

MOLECULAR ORGANIZATION OF PEPTIDE- BASED AMPHIPHILES AT ASSOCIATION COLLOIDAL INTERFACES

by

YONGLIANG ZHANG

A dissertation submitted to the

Graduate School - New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Department of Chemistry and Chemical Biology

Written under the direction of

Professor Laurence S. Romsted

And approved by

New Brunswick, New Jersey

October, 2016

ABSTRACT OF THE DISSERTATION

MOLECULAR ORGANIZATION OF PEPTIDE- BASED AMPHIPHILES AT ASSOCIATION COLLOIDAL INTERFACES

By YONGLIANG ZHANG

Dissertation Director:

Professor Laurence S. Romsted

Elucidating proteins/peptides structures at biomembrane interface is a great challenge. We are developing a chemical approach that reveals topological information of peptides at association colloidal interfaces. This is achieved by using a UV-active amphiphilic molecular probe, a hydrophobic alkyl arenediazonium ion that is oriented with its reactive diazo headgroup in the interfacial region of association colloids and spontaneously reacts with amide bonds, reactive sidechains, and interfacial water to give products. Analysis of various tagging and fragmentation patterns provides information on the locations and local concentrations of peptide backbone and reactive sidechains within the interfacial region.

In **Chapter II** we use this approach to estimate the local concentrations of weakly basic nucleophiles in the interfacial region of micelles of three different *N*-acyl amino acid amphiphiles: *N*-lauroylsarcosine, *N*-lauroylglycine and *N*-lauroylalanine, respectively. Our results demonstrate that the chemical probe, 4-hexadecyl-2,6-dimethyl-benzenediazonium cation, is trapped competitively and reproducibly by the amide oxygen and carboxylate groups of the amphiphile

headgroup, and by water in the interfacial region of the micelles. At a bulk pH of 6.8 - 7.0, the interfacial carboxylate group molarities are estimated to be ~ 1.6 M, but the concentration of the amide bonds for C₁₂Sar micelles is 2 - 5 times less (ca. 0.7 M) than that of C₁₂Ala (~ 1.5 M) and C₁₂Gly micelles (~ 3.0 M). These results indicate that the methyl group on the nitrogen of sarcosine partially buries the *N*-methyl amide bond in the micellar core, whereas glycine and alanine do not.

In **Chapter III**, we run the dediazonation reaction in a series of self-assembled di-peptide amphiphile aqueous solutions prepared at a pH range of 4.0 - 7.0. Product yields show that the local concentration of carboxylate sidechains are in the range at 1.0 - 2.5 M at a solution pH of 6 - 7, but are up to ~ 4.0 M at pH 4.5 when the dipeptide amphiphile forms a gel. Local concentrations of amide bonds are generally dependent on the peptide sequences, ranging across 1.0 - 3.0 M at a solution pH of 6 - 7, but are 2 to 3 times higher at gelation pHs.

In **Chapter IV**, a surprising minimum in the surface tension profile of sodium *N*-dodecanoyl sarcosinate is observed that is consistent with the literature. Our results show that the surface tension minimum in NaLS is caused in part by an excess of the protonated form of the *Z*-isomer. An intra-molecular hydrogen bonding interaction may exist between the amide oxygen and carboxylate headgroup of the protonated *Z* isomer, which contributes to its higher proton affinity than its *E* counterpart. Micellar dilution leads to a transfer of free acid unimer from micelles to water, concurrent with a decrease in the bulk pH.

Acknowledgement

Like many other works, the completion of this dissertation is the result of a collective effort made by people in, and also out of the chemical science including but not limited to:

Professor Laurence Romsted (research advisor and head of thesis committee)

Professor Ralf Warmuth (research collaborator and thesis committee member)

Professor Leslie Jimenez (thesis committee member)

Professor Qinglong Huang (thesis committee member)

Professor Kuang Yu Chen

Professor John Taylor

Professor Peter Lobel

Professor Aijaz Ahmad Dar

Professor Steve Bachofer

Current and former group members, including Tarek Awad, Yan Cai, Xiang Gao, Qing Gu, Gunaseelan Krishnan, Changyao Liu, Brian Regler, Jianbing Zhang and Lanzhen Zhuang.

