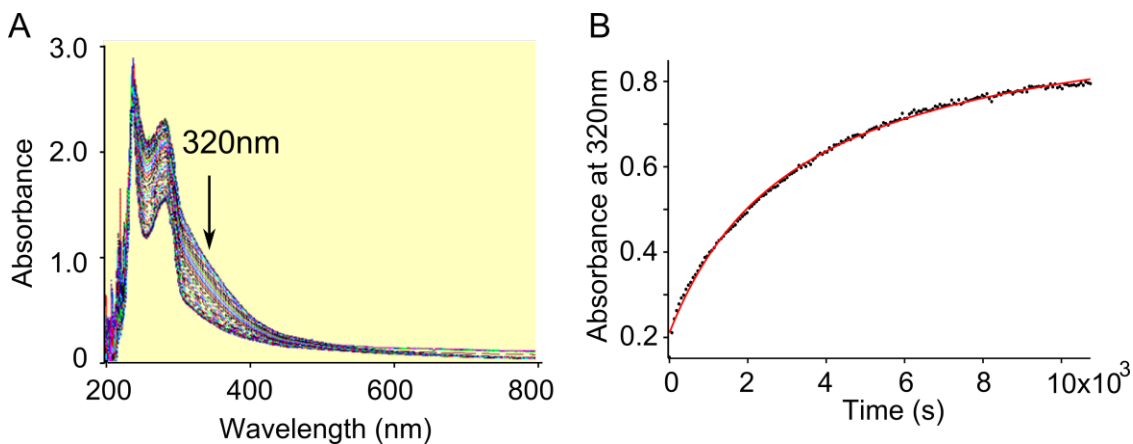
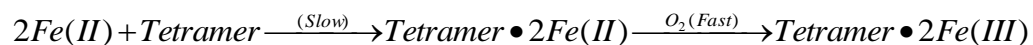


Figure 5.5 (A) Oxidation of Fe (II) in presence of MDF, (B) The time course of the intensity increase at 320nm was fit based on Scheme 2.

### **5.3 Conclusion and discussion**

In conclusion, we have successfully designed a membrane four-helix bundle, which tightly binds metals at a 2 metal cation: 1 tetramer stoichiometric ratio. Since MDF binds to iron cations, it allows the possibility of introducing iron-based redox reaction centers in the membrane phase and can serve as a template for further membrane metalloenzyme design.

## 5.4 Methods

### Peptide Synthesis, cleavage and purification

The sequence of MDF is KKWLLLLIASLHLLALLRYLVG. The peptide was synthesized at a 0.1 mmole scale on RINK amide resin (Novabiochem) on a Symphony peptide synthesizer as previously described<sup>139</sup>. The resin is further cleaved in a mixture of trifluoroacetic acid (TFA)/water/triisopropylsilane (94/3/3) and precipitated with cold diethyl ether. The crude product is purified using a reverse phase HPLC (Vydac, C4 Column) in a gradient of solvent A (water with 0.1% TFA) and solvent B (60% isopropanol, 30% acetonitrile, 10% water and 0.1% TFA).

### Circular dichroism (CD)

In order to prepare peptide detergent micelles, MDF and DPC were mixed in 2,2,2-trifluoroethanol (TFE), dried under nitrogen and rehydrated with the experimental buffer. CD measurements of MDF were carried out in 80 $\mu$ M MDF and 2mM DPC at pH 7.4 (10mM MOP and 10mM NaCl). The measurement of iron-bound form was acquired in a 0.1 cm quartz cell with a Jasco J-810 CD spectropolarimeter, using a 1-nm step at 25°C. The final CD spectra of MDF were averaged over three scans and the baseline (obtained by DPC in buffer) was subtracted. The CD spectra of the iron-bound form were collected from a sample containing 40 $\mu$ M iron cation, 80 $\mu$ M MDF and 2mM DPC, the mixture of which has been incubated for 12 hours prior to measurement.

### **Analytical ultracentrifugation (AUC)**

Sedimentation equilibrium experiments and data analysis were processed as previously demonstrated.<sup>82, 139, 141, 157</sup> The samples contained 320  $\mu\text{M}$  and 160  $\mu\text{M}$  MDF respectively in 8mM DPC at pH7.4 (100mM HEPES and 100 mM NaCl). 37%  $\text{D}_2\text{O}$  has been included in order to match the density of the detergent. The experiments were conducted at 25 °C using a Beckman XL-I analytical ultracentrifuge at respectively 30, 35, 40, 48 KRPM. Data obtained were globally fitted to a nonlinear least-squares curve by IGOR Pro (Wavemetrics).

### **Co (II) titration**

The spectra of Co(II) absorption were collected from 200 $\mu\text{M}$  MDF and 5mM DPC at pH7.4 (150mM MOP and 150 mM NaCl), using a Cary 300 UV spectrometer.  $\text{CoCl}_2$  was added to the sample and allowed to equilibrate at room temperature. Data obtained were fitted to scheme xx by IGOR Pro (Wavemetrics).

### **Fe (II) binding kinetics**

The kinetics of Fe (II) binding was measured at 100 $\mu\text{M}$  MDF and 2.5mM DPC at pH7.4 (150mM MOP and 150 mM NaCl), using a Hewlett Packard model 8453 diode array spectrometer. The reaction was initiated by addition of Fe(II) from a stock solution

(10mM  $\text{Fe}(\text{NH}_4)_2\text{SO}_4$  and 0.01%  $\text{H}_2\text{SO}_4$ ), reaching a final concentration of 50 $\mu\text{M}$  Fe(II).  
Data obtained were fit to scheme xx by IGOR Pro (Wavemetrics).

## **5.5 Contributions**

William F. DeGrado and Yao Zhang designed MDF. Yao Zhang performed all the experiments and related analysis in this chapter.

## Chapter 6: Conclusions and Discussions

This thesis has used helix association as a model system and contributed to our knowledge in: 1. identifying the role of membrane protein folding in type I envelope viral fusion; 2. understanding the driving forces behind membrane protein folding; 3. designing and characterizing novel therapeutic agents.

Chapter 2 introduces the importance of membrane protein folding in viral fusion, where I have discovered that the N-terminal hydrophobic fusion peptides of paramyxovirus fusion proteins form a six-helical TM bundle. This is the first experimental evidence that suggests the fusion peptide of type I viral fusion protein adopts a transmembrane alpha helix secondary structure in lipid bilayers. Further experiments show that the C-terminal membrane segment of the fusion protein strongly interacts with the corresponding N-terminal fusion peptide in a 1:1 ratio. All the data point to a cooperative folding of fusion proteins in both the membrane, in addition to the previous discovery about folding in the water-soluble regions, which is predicted to play a critical role in overcoming the energy barrier of lipid mixing between virus and host cell and driving viral fusion. Based on this discovery, we proposed a novel protein-folding-centric viral fusion model.

In Chapter 3, using a *de novo* designed membrane model peptide MS1, I have identified some of the key forces that drive membrane helix association in chapter 5. My experimental data demonstrated that as the size of apolar residue in the putative **a** position decreases (from isoleucine, to valine, to alanine, to glycine), the helix association affinity significantly increases, the trend of which is opposite to that in aqueous systems, in which hydrophobic interaction is the driving force for folding. Instead, van der Waals interaction appear to plays the crucial roles in bringing membrane helices together and small apolar residues are preferred at the helix interface. Moreover, apolar residues in the helix interfaces also determine the orientation of helix-helix interaction. As the apolar residues in the interface get smaller, the helices are able to get closer to each other, and the role of electrostatic interaction is significantly increased. In order to satisfy the dipole-dipole interaction between the backbones and stabilize the helix bundle, an anti-parallel orientation is preferred. This is the first time that role of these two driving forces have been experimentally identified in membrane helix interactions.