Department and graduate program administrative staff: Ann Doffinger, Karen Fowler, Melissa Grunwerg, Allison Larkin and Kristina Wetter.

My other colleagues: Sander de Jong, Alexei Ermakov, Rojay Gordon, Seho Kim, Miaoxin Lin, Longle Ma, Brigid McDonald, Murali Nagarayan, Yuan Tian, Huaimin Wang, Qiuyang Xia, Jie Xiao, Letao Yang, Libing Yu, Jingming Zhang, Yixiao Zhang and Haiyan Zheng.

We all give birth to this dissertation. I thank all of these people.

Dedication

Dedicated to my wife, my parents, and Professor Laurence Romsted.

Table of Contents

Abstract.....	ii
Acknowledgement.....	iv
Dedication	v
List of Figures.....	viii
List of Tables	xiv
List of Schemes.....	xvii
Abbreviations	xviii
Chapter I. Introduction.....	1
1.1 MEMBRANE PROTEINS AND THEIR STRUCTURAL ELUCIDATION	1
1.2. INTRODUCTION TO SURFACTANTS.....	9
1.3. CHEMICAL REACTIVITY IN ASSOCIATION COLLOIDS	11
1.4. THE RATIONALE OF THE PRESENT WORK.....	15
Bibliography	21
Chapter II. Simultaneous Determination of Local Concentrations of Amide Bonds, Carboxylate Groups, and Water in Micelles of Amino Acid Amphiphiles Containing Peptide Bond Models.	23
2.1. INTRODUCTION	23
2.2. EXPERIMENTAL SECTION	26
2.2.1. <i>Materials.</i>	26
2.2.2. <i>Methods.</i>	27
2.2.3. <i>Calibration Curves for Dediazonation Reaction Products Formed in C₁₂Ala, C₁₂Sar and C₁₂Gly Micelles.</i>	27
2.2.4. <i>Critical Micelle Concentrations of C₁₂Sar & C₁₂Gly.</i>	28
2.2.5. <i>Dediazonation Kinetics in C₁₂Sar and C₁₂Gly Micelles.</i>	29
2.2.6. <i>Chemical Imaging in C₁₂Ala, C₁₂Sar & C₁₂Gly Micelles.</i>	29
2.3. RESULTS.....	29
2.3.1. <i>Cmcs of C₁₂Sar and C₁₂Gly.</i>	30
2.3.2. <i>Determination of k_{obs} for Dediazonation in C₁₂Sar and C₁₂Gly Micelles.</i>	31
2.3.3. <i>Chemical Imaging Experiments in C₁₂Ala, C₁₂Sar and C₁₂Gly.</i>	32
2.3.4. <i>Estimates of the Interfacial Molarities of Water and Amide and Carboxylate Groups in Amino Acid Amphiphile Micelles.</i>	42
2.4. DISCUSSION	43
2.5. SUMMARY	46
Bibliography	47
Appendix.....	49

SECTION S1. SYNTHESIS AND/OR PURIFICATION PROCEDURES FOR THE PREPARATION OF ARENEDIAZONIUM ION, DEDIAZONIATION PRODUCTS, AMINO ACID AMPHIPHILES AND OTHER RELATED PRODUCTS, AND THEIR ¹ H NMR SPECTRA, RESPECTIVELY.	49
<i>a. Synthetic routes for some of the products.</i>	49
<i>b. 1-n-Hexadecyl-3, 5-dimethylbenzene, 16-ArH.</i>	49
<i>c. 4-n-Hexadecyl-2,6-dimethylphenyllaurate, 16-ArEC₁₂.</i>	50
<i>d. 4-n-Hexadecyl-2, 6-dimethylphenyl-N-lauroyl sarcosinate, 16-ArSE.</i>	51
<i>e. 4-n-Hexadecyl-2, 6-dimethylphenyl-N-lauroyl glycinate, 16-ArGE.</i>	52
<i>f. 4-n-hexadecyl-2, 6-dimethylphenyl-N-lauroylalaninate ester, 16-ArAE.</i>	54
<i>g. Sodium N-lauroyl glycinate, NaLG.</i>	55
<i>h. N-lauroyl-L-alanine, C₁₂Ala.</i>	56
<i>i. Purification of NaLS, its ¹H-NMR and its cmc determination in buffer.</i>	58
SECTION S2. CALIBRATION CURVES FOR REACTION PRODUCTS.	59
SECTION S3. ANALYSIS OF DEDIAZONIATION KINETIC DATA.	64
SECTION S4. SELECTIVITY OF THE DEDIAZONIATION REACTION TOWARD THE AMIDE OXYGEN AND THE CARBOXYLATE GROUP.	68
SECTION S5. ANALYSIS OF 16-AREC12 AND 16-AROHH YIELDS FROM HYDROLYSIS OF 16-AROI FROM ACETAMIDES BASED ON RESULTS PUBLISHED PREVIOUSLY.	69
Bibliography	71
Chapter III. Chemical “Fingerprints” of Peptide-based Amphiphiles - Laying the Foundation for Membrane Peptide Structure Determination	72
3.1. INTRODUCTION	72
3.2. EXPERIMENTAL SECTION	75
3.2.1. <i>Materials.</i>	75
3.2.2. <i>Methods.</i>	75
3.3. RESULTS	76
3.3.1. <i>Dediazoniation Reactions in the Aqueous Solution of Water-Soluble Short Peptide Model.</i>	76
3.3.2. <i>Dediazoniation Reactions at Self-Assembled Aggregate Interfaces.</i>	77
3.3.2.1. <i>Surface Tension Determination.</i>	78
3.3.2.2. <i>Reaction Kinetics.</i>	79
3.3.2.3. <i>Chemical Imaging Experiments by HPLC.</i>	79
3.3.2.4. <i>Peptide Hydrogelation.</i>	86
3.3.2.5. <i>Product Identifications by MALDI-TOF Mass Spectrometry.</i>	87
3.3.2.6. <i>Estimated Local Concentrations of Amide Bonds and Carboxylate Headgroups of PAs.</i>	92
3.4. DISCUSSION	94
3.5. SUMMARY	97
Bibliography	99
Appendix	100
SECTION S1. SYNTHESSES OF PEPTIDE AMPHIPHILES.	100
<i>a. N-lauroyl-L-alanyl-L-glutamic acid. C₁₂AE.</i>	102
<i>b. N-lauroylglycyl-L-glutamic acid. C₁₂GE.</i>	103
<i>c. N-lauroyl-L-alanyl-L-phenylalanine. C₁₂AF.</i>	104
<i>d. N-lauroylglycyl-L-phenylalanine. C₁₂GF.</i>	105
<i>e. N-lauroyl-L-alanyl-L-phenylalanyl-L-glutamyl-L-glutamyl-L-glutamic acid, C₁₂AFEEE.</i>	106