In light of the discovery in Chapter 3, we have designed a group of pH-switchable helix-association-based intracellular delivery agents in Chapter 4. This design has touched a few challenging points: 1. The designed peptide forms stable water-soluble helix in water-soluble system at pH7.4; 2. As the pH decreases to 5.5, which corresponds to the pH in endosome, the stability of water-soluble form was destabilized and the peptide is driven to the membrane insertion; 3. The peptide is also designed to form



stable TM helix bundle at pH 5.5 and lead to endosomal release. The first round of design chose Glu, Asp and His as the pH-dependent switch trigger. Leucine zipper has been selected to stabilize water-soluble form. Small residues such as Gly and Als have been used to stabilize TM helical interface, based on the results from Chapter 3. Furthermore, Lys and Arg have been screened to be water-exposure phase in both water-soluble system (outer-surface) and membrane system (the interior of channel). The top candidate PSPF-EKG in the first round design has successfully induced ATP and miRNA release upon pH switch. Results from the following biophysical assays also point to a mechanism model that agrees well with the original design.

Furthermore, the design and characterization of a membrane metalloprotein has been demonstrated in Chapter 5. This design took the self-assembly of a membrane four-helix bundle as scaffold and introduced the metal binding ( $\text{Co}^{2+}$  and  $\text{Fe}^{2+}$ ) via Glu and His chelating. The design has been validated and can serve as a scaffold for future membrane metalloenzyme design.

In summation, this thesis covers a wide array of diverse applications related to the helix-association model, from the mechanism identification of diseases systems, to the fundamental study of protein folding, and finally to the design of novel functions. The work presented here has contributed to our understanding of protein folding and explored the application of protein folding, in both aqueous and membrane system.

## Chapter 7: References

1. Voet, D.; Voet, J., Biochemistry 3rd edition. Wiley **2004**.
2. Carter, M. D.; Simms, G. A.; Weaver, D. F., The development of new therapeutics for Alzheimer's disease. *Clin Pharmacol Ther* **88** (4), 475-86.
3. Collawn, J. F.; Fu, L.; Bebok, Z., Targets for cystic fibrosis therapy: proteomic analysis and correction of mutant cystic fibrosis transmembrane conductance regulator. *Expert Rev Proteomics* **7** (4), 495-506.
4. Guest, W. C.; Plotkin, S. S.; Cashman, N. R., Toward a mechanism of prion misfolding and structural models of PrP(Sc): current knowledge and future directions. *J Toxicol Environ Health A* **74** (2-4), 154-60.
5. Anfinsen, C. B., Principles that govern the folding of protein chains. *Science* **1973**, **181** (96), 223-30.
6. Branden, C.; Tooze, J., Introduction to Protein Structure 2nd edition. **1999**.
7. Creighton, T. E., Proteins: Structures and Molecular Properties, second edition. **1992**.
8. Tapley, T. L.; Franzmann, T. M.; Chakraborty, S.; Jakob, U.; Bardwell, J. C., Protein refolding by pH-triggered chaperone binding and release. *Proc Natl Acad Sci U S A* **107** (3), 1071-6.
9. Klein, P.; Mattoon, D.; Lemmon, M. A.; Schlessinger, J., A structure-based model for ligand binding and dimerization of EGF receptors. *Proc Natl Acad Sci U S A* **2004**, **101** (4), 929-34.
10. Fan, D. J.; Ding, Y. W.; Pan, X. M.; Zhou, J. M., Thermal unfolding of Escherichia coli trigger factor studied by ultra-sensitive differential scanning calorimetry. *Biochim Biophys Acta* **2008**, **1784** (11), 1728-34.
11. Lee, J. C.; Gray, H. B.; Winkler, J. R., Cytochrome c' folding triggered by electron transfer: fast and slow formation of four-helix bundles. *Proc Natl Acad Sci U S A* **2001**, **98** (14), 7760-4.
12. Lu, D.; Wu, J.; Liu, Z., Dynamic control of protein folding pathway with a polymer of tunable hydrophobicity. *J Phys Chem B* **2007**, **111** (42), 12303-9.
13. Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S., Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **1997**, **89** (2), 263-73.
14. Stouffer, A. L.; Acharya, R.; Salom, D.; Levine, A. S.; Di Costanzo, L.; Soto, C. S.; Tereshko, V.; Nanda, V.; Stayrook, S.; DeGrado, W. F., Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* **2008**, **451** (7178), 596-9.
15. Sutton, R. B.; Fasshauer, D.; Jahn, R.; Brunger, A. T., Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **1998**, **395** (6700), 347-53.
16. Cai, L.; Jiang, S., Development of peptide and small-molecule HIV-1 fusion inhibitors that target gp41. *ChemMedChem* **5** (11), 1813-24.

17. Walshaw, J.; Woolfson, D. N., Socket: a program for identifying and analysing coiled-coil motifs within protein structures. *J Mol Biol* **2001**, *307* (5), 1427-50.
18. Crick, F. H. C., The packing of alpha-helices: simple coiled-coils. *Acta Crystallogr B* **1953**, *6* (8), 9.
19. Mason, J. M.; Arndt, K. M., Coiled coil domains: stability, specificity, and biological implications. *ChemBiochem* **2004**, *5* (2), 170-6.
20. Yu, Y. B., Coiled-coils: stability, specificity, and drug delivery potential. *Adv Drug Deliv Rev* **2002**, *54* (8), 1113-29.
21. **Strathern, J. N.; Jones, E. W.; Broach, J. R.**, The Molecular Biology of the Yeast *Saccharomyces*: Metabolism and Gene Expression. *Cold Spring Harbor Laboratory Pr* **1982**.
22. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T., A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **1993**, *262* (5138), 1401-7.
23. Harbury, P. B.; Kim, P. S.; Alber, T., Crystal structure of an isoleucine-zipper trimer. *Nature* **1994**, *371* (6492), 80-3.
24. Oakley, M. G.; Hollenbeck, J. J., The design of antiparallel coiled coils. *Curr Opin Struct Biol* **2001**, *11* (4), 450-7.
25. Liu, J.; Zheng, Q.; Deng, Y.; Cheng, C. S.; Kallenbach, N. R.; Lu, M., A seven-helix coiled coil. *Proc Natl Acad Sci U S A* **2006**, *103* (42), 15457-62.
26. Stetefeld, J.; Jenny, M.; Schulthess, T.; Landwehr, R.; Engel, J.; Kammerer, R. A., Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat Struct Biol* **2000**, *7* (9), 772-6.
27. Kuhnel, K.; Jarchau, T.; Wolf, E.; Schlichting, I.; Walter, U.; Wittinghofer, A.; Strelkov, S. V., The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. *Proc Natl Acad Sci U S A* **2004**, *101* (49), 17027-32.
28. O'Shea, E. K.; Lumb, K. J.; Kim, P. S., Peptide 'Velcro': design of a heterodimeric coiled coil. *Curr Biol* **1993**, *3* (10), 658-67.
29. Kauzmann, W., Some Factors in the Interpretation of Protein Denaturation. *Advances in Protein Chemistry* **1959**, *14*, 63.
30. Yutani, K.; Ogasahara, K.; Tsujita, T.; Sugino, Y., Dependence of conformational stability on hydrophobicity of the amino acid residue in a series of variant proteins substituted at a unique position of tryptophan synthase alpha subunit. *Proc Natl Acad Sci U S A* **1987**, *84* (13), 4441-4.
31. Conway, J. F.; Parry, D. A., Structural features in the heptad substructure and longer range repeats of two-stranded alpha-fibrous proteins. *Int J Biol Macromol* **1990**, *12* (5), 328-34.
32. Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, I. W.; DeGrado, W. F.; Tirrell, D. A., Stabilization of coiled-coil peptide domains by introduction of trifluoroleucine. *Biochemistry* **2001**, *40* (9), 2790-6.
33. Bilgicer, B.; Fichera, A.; Kumar, K., A coiled coil with a fluororous core. *J Am Chem Soc* **2001**, *123* (19), 4393-9.