SECTION S2. SYNTHESIS AND PURIFICATION PROCEDURES FOR THE PREPARATION OF ARENEDIAZONIUM ION, DEDIAZONIATION PRODUCTS AND OTHER RELATED PRODUCTS, AND THEIR ^1H NMR AND ESI-MS SPECTRA, RESPECTIVELY.	108
<i>a. Synthetic routes for some of the products, 1-ArN₂⁺, 1-ArOH, 1-ArOAc, 16-ArNH₂, 16-ArN₂⁺, 16-ArOH, 16-ArNHAc and 16-ArInd, have been published. See Chapter II.</i>	108
<i>b. 4-n-Hexadecyl-2, 6-dimethylphenyl-N-lauroyl-L-alanyl-L-phenylalaninate, 16-ArEFAC₁₂.</i>	108
<i>c. 4-n-Hexadecyl-2, 6-dimethylphenyl-N-lauroyl-L-alanyl-L-glutamate, 16-ArEEAC₁₂ (alpha ester).</i>	109
<i>d. 4-n-Hexadecyl-2, 6-dimethylphenyl-N-lauroyl-L-alanyl-L-glutamate, 16-ArEEAC₁₂ (gamma ester).</i>	110
<i>e. 2, 4, 6-Trimethylphenyl-N-acetyl glycine, 1-ArOGAc.</i>	111
<i>f. 2, 4, 6-Trimethylphenyl-N-acetyl glycyglycine, 1-ArOGGAc.</i>	112
SECTION S3. PROTOCOL AND CONDITIONS OF THE SYNTHESIZED CHEMICAL IMAGING PRODUCT CHARACTERIZATION VIA ESI-MS AND MALDI-TOF MS.	113
SECTION S4. CALIBRATION CURVES FOR REACTION PRODUCTS.	115
SECTION S5. ANALYSES OF DEDIAZONIATION KINETIC DATA.	119
SECTION S6. SUPPLEMENTAL INFORMATION OF HPLC OBSERVED AND NORMALIZED PERCENT YIELDS OF REACTIONS OF 16-ArN ₂ ⁺ WITH DIPEPTIDE AMPHIPHILE AQUEOUS SOLUTIONS.	121
SECTION S7. GELATION PROPERTIES OF C ₁₂ GE, C ₁₂ AE AND C ₁₂ AF.	122
Chapter IV. Interpretation of the Minimum at Surface Tension Curves of Aqueous N-lauroyl Sarcosinate Solutions.	124
4.1. INTRODUCTION	124
4.2. EXPERIMENTAL SECTION	130
4.3. RESULTS	133
4.4. DISCUSSION	144
4.5. SUMMARY	148
Bibliography	149
Appendix	151
Chapter V. Conclusions and Perspectives	157
Bibliography	159

List of Figures

FIGURE 1.1. Molecular Model of the Cross Section of an Average Synaptic Vesicle Showing the Orientations of Different Types of Proteins and the Membrane Phospholipids.....	2
FIGURE 1.2. Flow Chart of Protein Structural Elucidation by X-Ray Diffraction.....	3
FIGURE 1.3. Examples of Different Pure Secondary Structures Detecting by Circular Dichrism.....	4
FIGURE 1.4. NMR Spectroscopy and Protein Structure Determination.....	5
FIGURE 1.5. Principle of Titration of Basic Amino Acid Residues by Sulfonates.....	6
FIGURE 1.6. General Workflow for Covalent Labeling of Proteins with Subsequent Analysis by Limited Proteolysis, Liquid Chromatography and Electrospray Mass Spectrometry.....	8
FIGURE 1.7. Representative Surfactants and their Structures.....	10
FIGURE 1.8. Representative Self-assembled Aggregates in Association Colloids.....	11
FIGURE 1.9. Illustration of a Small Section of the Immediate Vicinity of the Interfacial Region of Membrane Mimic in Aqueous Solution.....	18
FIGURE 1.10. An α -Helix Peptide Segment Oriented Either Perpendicular or Parallel to a Small Section of Membrane Bilayer that Contains Molecular Probes with Reactive Functional Groups Located in the Interfacial Region.....	19
FIGURE 2.1. Principles of Chemical Imaging Approach.....	25
FIGURE 2.2. Surface Tension versus the Natural Logarithm of $C_{12}Sar$ and $C_{12}Gly$ Concentrations at Ambient Temperature.....	30
FIGURE 2.3. HPLC Chromatogram for Dediazonation of $16-ArN_2^+$ in 0.098M $C_{12}Sar$ Aqueous Solution.....	33
FIGURE 2.4. HPLC Chromatogram for Dediazonation of $16-ArN_2^+$ in 0.098M $C_{12}Gly$ Aqueous Solution.....	33
FIGURE 2.5. HPLC Chromatogram for Dediazonation of $16-ArN_2^+$ in 0.098M $C_{12}Ala$ Aqueous Solution.....	34
FIGURE 2.6. Cartoons of a Small Cross Section of $C_{12}Sar$ and $C_{12}Gly$ Micelles Containing the Chemical Imaging Probe $16-ArN_2^+$	45
FIGURE 2S1-B. 1H -NMR Spectrum for $16-ArH$	50
FIGURE 2S1-C. 1H -NMR Spectrum for $16-ArEC_{12}$	51