34. Potekhin, S. A.; Medvedkin, V. N.; Kashparov, I. A.; Venyaminov, S., Synthesis and properties of the peptide corresponding to the mutant form of the leucine zipper of the transcriptional activator GCN4 from yeast. *Protein Eng* **1994**, *7* (9), 1097-101.
35. Donald, J. E.; Kulp, D. W.; DeGrado, W. F., Salt bridges: geometrically specific, designable interactions. *Proteins* **79** (3), 898-915.
36. Zhou, N. E.; Kay, C. M.; Hodges, R. S., The role of interhelical ionic interactions in controlling protein folding and stability. De novo designed synthetic two-stranded alpha-helical coiled-coils. *J Mol Biol* **1994**, *237* (4), 500-12.
37. Krylov, D.; Barchi, J.; Vinson, C., Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids. *J Mol Biol* **1998**, *279* (4), 959-72.
38. Spek, E. J.; Bui, A. H.; Lu, M.; Kallenbach, N. R., Surface salt bridges stabilize the GCN4 leucine zipper. *Protein Sci* **1998**, *7* (11), 2431-7.
39. Yu, Y.; Monera, O. D.; Hodges, R. S.; Privalov, P. L., Investigation of electrostatic interactions in two-stranded coiled-coils through residue shuffling. *Biophys Chem* **1996**, *59* (3), 299-314.
40. Deng, Y.; Liu, J.; Zheng, Q.; Eliezer, D.; Kallenbach, N. R.; Lu, M., Antiparallel four-stranded coiled coil specified by a 3-3-1 hydrophobic heptad repeat. *Structure* **2006**, *14* (2), 247-55.
41. Monera, O. D.; Zhou, N. E.; Kay, C. M.; Hodges, R. S., Comparison of antiparallel and parallel two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization. *J Biol Chem* **1993**, *268* (26), 19218-27.
42. Monera, O. D.; Kay, C. M.; Hodges, R. S., Electrostatic interactions control the parallel and antiparallel orientation of alpha-helical chains in two-stranded alpha-helical coiled-coils. *Biochemistry* **1994**, *33* (13), 3862-71.
43. Ghosh, I.; Hamilton, A.; Regan, L., Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *J Am Chem Soc* **2000**, *122*, 2.
44. Betz, S.; Fairman, R.; O'Neil, K.; Lear, J.; DeGrado, W., Design of two-stranded and three-stranded coiled-coil peptides. *Philos Trans R Soc Lond B Biol Sci* **1995**, *348* (1323), 81-8.
45. Lombardi, A.; Bryson, J. W.; DeGrado, W. F., De novo design of heterotrimeric coiled coils. *Biopolymers* **1996**, *40* (5), 495-504.
46. Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F., From coiled coils to small globular proteins: design of a native-like three-helix bundle. *Protein Sci* **1998**, *7* (6), 1404-14.
47. Ogihara, N. L.; Weiss, M. S.; DeGrado, W. F.; Eisenberg, D., The crystal structure of the designed trimeric coiled coil coil-VaLd: implications for engineering crystals and supramolecular assemblies. *Protein Sci* **1997**, *6* (1), 80-8.
48. Betz, S. F.; DeGrado, W. F., Controlling topology and native-like behavior of de novo-designed peptides: design and characterization of antiparallel four-stranded coiled coils. *Biochemistry* **1996**, *35* (21), 6955-62.
49. Wheeldon, I. R.; Gallaway, J. W.; Barton, S. C.; Banta, S., Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks. *Proc Natl Acad Sci U S A* **2008**, *105* (40), 15275-80.

50. Bianchi, E.; Joyce, J. G.; Miller, M. D.; Finnefrock, A. C.; Liang, X.; Finotto, M.; Ingallinella, P.; McKenna, P.; Citron, M.; Ottinger, E.; Hepler, R. W.; Hrin, R.; Nahas, D.; Wu, C.; Montefiori, D.; Shiver, J. W.; Pessi, A.; Kim, P. S., Vaccination with peptide mimetics of the gp41 prehairpin fusion intermediate yields neutralizing antisera against HIV-1 isolates. *Proc Natl Acad Sci U S A* **107** (23), 10655-60.
51. Pimentel, T. A.; Yan, Z.; Jeffers, S. A.; Holmes, K. V.; Hodges, R. S.; Burkhard, P., Peptide nanoparticles as novel immunogens: design and analysis of a prototypic severe acute respiratory syndrome vaccine. *Chem Biol Drug Des* **2009**, *73* (1), 53-61.
52. Papapostolou, D.; Smith, A. M.; Atkins, E. D.; Oliver, S. J.; Ryadnov, M. G.; Serpell, L. C.; Woolfson, D. N., Engineering nanoscale order into a designed protein fiber. *Proc Natl Acad Sci U S A* **2007**, *104* (26), 10853-8.
53. Slovic, A. M.; Kono, H.; Lear, J. D.; Saven, J. G.; DeGrado, W. F., Computational design of water-soluble analogues of the potassium channel KcsA. *Proc Natl Acad Sci U S A* **2004**, *101* (7), 1828-33.
54. Korendovych, I. V.; Kim, Y. H.; Ryan, A. H.; Lear, J. D.; DeGrado, W. F.; Shandler, S. J., Computational design of a self-assembling beta-peptide oligomer. *Org Lett* **12** (22), 5142-5.
55. Grigoryan, G.; Reinke, A. W.; Keating, A. E., Design of protein-interaction specificity gives selective bZIP-binding peptides. *Nature* **2009**, *458* (7240), 859-64.
56. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **2001**, *305* (3), 567-80.
57. White, S. H.; Wimley, W. C., Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* **1999**, *28*, 319-65.
58. Popot, J. L.; Gerchman, S. E.; Engelman, D. M., Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process. *J Mol Biol* **1987**, *198* (4), 655-76.
59. DeGrado, W. F.; Gratkowski, H.; Lear, J. D., How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. *Protein Sci* **2003**, *12* (4), 647-65.
60. Choma, C.; Gratkowski, H.; Lear, J. D.; DeGrado, W. F., Asparagine-mediated self-association of a model transmembrane helix. *Nat Struct Biol* **2000**, *7* (2), 161-6.
61. Gratkowski, H.; Lear, J. D.; DeGrado, W. F., Polar side chains drive the association of model transmembrane peptides. *Proc Natl Acad Sci U S A* **2001**, *98* (3), 880-5.
62. Tatko, C. D.; Nanda, V.; Lear, J. D.; DeGrado, W. F., Polar networks control oligomeric assembly in membranes. *J Am Chem Soc* **2006**, *128* (13), 4170-1.
63. Lear, J. D.; Gratkowski, H.; DeGrado, W. F., De novo design, synthesis and characterization of membrane-active peptides. *Biochem Soc Trans* **2001**, *29* (Pt 4), 559-64.
64. Zhou, F. X.; Cocco, M. J.; Russ, W. P.; Brunger, A. T.; Engelman, D. M., Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat Struct Biol* **2000**, *7* (2), 154-60.