FIGURE 2S1-D. ^1H -NMR and ESI-MS Spectrum for 16-ArSE.....	52
FIGURE 2S1-E. ^1H -NMR and ESI-MS Spectrum for 16-ArGE.....	53
FIGURE 2S1-F. ^1H -NMR and ESI-MS Spectrum for 16-ArAE.....	54
FIGURE 2S1-G. ^1H -NMR and ESI-MS Spectrum for Sodium <i>N</i> -Lauroyl Glycinate.....	56
FIGURE 2S1-H. ^1H -NMR and ESI-MS Spectrum for <i>N</i> -Lauroylalanine.	57
FIGURE 2S1-I. ^1H -NMR and ESI-MS Spectrum for Sodium <i>N</i> -Lauroyl Sarcosinate.....	58
FIGURE 2S2-A. Calibration Curve for 16-ArOH.....	60
FIGURE 2S2-B. Calibration Curve for 16-ArNHAc.....	60
FIGURE 2S2-C. Calibration Curve for 16-ArH.....	61
FIGURE 2S2-D. Calibration Curve for 16-ArF.....	61
FIGURE 2S2-E. Calibration Curve for 16-ArSE.....	62
FIGURE 2S2-F. Calibration Curve for 16-ArGE.....	62
FIGURE 2S2-G. Calibration Curve for 16-ArAE.....	63
FIGURE 2S2-H. Calibration Curve for 16-ArEC ₁₂	63
FIGURE 2S2-I. Calibration Curve for 16-ArInd.....	64
FIGURE 2S3-A. Decrease in Absorbance for Dediazonation of 2×10^{-4} M 16-ArN ₂ ⁺ in 0.098 M C ₁₂ Sar and C ₁₂ Gly Micelles at 285.5 nm at 40°C over 7 hours.....	65
FIGURE 2S3-B. Decrease in Absorbance for Dediazonation of 2×10^{-4} M 16-ArN ₂ ⁺ in 0.069 M C ₁₂ Sar and C ₁₂ Gly Micelles at 285.5 nm at 40°C over 7 hours.....	65
FIGURE 2S3-C. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN ₂ ⁺ in 0.098 M C ₁₂ Sar.....	66
FIGURE 2S3-D. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN ₂ ⁺ in 0.098 M C ₁₂ Gly.....	66
FIGURE 2S3-E. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN ₂ ⁺ in 0.069 M C ₁₂ Sar.....	67
FIGURE 2S3-F. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN ₂ ⁺ in 0.069 M C ₁₂ Gly.....	67
FIGURE 3.1. Representative HPLC Chromatogram for Reaction of Ac-Gly-Gly at a Concentration of 5.0 M at pH 4.0.....	76

FIGURE 3.2. Surface Tension Measurement of C ₁₂ AENa at 27°C.....	78
FIGURE 3.3. HPLC Chromatogram for Dediazonation of 16-ArN ₂ ⁺ in 50 mM C ₁₂ Gly-Glu Aqueous Solution at pH 6.0 at Two Different HPLC Eluent Conditions.....	80
FIGURE 3.4. HPLC Chromatogram for Dediazonation of 16-ArN ₂ ⁺ in 0.1 M C ₁₂ Ala-Glu Aqueous Solution at pH 6.0 at Two Different HPLC Eluent Conditions.....	82
FIGURE 3.5. HPLC Chromatogram for Dediazonation of 16-ArN ₂ ⁺ in 0.08 M C ₁₂ Ala-Phe Aqueous Solution at pH 6.8.....	84
FIGURE 3.6. HPLC Chromatogram for Dediazonation of 16-ArN ₂ ⁺ in 0.05 M C ₁₂ Gly-Phe Aqueous Solution at pH 6.5.....	84
FIGURE 3.7. HPLC Chromatogram for Dediazonation of 16-ArN ₂ ⁺ in 0.08 M C ₁₂ Ala-Phe-Glu-Glu Aqueous Solution at pH 6.0.....	86
FIGURE 3.8. MALDI-TOF Spectrum of the Ester Product, 16-ArEGC ₁₂	88
FIGURE 3.9. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEEGC ₁₂ Gamma Ester.....	89
FIGURE 3.10. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEEAC ₁₂ Alpha Ester...	89
FIGURE 3.11. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEEAC ₁₂ Gamma Ester.....	90
FIGURE 3.12. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEAC ₁₂	90
FIGURE 3.13. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEFAC ₁₂	91
FIGURE 3.14. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEGC ₁₂	91
FIGURE 3.15. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEFGC ₁₂	96
FIGURE 3.16. Chemical Structure of Short Peptide Amphiphile, C ₁₂ Ala-Glu and Predicted Orientation of Cleavable Amide Bonds and Carboxylate Sidechains at Interface of Self-Assemblies.....	97
FIGURE 3.17. Chemical Structure of Short Peptide Amphiphile, C ₁₂ Gly-Glu and Predicted Orientation of Cleavable Amide Bonds and Carboxylate Sidechains at Interface of Self-Assemblies.....	99
FIGURE 3S1-A. ¹ H-NMR and ESI-MS Spectrum for <i>N</i> -Lauroylalanyl Glutamic Acid, C ₁₂ AE.....	103
FIGURE 3S1-B. ¹ H-NMR and ESI-MS Spectrum for <i>N</i> -Lauroylglycyl Glutamic Acid, C ₁₂ GE.....	104
FIGURE 3S1-C. ¹ H-NMR and ESI-MS Spectrum for <i>N</i> -Lauroylalanyl Phenylalanine, C ₁₂ AF.....	105

FIGURE 3S1-D. ^1H -NMR and ESI-MS Spectrum for <i>N</i> -Lauroylglycyl Phenylalanine, C_{12}GF	106
FIGURE 3S1-E. ^1H -NMR and ESI-MS Spectrum for <i>N</i> -Lauroyl- <i>L</i> -Alanyl- <i>L</i> -Phenylalanyl- <i>L</i> -Glutamyl- <i>L</i> -Glutamyl- <i>L</i> -Glutamic Acid, $\text{C}_{12}\text{AFEEEE}$	107
FIGURE 3S2-B. ^1H -NMR and ESI-MS Spectrum for 4- <i>n</i> -Hexadecyl-2, 6-dimethylphenyl- <i>N</i> -lauroylalanyl phenylalaninate, 16-ArEFAC $_{12}$	109
FIGURE 3S2-C. ^1H -NMR and ESI-MS Spectrum for 4- <i>n</i> -Hexadecyl-2, 6-dimethylphenyl- <i>N</i> -lauroylalanyl glutamate, 16-ArEEAC $_{12}$ (Alpha Ester).....	110
FIGURE 3S2-D. ESI-MS Spectrum for 4- <i>n</i> -Hexadecyl-2, 6-dimethylphenyl- <i>N</i> -lauroylalanyl glutamate, 16-ArEEAC $_{12}$ (Gamma Ester).....	111
FIGURE 3S2-E. ^1H -NMR and ESI-MS Spectrum for 2, 4, 6-Trimethylphenyl- <i>N</i> -acetyl glycine, 1ArOGAc.....	112
FIGURE 3S2-F. ^1H -NMR and ESI-MS Spectrum for 2, 4, 6-Trimethylphenyl- <i>N</i> -acetyl glycylglycine, 1-ArOGGAc.....	113
FIGURE 3S4-1. Calibration Curve for 1-ArOGAc.....	116
FIGURE 3S4-2. Calibration Curve for 1-ArOGGAc.....	117
FIGURE 3S4-3. Calibration Curve for 16-ArEFAC $_{12}$	117
FIGURE 3S4-4. Calibration Curve for 16-ArEEAC $_{12}$ Alpha Ester.....	118
FIGURE 3S4-5. Calibration Curve for 16-ArEEAC $_{12}$ Gamma Ester.....	118
FIGURE 3S5-1. Ln Plot of UV Absorbance for the Dediazonation of 6×10^{-4} M 16-ArN $_2^+$ in 0.098 M $\text{C}_{12}\text{Ala-Glu}$ at pH 4.8 at 40°C. S Stands for Time in Seconds.....	119
FIGURE 3S5-2. Ln Plot of UV Absorbance for the Dediazonation of 6×10^{-4} M 16-ArN $_2^+$ in 0.098 M $\text{C}_{12}\text{Ala-Glu}$ at pH 6.2 at 40°C. S Stands for Time in Seconds.....	120
FIGURE 3S8-1. Appearance of Six PA Samples at Various Solution pHs.....	123
FIGURE 4.1. Surface Tension Plots of Aqueous Solutions of NaLS at Ambient Temperature.....	135
FIGURE 4.2. Surface Tension Measurements of Aqueous Solutions of MLS with Different Counterion Types.....	137
FIGURE 4.3. Change of Chemical Shifts of Protons at Different Chemical Environments of the NaLS Molecules at Various Concentrations in their Micellar or Nonmicellar Aqueous Solution.....	142