65. Partridge, A. W.; Therien, A. G.; Deber, C. M., Polar mutations in membrane proteins as a biophysical basis for disease. *Biopolymers* **2002**, *66* (5), 350-8.
66. Partridge, A. W.; Therien, A. G.; Deber, C. M., Missense mutations in transmembrane domains of proteins: phenotypic propensity of polar residues for human disease. *Proteins* **2004**, *54* (4), 648-56.
67. Joh, N. H.; Min, A.; Faham, S.; Whitelegge, J. P.; Yang, D.; Woods, V. L.; Bowie, J. U., Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* **2008**, *453* (7199), 1266-70.
68. Bowie, J. U., Membrane protein folding: how important are hydrogen bonds? *Curr Opin Struct Biol* *21* (1), 42-9.
69. Lemmon, M. A.; Flanagan, J. M.; Hunt, J. F.; Adair, B. D.; Bormann, B. J.; Dempsey, C. E.; Engelman, D. M., Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J Biol Chem* **1992**, *267* (11), 7683-9.
70. Lemmon, M. A.; Flanagan, J. M.; Treutlein, H. R.; Zhang, J.; Engelman, D. M., Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry* **1992**, *31* (51), 12719-25.
71. Lemmon, M. A.; Treutlein, H. R.; Adams, P. D.; Brunger, A. T.; Engelman, D. M., A dimerization motif for transmembrane alpha-helices. *Nat Struct Biol* **1994**, *1* (3), 157-63.
72. MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M., A transmembrane helix dimer: structure and implications. *Science* **1997**, *276* (5309), 131-3.
73. Russ, W. P.; Engelman, D. M., TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc Natl Acad Sci U S A* **1999**, *96* (3), 863-8.
74. Senes, A.; Gerstein, M.; Engelman, D. M., Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol* **2000**, *296* (3), 921-36.
75. Walters, R. F.; DeGrado, W. F., Helix-packing motifs in membrane proteins. *Proc Natl Acad Sci U S A* **2006**, *103* (37), 13658-63.
76. DeGrado, W. F.; Lear, J. D., Conformationally constrained alpha-helical peptide models for protein ion channels. *Biopolymers* **1990**, *29* (1), 205-13.
77. North, B.; Cristian, L.; Fu Stowell, X.; Lear, J. D.; Saven, J. G.; DeGrado, W. F., Characterization of a membrane protein folding motif, the Ser zipper, using designed peptides. *J Mol Biol* **2006**, *359* (4), 930-9.
78. Moore, D. T.; Berger, B. W.; DeGrado, W. F., Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Structure* **2008**, *16* (7), 991-1001.
79. Elofsson, A.; von Heijne, G., Membrane protein structure: prediction versus reality. *Annu Rev Biochem* **2007**, *76*, 125-40.
80. Barth, P.; Schonbrun, J.; Baker, D., Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc Natl Acad Sci U S A* **2007**, *104* (40), 15682-7.
81. Barth, P.; Wallner, B.; Baker, D., Prediction of membrane protein structures with complex topologies using limited constraints. *Proc Natl Acad Sci U S A* **2009**, *106* (5), 1409-14.

82. Yin, H.; Slusky, J. S.; Berger, B. W.; Walters, R. S.; Vilaire, G.; Litvinov, R. I.; Lear, J. D.; Caputo, G. A.; Bennett, J. S.; DeGrado, W. F., Computational design of peptides that target transmembrane helices. *Science* **2007**, *315* (5820), 1817-22.
83. An, M.; Wijesinghe, D.; Andreev, O. A.; Reshetnyak, Y. K.; Engelman, D. M., pH-(low)-insertion-peptide (pHLIP) translocation of membrane impermeable phalloidin toxin inhibits cancer cell proliferation. *Proc Natl Acad Sci U S A* **107** (47), 20246-50.
84. Korendovych, I. V.; Senes, A.; Kim, Y. H.; Lear, J. D.; Fry, H. C.; Therien, M. J.; Blasie, J. K.; Walker, F. A.; DeGrado, W. F., De novo design and molecular assembly of a transmembrane diporphyrin-binding protein complex. *J Am Chem Soc* **132** (44), 15516-8.
85. Harrison, S. C., Viral membrane fusion. *Nat Struct Mol Biol* **2008**, *15* (7), 690-8.
86. White, J. M.; Delos, S. E.; Brecher, M.; Schornberg, K., Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* **2008**, *43* (3), 189-219.
87. Lamb, R. A.; Jardetzky, T. S., Structural basis of viral invasion: lessons from paramyxovirus F. *Curr Opin Struct Biol* **2007**, *17* (4), 427-36.
88. Colman, P. M.; Lawrence, M. C., The structural biology of type I viral membrane fusion. *Nat Rev Mol Cell Biol* **2003**, *4* (4), 309-19.
89. Wilson, I. A.; Skehel, J. J.; Wiley, D. C., Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **1981**, *289* (5796), 366-73.
90. Stegmann, T.; Delfino, J. M.; Richards, F. M.; Helenius, A., The HA2 subunit of influenza hemagglutinin inserts into the target membrane prior to fusion. *J Biol Chem* **1991**, *266* (27), 18404-10.
91. Blumenthal, R.; Sarkar, D. P.; Durell, S.; Howard, D. E.; Morris, S. J., Dilation of the influenza hemagglutinin fusion pore revealed by the kinetics of individual cell-cell fusion events. *J Cell Biol* **1996**, *135* (1), 63-71.
92. Cross, K. J.; Wharton, S. A.; Skehel, J. J.; Wiley, D. C.; Steinhauer, D. A., Studies on influenza haemagglutinin fusion peptide mutants generated by reverse genetics. *EMBO J* **2001**, *20* (16), 4432-42.
93. Durrer, P.; Galli, C.; Hoenke, S.; Corti, C.; Gluck, R.; Vorherr, T.; Brunner, J., H<sup>+</sup>-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. *J Biol Chem* **1996**, *271* (23), 13417-21.
94. Freitas, M. S.; Gaspar, L. P.; Lorenzoni, M.; Almeida, F. C.; Tinoco, L. W.; Almeida, M. S.; Maia, L. F.; Degreve, L.; Valente, A. P.; Silva, J. L., Structure of the Ebola fusion peptide in a membrane-mimetic environment and the interaction with lipid rafts. *J Biol Chem* **2007**, *282* (37), 27306-14.
95. Ge, M.; Freed, J. H., Fusion peptide from influenza hemagglutinin increases membrane surface order: an electron-spin resonance study. *Biophys J* **2009**, *96* (12), 4925-34.
96. Han, X.; Bushweller, J. H.; Cafiso, D. S.; Tamm, L. K., Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat Struct Biol* **2001**, *8* (8), 715-20.

97. Han, X.; Steinhauer, D. A.; Wharton, S. A.; Tamm, L. K., Interaction of mutant influenza virus hemagglutinin fusion peptides with lipid bilayers: probing the role of hydrophobic residue size in the central region of the fusion peptide. *Biochemistry* **1999**, *38* (45), 15052-9.
98. Macosko, J. C.; Kim, C. H.; Shin, Y. K., The membrane topology of the fusion peptide region of influenza hemagglutinin determined by spin-labeling EPR. *J Mol Biol* **1997**, *267* (5), 1139-48.
99. Jaroniec, C. P.; Kaufman, J. D.; Stahl, S. J.; Viard, M.; Blumenthal, R.; Wingfield, P. T.; Bax, A., Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry* **2005**, *44* (49), 16167-80.
100. Lai, A. L.; Park, H.; White, J. M.; Tamm, L. K., Fusion peptide of influenza hemagglutinin requires a fixed angle boomerang structure for activity. *J Biol Chem* **2006**, *281* (9), 5760-70.
101. Li, Y.; Han, X.; Lai, A. L.; Bushweller, J. H.; Cafiso, D. S.; Tamm, L. K., Membrane structures of the hemifusion-inducing fusion peptide mutant G1S and the fusion-blocking mutant G1V of influenza virus hemagglutinin suggest a mechanism for pore opening in membrane fusion. *J Virol* **2005**, *79* (18), 12065-76.
102. Lorieau, J. L.; Louis, J. M.; Bax, A., The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. *Proc Natl Acad Sci U S A* **107** (25), 11341-6.
103. Sackett, K.; Nethercott, M. J.; Epand, R. F.; Epand, R. M.; Kindra, D. R.; Shai, Y.; Weliky, D. P., Comparative analysis of membrane-associated fusion peptide secondary structure and lipid mixing function of HIV gp41 constructs that model the early pre-hairpin intermediate and final hairpin conformations. *J Mol Biol* **397** (1), 301-15.
104. Schoch, C.; Blumenthal, R., Role of the fusion peptide sequence in initial stages of influenza hemagglutinin-induced cell fusion. *J Biol Chem* **1993**, *268* (13), 9267-74.
105. Spruce, A. E.; Iwata, A.; Almers, W., The first milliseconds of the pore formed by a fusogenic viral envelope protein during membrane fusion. *Proc Natl Acad Sci U S A* **1991**, *88* (9), 3623-7.
106. Steinhauer, D. A.; Wharton, S. A.; Skehel, J. J.; Wiley, D. C., Studies of the membrane fusion activities of fusion peptide mutants of influenza virus hemagglutinin. *J Virol* **1995**, *69* (11), 6643-51.
107. Tse, F. W.; Iwata, A.; Almers, W., Membrane flux through the pore formed by a fusogenic viral envelope protein during cell fusion. *J Cell Biol* **1993**, *121* (3), 543-52.
108. Tsurudome, M.; Gluck, R.; Graf, R.; Falchetto, R.; Schaller, U.; Brunner, J., Lipid interactions of the hemagglutinin HA2 NH2-terminal segment during influenza virus-induced membrane fusion. *J Biol Chem* **1992**, *267* (28), 20225-32.
109. Han, X.; Tamm, L. K., A host-guest system to study structure-function relationships of membrane fusion peptides. *Proc Natl Acad Sci U S A* **2000**, *97* (24), 13097-102.
110. Qiang, W.; Sun, Y.; Weliky, D. P., A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. *Proc Natl Acad Sci U S A* **2009**, *106* (36), 15314-9.



111. Li, Y.; Tamm, L. K., Structure and plasticity of the human immunodeficiency virus gp41 fusion domain in lipid micelles and bilayers. *Biophys J* **2007**, *93* (3), 876-85.
112. Rafalski, M.; Lear, J. D.; DeGrado, W. F., Phospholipid interactions of synthetic peptides representing the N-terminus of HIV gp41. *Biochemistry* **1990**, *29* (34), 7917-22.
113. Martin, I.; Schaal, H.; Scheid, A.; Ruyschaert, J. M., Lipid membrane fusion induced by the human immunodeficiency virus type 1 gp41 N-terminal extremity is determined by its orientation in the lipid bilayer. *J Virol* **1996**, *70* (1), 298-304.
114. Martin, I.; Defrise-Quertain, F.; Decroly, E.; Vandenbranden, M.; Brasseur, R.; Ruyschaert, J. M., Orientation and structure of the NH<sub>2</sub>-terminal HIV-1 gp41 peptide in fused and aggregated liposomes. *Biochim Biophys Acta* **1993**, *1145* (1), 124-33.
115. Nieva, J. L.; Nir, S.; Muga, A.; Goni, F. M.; Wilschut, J., Interaction of the HIV-1 fusion peptide with phospholipid vesicles: different structural requirements for fusion and leakage. *Biochemistry* **1994**, *33* (11), 3201-9.
116. Ishiguro, R.; Kimura, N.; Takahashi, S., Orientation of fusion-active synthetic peptides in phospholipid bilayers: determination by Fourier transform infrared spectroscopy. *Biochemistry* **1993**, *32* (37), 9792-7.
117. Luneberg, J.; Martin, I.; Nussler, F.; Ruyschaert, J. M.; Herrmann, A., Structure and topology of the influenza virus fusion peptide in lipid bilayers. *J Biol Chem* **1995**, *270* (46), 27606-14.
118. Gray, C.; Tatulian, S. A.; Wharton, S. A.; Tamm, L. K., Effect of the N-terminal glycine on the secondary structure, orientation, and interaction of the influenza hemagglutinin fusion peptide with lipid bilayers. *Biophys J* **1996**, *70* (5), 2275-86.
119. Miyauchi, K.; Curran, R.; Matthews, E.; Komano, J.; Hoshino, T.; Engelman, D. M.; Matsuda, Z., Mutations of conserved glycine residues within the membrane-spanning domain of human immunodeficiency virus type 1 gp41 can inhibit membrane fusion and incorporation of Env onto virions. *Jpn J Infect Dis* **2006**, *59* (2), 77-84.
120. Miyauchi, K.; Komano, J.; Yokomaku, Y.; Sugiura, W.; Yamamoto, N.; Matsuda, Z., Role of the specific amino acid sequence of the membrane-spanning domain of human immunodeficiency virus type 1 in membrane fusion. *J Virol* **2005**, *79* (8), 4720-9.
121. Owens, R. J.; Burke, C.; Rose, J. K., Mutations in the membrane-spanning domain of the human immunodeficiency virus envelope glycoprotein that affect fusion activity. *J Virol* **1994**, *68* (1), 570-4.
122. Taylor, G. M.; Sanders, D. A., The role of the membrane-spanning domain sequence in glycoprotein-mediated membrane fusion. *Mol Biol Cell* **1999**, *10* (9), 2803-15.
123. Kemble, G. W.; Danieli, T.; White, J. M., Lipid-anchored influenza hemagglutinin promotes hemifusion, not complete fusion. *Cell* **1994**, *76* (2), 383-91.
124. Markosyan, R. M.; Cohen, F. S.; Melikyan, G. B., The lipid-anchored ectodomain of influenza virus hemagglutinin (GPI-HA) is capable of inducing nonenlarging fusion pores. *Mol Biol Cell* **2000**, *11* (4), 1143-52.
125. Melikyan, G. B.; White, J. M.; Cohen, F. S., GPI-anchored influenza hemagglutinin induces hemifusion to both red blood cell and planar bilayer membranes. *J Cell Biol* **1995**, *131* (3), 679-91.

126. Tong, S.; Compans, R. W., Oligomerization, secretion, and biological function of an anchor-free parainfluenza virus type 2 (PI2) fusion protein. *Virology* **2000**, *270* (2), 368-76.
127. Nobusawa, E.; Aoyama, T.; Kato, H.; Suzuki, Y.; Tateno, Y.; Nakajima, K., Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology* **1991**, *182* (2), 475-85.
128. Senes, A.; Engel, D. E.; DeGrado, W. F., Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* **2004**, *14* (4), 465-79.
129. Kim, S.; Jeon, T. J.; Oberai, A.; Yang, D.; Schmidt, J. J.; Bowie, J. U., Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. *Proc Natl Acad Sci U S A* **2005**, *102* (40), 14278-83.
130. Liu, Z.; Gandhi, C. S.; Rees, D. C., Structure of a tetrameric MscL in an expanded intermediate state. *Nature* **2009**, *461* (7260), 120-4.
131. Stein, A.; Weber, G.; Wahl, M. C.; Jahn, R., Helical extension of the neuronal SNARE complex into the membrane. *Nature* **2009**, *460* (7254), 525-8.
132. Armstrong, R. T.; Kushnir, A. S.; White, J. M., The transmembrane domain of influenza hemagglutinin exhibits a stringent length requirement to support the hemifusion to fusion transition. *J Cell Biol* **2000**, *151* (2), 425-37.
133. Tamm, L. K.; Crane, J.; Kiessling, V., Membrane fusion: a structural perspective on the interplay of lipids and proteins. *Curr Opin Struct Biol* **2003**, *13* (4), 453-66.
134. Tamm, L. K., Hypothesis: spring-loaded boomerang mechanism of influenza hemagglutinin-mediated membrane fusion. *Biochim Biophys Acta* **2003**, *1614* (1), 14-23.
135. Yin, H. S.; Paterson, R. G.; Wen, X.; Lamb, R. A.; Jardetzky, T. S., Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proc Natl Acad Sci U S A* **2005**, *102* (26), 9288-93.
136. Yin, H. S.; Wen, X.; Paterson, R. G.; Lamb, R. A.; Jardetzky, T. S., Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* **2006**, *439* (7072), 38-44.
137. Baker, K. A.; Dutch, R. E.; Lamb, R. A.; Jardetzky, T. S., Structural basis for paramyxovirus-mediated membrane fusion. *Mol Cell* **1999**, *3* (3), 309-19.
138. Bissonnette, M. L.; Donald, J. E.; DeGrado, W. F.; Jardetzky, T. S.; Lamb, R. A., Functional analysis of the transmembrane domain in paramyxovirus F protein-mediated membrane fusion. *J Mol Biol* **2009**, *386* (1), 14-36.
139. Zhang, Y.; Kulp, D. W.; Lear, J. D.; DeGrado, W. F., Experimental and computational evaluation of forces directing the association of transmembrane helices. *J Am Chem Soc* **2009**, *131* (32), 11341-3.
140. Horvath, C. M.; Lamb, R. A., Studies on the fusion peptide of a paramyxovirus fusion glycoprotein: roles of conserved residues in cell fusion. *J Virol* **1992**, *66* (4), 2443-55.
141. Cristian, L.; Lear, J. D.; DeGrado, W. F., Determination of membrane protein stability via thermodynamic coupling of folding to thiol-disulfide interchange. *Protein Sci* **2003**, *12* (8), 1732-40.

142. Mukherjee, S.; Chowdhury, P.; Gai, F., Infrared study of the effect of hydration on the amide I band and aggregation properties of helical peptides. *J Phys Chem B* **2007**, *111* (17), 4596-602.
143. Menikh, A.; Saleh, M. T.; Garipey, J.; Boggs, J. M., Orientation in lipid bilayers of a synthetic peptide representing the C-terminus of the A1 domain of shiga toxin. A polarized ATR-FTIR study. *Biochemistry* **1997**, *36* (50), 15865-72.
144. Tucker, M. J.; Getahun, Z.; Nanda, V.; DeGrado, W. F.; Gai, F., A new method for determining the local environment and orientation of individual side chains of membrane-binding peptides. *J Am Chem Soc* **2004**, *126* (16), 5078-9.
145. Kukol, A.; Adams, P. D.; Rice, L. M.; Brunger, A. T.; Arkin, T. I., Experimentally based orientational refinement of membrane protein models: A structure for the Influenza A M2 H+ channel. *J Mol Biol* **1999**, *286* (3), 951-62.
146. Crick, F. H. C., The Fourier Transform of a Coiled-Coil. *Acta Cryst.* **1953**, *6*, 685-689.
147. MacKerell Jr., A. D., Bashford, D., Bellott, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchner, L, Kuczera, K, Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher III, W.E., Roux, B, Schlenkrich, M, Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., Karplus, M., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
148. Lazaridis, T., Effective energy function for proteins in lipid membranes. *Proteins* **2003**, *52* (2), 176-92.
149. Torii, H.; Tasumi, M., Model calculations on the amide-I infrared bands of globular proteins. *J Chem Phys* **1992**, *96* (5), 3379-87.
150. Manor, J.; Mukherjee, P.; Lin, Y. S.; Leonov, H.; Skinner, J. L.; Zanni, M. T.; Arkin, I. T., Gating mechanism of the influenza A M2 channel revealed by 1D and 2D IR spectroscopies. *Structure* **2009**, *17* (2), 247-54.
151. Chang, D. K.; Cheng, S. F.; Kantchev, E. A.; Lin, C. H.; Liu, Y. T., Membrane interaction and structure of the transmembrane domain of influenza hemagglutinin and its fusion peptide complex. *BMC Biol* **2008**, *6*, 2.
152. Brasseur, R.; Pillot, T.; Lins, L.; Vandekerckhove, J.; Rosseneu, M., Peptides in membranes: tipping the balance of membrane stability. *Trends Biochem Sci* **1997**, *22* (5), 167-71.
153. Chernomordik, L. V.; Kozlov, M. M., Mechanics of membrane fusion. *Nat Struct Mol Biol* **2008**, *15* (7), 675-83.
154. Kasson, P. M.; Lindahl, E.; Pande, V. S., Atomic-resolution simulations predict a transition state for vesicle fusion defined by contact of a few lipid tails. *PLoS Comput Biol* **6** (6), e1000829.
155. Lorizate, M.; Huarte, N.; Saez-Cirion, A.; Nieva, J. L., Interfacial pre-transmembrane domains in viral proteins promoting membrane fusion and fission. *Biochim Biophys Acta* **2008**, *1778* (7-8), 1624-39.

156. Yue, L.; Shang, L.; Hunter, E., Truncation of the membrane-spanning domain of human immunodeficiency virus type 1 envelope glycoprotein defines elements required for fusion, incorporation, and infectivity. *J Virol* **2009**, *83* (22), 11588-98.
157. Kochendoerfer, G. G.; Salom, D.; Lear, J. D.; Wilk-Orescan, R.; Kent, S. B.; DeGrado, W. F., Total chemical synthesis of the integral membrane protein influenza A virus M2: role of its C-terminal domain in tetramer assembly. *Biochemistry* **1999**, *38* (37), 11905-13.
158. Komiya, H.; Yeates, T. O.; Rees, D. C.; Allen, J. P.; Feher, G., Structure of the reaction center from *Rhodobacter sphaeroides* R-26 and 2.4.1: symmetry relations and sequence comparisons between different species. *Proc Natl Acad Sci U S A* **1988**, *85* (23), 9012-6.
159. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J Mol Biol* **1990**, *215* (3), 403-10.
160. Torchilin, V. P., Recent approaches to intracellular delivery of drugs and DNA and organelle targeting. *Annu Rev Biomed Eng* **2006**, *8*, 343-75.
161. North, B.; Summa, C. M.; Ghirlanda, G.; DeGrado, W. F., D(n)-symmetrical tertiary templates for the design of tubular proteins. *J Mol Biol* **2001**, *311* (5), 1081-90.
162. Xiang, Z.; Honig, B., Extending the accuracy limits of prediction for side-chain conformations. *J Mol Biol* **2001**, *311* (2), 421-30.
163. Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C., Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol* **1999**, *285* (4), 1735-47.
164. MacKerell Jr., A. D.; Banavali, N., All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* **2000**, *21*, 86-104.
165. Jorgensen, W. L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926-935.
166. Henin, J., Shinoda, W., Klein, M.L., United-atom acyl chains for CHARMM phospholipids. *J. Phys. Chem. B* **2008**, *112* (23), 7008-7015.
167. Essman, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L., A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103* (19), 8577-8593.
168. Adelman, S. A., Doll, J.D., Generalized Langevin Equation Approach for Atom-Solid-Surface Scattering - General Formulation for Classical Scattering Off Harmonic Solids. *J. Chem. Phys.* **1976**, *64* (6), 2375-2388.
169. Martyna, G. J., Tobias, D.J., Klein, M.L., Constant-Pressure Molecular-Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177-4189.
170. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., Schulten, K., Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **2005**, *26* (16), 1781-1802.
171. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J Mol Graph* **1996**, *14* (1), 33-8, 27-8.

172. Bonnafeous, P.; Stegmann, T., Membrane perturbation and fusion pore formation in influenza hemagglutinin-mediated membrane fusion. A new model for fusion. *J Biol Chem* **2000**, *275* (9), 6160-6.
173. Soto, C. S.; Fasnacht, M.; Zhu, J.; Forrest, L.; Honig, B., Loop modeling: Sampling, filtering, and scoring. *Proteins* **2008**, *70* (3), 834-43.
174. Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D., Design of a novel globular protein fold with atomic-level accuracy. *Science* **2003**, *302* (5649), 1364-8.
175. Gront, D.; Kmiecik, S.; Kolinski, A., Backbone building from quadrilaterals: a fast and accurate algorithm for protein backbone reconstruction from alpha carbon coordinates. *J Comput Chem* **2007**, *28* (9), 1593-7.
176. O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T., X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991**, *254* (5031), 539-44.
177. Acharya, A.; Rishi, V.; Vinson, C., Stability of 100 homo and heterotypic coiled-coil a-a' pairs for ten amino acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry* **2006**, *45* (38), 11324-32.
178. Faham, S.; Yang, D.; Bare, E.; Yohannan, S.; Whitelegge, J. P.; Bowie, J. U., Side-chain contributions to membrane protein structure and stability. *J Mol Biol* **2004**, *335* (1), 297-305.
179. Eilers, M.; Patel, A. B.; Liu, W.; Smith, S. O., Comparison of helix interactions in membrane and soluble alpha-bundle proteins. *Biophys J* **2002**, *82* (5), 2720-36.
180. Liu, W.; Eilers, M.; Patel, A. B.; Smith, S. O., Helix packing moments reveal diversity and conservation in membrane protein structure. *J Mol Biol* **2004**, *337* (3), 713-29.
181. Adamian, L.; Liang, J., Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J Mol Biol* **2001**, *311* (4), 891-907.
182. MacKenzie, K. R.; Fleming, K. G., Association energetics of membrane spanning alpha-helices. *Curr Opin Struct Biol* **2008**, *18* (4), 412-9.
183. Barth, P., Modulating membrane protein stability and association by design. *Curr Opin Struct Biol* **2007**, *17* (4), 460-6.
184. Russ, W. P.; Engelman, D. M., The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol* **2000**, *296* (3), 911-9.
185. Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. S., The role of position a in determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Sci* **1999**, *8* (11), 2312-29.
186. Kim, S.; Chamberlain, A. K.; Bowie, J. U., A simple method for modeling transmembrane helix oligomers. *J Mol Biol* **2003**, *329* (4), 831-40.
187. Kharakoz, D. P., Partial volumes and compressibilities of extended polypeptide chains in aqueous solution: additivity scheme and implication of protein unfolding at normal and high pressure. *Biochemistry* **1997**, *36* (33), 10276-85.
188. Crick, F. H. C., The packing of [alpha]-helices: simple coiled-coils. *Acta Cryst.* **1953**, *6*, 689-697.
189. Senes, A.; Tidor, B., *Unpublished work*.

190. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M., CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **1983**, *4*, 187-217
191. Kingsford, C. L.; Chazelle, B.; Singh, M., Solving and analyzing side-chain positioning problems using linear and integer programming. *Bioinformatics* **2005**, *21* (7), 1028-36.
192. *The GLPK (GNU Linear Programming Kit) package* **2008**, <http://www.gnu.org/software/glpk/>.
193. Whittlesey, K. J.; Shea, L. D., Delivery systems for small molecule drugs, proteins, and DNA: the neuroscience/biomaterial interface. *Exp Neurol* **2004**, *190* (1), 1-16.
194. Veldhoen, S.; Laufer, S. D.; Restle, T., Recent developments in peptide-based nucleic acid delivery. *Int J Mol Sci* **2008**, *9* (7), 1276-320.
195. Camerino, D. C.; Tricarico, D.; Desaphy, J. F., Ion channel pharmacology. *Neurotherapeutics* **2007**, *4* (2), 184-98.
196. Saier, M. H., Jr., Families of proteins forming transmembrane channels. *J Membr Biol* **2000**, *175* (3), 165-80.
197. Lacy, D. B.; Stevens, R. C., Unraveling the structures and modes of action of bacterial toxins. *Curr Opin Struct Biol* **1998**, *8* (6), 778-84.
198. O'Neil, K. T.; DeGrado, W. F., A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **1990**, *250* (4981), 646-51.
199. Betz, S. F.; Liebman, P. A.; DeGrado, W. F., De novo design of native proteins: characterization of proteins intended to fold into antiparallel, rop-like, four-helix bundles. *Biochemistry* **1997**, *36* (9), 2450-8.
200. Daniel P. Raleigh; Stephen F. Betz; DeGrado, W. F., A de Novo Designed Protein Mimics the Native State of Natural Proteins. *J. Am. Chem. Soc.* **1995**, *117* (28), 2.
201. Donald, J. E.; Zhang, Y.; Fiorin, G.; Carnevale, V.; Slochower, D. R.; Gai, F.; Klein, M. L.; DeGrado, W. F., Transmembrane orientation and possible role of the fusogenic peptide from parainfluenza virus 5 (PIV5) in promoting fusion. *Proc Natl Acad Sci U S A* *108* (10), 3958-63.
202. Zhang, H.; Wang, G.; Yang, H., Drug delivery systems for differential release in combination therapy. *Expert Opin Drug Deliv* *8* (2), 171-90.
203. Lammers, T.; Kiessling, F.; Hennink, W. E.; Storm, G., Nanotheranostics and image-guided drug delivery: current concepts and future directions. *Mol Pharm* *7* (6), 1899-912.
204. Gene therapy deserves a fresh chance. *Nature* **2009**, *461* (7268), 1173.
205. Mueller, C.; Flotte, T. R., Clinical gene therapy using recombinant adeno-associated virus vectors. *Gene Ther* **2008**, *15* (11), 858-63.
206. Shim, M. S.; Kwon, Y. J., Efficient and targeted delivery of siRNA in vivo. *FEBS J* *277* (23), 4814-27.
207. Pathak, A.; Patnaik, S.; Gupta, K. C., Recent trends in non-viral vector-mediated gene delivery. *Biotechnol J* **2009**, *4* (11), 1559-72.
208. Zimmermann, T. S.; Lee, A. C.; Akinc, A.; Bramlage, B.; Bumcrot, D.; Fedoruk, M. N.; Harborth, J.; Heyes, J. A.; Jeffs, L. B.; John, M.; Judge, A. D.; Lam, K.;

- McClintock, K.; Nechev, L. V.; Palmer, L. R.; Racie, T.; Rohl, I.; Seiffert, S.; Shanmugam, S.; Sood, V.; Soutschek, J.; Toudjarska, I.; Wheat, A. J.; Yaworski, E.; Zedalis, W.; Koteliansky, V.; Manoharan, M.; Vornlocher, H. P.; MacLachlan, I., RNAi-mediated gene silencing in non-human primates. *Nature* **2006**, *441* (7089), 111-4.
209. Boussif, O.; Lezoualc'h, F.; Zanta, M. A.; Mergny, M. D.; Scherman, D.; Demeneix, B.; Behr, J. P., A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A* **1995**, *92* (16), 7297-301.
210. Watanabe, K.; Harada-Shiba, M.; Suzuki, A.; Gokuden, R.; Kurihara, R.; Sugao, Y.; Mori, T.; Katayama, Y.; Niidome, T., In vivo siRNA delivery with dendritic poly(L-lysine) for the treatment of hypercholesterolemia. *Mol Biosyst* **2009**, *5* (11), 1306-10.
211. Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R., Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol* *5* (8), 747-52.
212. Ruan, L.; Zhang, H.; Luo, H.; Liu, J.; Tang, F.; Shi, Y. K.; Zhao, X., Designed amphiphilic peptide forms stable nanoweb, slowly releases encapsulated hydrophobic drug, and accelerates animal hemostasis. *Proc Natl Acad Sci U S A* **2009**, *106* (13), 5105-10.
213. Hasadsri, L.; Kreuter, J.; Hattori, H.; Iwasaki, T.; George, J. M., Functional protein delivery into neurons using polymeric nanoparticles. *J Biol Chem* **2009**, *284* (11), 6972-81.
214. Chan, J. M.; Zhang, L.; Tong, R.; Ghosh, D.; Gao, W.; Liao, G.; Yuet, K. P.; Gray, D.; Rhee, J. W.; Cheng, J.; Golomb, G.; Libby, P.; Langer, R.; Farokhzad, O. C., Spatiotemporal controlled delivery of nanoparticles to injured vasculature. *Proc Natl Acad Sci U S A* *107* (5), 2213-8.
215. Qiu, L. Y.; Wang, R. J.; Zheng, C.; Jin, Y.; Jin le, Q., Beta-cyclodextrin-centered star-shaped amphiphilic polymers for doxorubicin delivery. *Nanomedicine (Lond)* *5* (2), 193-208.
216. Kwon, E. J.; Bergen, J. M.; Pun, S. H., Application of an HIV gp41-derived peptide for enhanced intracellular trafficking of synthetic gene and siRNA delivery vehicles. *Bioconjug Chem* **2008**, *19* (4), 920-7.
217. Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M., The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A* **2001**, *98* (16), 9056-61.
218. Henry, S. M.; El-Sayed, M. E.; Pirie, C. M.; Hoffman, A. S.; Stayton, P. S., pH-responsive poly(styrene-alt-maleic anhydride) alkylamide copolymers for intracellular drug delivery. *Biomacromolecules* **2006**, *7* (8), 2407-14.
219. Bartz, R.; Fan, H.; Zhang, J.; Innocent, N.; Cherrin, C.; Beck, S. C.; Pei, Y.; Momose, A.; Jadhav, V.; Tellers, D. M.; Meng, F.; Crocker, L. S.; Sepp-Lorenzino, L.; Barnett, S. F., Effective siRNA delivery and target mRNA degradation using an amphipathic peptide to facilitate pH-dependent endosomal escape. *Biochem J* **2011**.
220. Abrams, M. T.; Koser, M. L.; Seitzer, J.; Williams, S. C.; DiPietro, M. A.; Wang, W.; Shaw, A. W.; Mao, X.; Jadhav, V.; Davide, J. P.; Burke, P. A.; Sachs, A. B.; Stirdivant, S. M.; Sepp-Lorenzino, L., Evaluation of efficacy, biodistribution, and

inflammation for a potent siRNA nanoparticle: effect of dexamethasone co-treatment. *Mol Ther* **18** (1), 171-80.

221. Tang, J.; Signarvic, R. S.; DeGrado, W. F.; Gai, F., Role of helix nucleation in the kinetics of binding of mastoparan X to phospholipid bilayers. *Biochemistry* **2007**, *46* (48), 13856-63.

222. Summa, C. M.; Lombardi, A.; Lewis, M.; DeGrado, W. F., Tertiary templates for the design of diiron proteins. *Curr Opin Struct Biol* **1999**, *9* (4), 500-8.

223. Lombardi, A.; Summa, C. M.; Geremia, S.; Randaccio, L.; Pavone, V.; DeGrado, W. F., Retrostructural analysis of metalloproteins: application to the design of a minimal model for diiron proteins. *Proc Natl Acad Sci U S A* **2000**, *97* (12), 6298-305.

224. Marsh, E. N.; DeGrado, W. F., Noncovalent self-assembly of a heterotetrameric diiron protein. *Proc Natl Acad Sci U S A* **2002**, *99* (8), 5150-4.

225. Calhoun, J. R.; Kono, H.; Lahr, S.; Wang, W.; DeGrado, W. F.; Saven, J. G., Computational design and characterization of a monomeric helical dinuclear metalloprotein. *J Mol Biol* **2003**, *334* (5), 1101-15.

226. Kaplan, J.; DeGrado, W. F., De novo design of catalytic proteins. *Proc Natl Acad Sci U S A* **2004**, *101* (32), 11566-70.

227. Rosenzweig, A. C.; Nordlund, P.; Takahara, P. M.; Frederick, C. A.; Lippard, S. J., Geometry of the soluble methane monooxygenase catalytic diiron center in two oxidation states. *Chem Biol* **1995**, *2* (9), 409-18.

228. Logan, D. T.; Su, X. D.; Aberg, A.; Regnstrom, K.; Hajdu, J.; Eklund, H.; Nordlund, P., Crystal structure of reduced protein R2 of ribonucleotide reductase: the structural basis for oxygen activation at a dinuclear iron site. *Structure* **1996**, *4* (9), 1053-64.

229. Pasternak, A.; Kaplan, J.; Lear, J. D.; DeGrado, W. F., Proton and metal ion-dependent assembly of a model diiron protein. *Protein Sci* **2001**, *10* (5), 958-69.

230. Wade, H.; Stayrook, S. E.; DeGrado, W. F., The structure of a designed diiron(III) protein: implications for cofactor stabilization and catalysis. *Angew Chem Int Ed Engl* **2006**, *45* (30), 4951-4.

231. Bertini, I.; Luchinat, C., High spin cobalt(II) as a probe for the investigation of metalloproteins. *Adv Inorg Biochem* **1984**, *6*, 71-111.