FIGURE 4.4. ^1H -NMR Chemical Shifts of the Carboxylate Methylene Protons, $\text{C}_{11}\text{H}_{23}\text{CON}(\text{CH}_3)\text{CH}_2\text{COONa}$, in Increasing Concentrations of NaLS in D_2O Solution.....	143
FIGURE 4S-1a. ^1H -NMR spectrum of 47.7 mM sodium <i>N</i> -lauroyl sarcosinate micelles (in D_2O). Note that the <i>Z</i> and <i>E</i> rotamer signal labels are this and follow ^1H NMR spectra. Average <i>Z</i> : <i>E</i> ratio = 1.86.....	151
FIGURE 4S-1b. ^1H -NMR spectrum of 6.8 mM sodium <i>N</i> -lauroyl sarcosinate monomer in D_2O . Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2^-$ are shown. Average <i>Z</i> : <i>E</i> ratio = 0.91.....	151
FIGURE 4S-1c. ^1H -NMR spectrum of 6.8 mM sodium <i>N</i> -lauroyl sarcosinate at pD 11 in D_2O (Titrated with NaOD). Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2^-$ are shown. Average <i>Z</i> : <i>E</i> ratio = 0.88.....	152
FIGURE 4S-1d. ^1H -NMR spectrum of 0.22 mM <i>N</i> -lauroylsarcosine at pD 2.2 in D_2O (titrated with DCl). Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2^-$ are shown. Average <i>Z</i> : <i>E</i> ratio = 1.30.....	152
FIGURE 4S-1e. ^1H -NMR spectrum of 1 mM sodium <i>N</i> -lauroyl sarcosinate in CD_3OD	153
FIGURE 4S-1f. ^1H -NMR spectrum of <i>N</i> -lauroylsarcosine (1 mM) in CD_3OD	153
FIGURE 4S-1g. ^1H -NMR spectrum of sodium <i>N</i> -lauroyl sarcosinate (6.8 mM) in CDCl_3	153
FIGURE 4S-1h. ^1H -NMR spectrum of <i>N</i> -lauroylsarcosine (6.8 mM) in CDCl_3	153
FIGURE 4S-2. Acid-induced precipitation of sodium <i>N</i> -lauroyl sarcosinate from aqueous solutions: (A) 27.2 mM, pH = 6.0; (B) 5.0 mM, pH = 5.0.....	154
FIGURE 4S-3. Titration of <i>N</i> -lauroylsarcosine, HLS, aqueous solution with 200 mM NaOH..	154
FIGURE 4S-4. ^1H -NMR chemical shifts of the <i>N</i> -methyl proton, $\text{C}_{11}\text{H}_{23}\text{CON}(\text{CH}_3)^*\text{CH}_2\text{COONa}$, at various concentrations of NaLS in D_2O solution.....	155
FIGURE 4S-5. ^1H -NMR chemical shifts of the α -methylene protons, $\text{C}_{10}\text{H}_{21}\text{CH}_2^*\text{CON}(\text{CH}_3)\text{CH}_2\text{COONa}$, at various concentrations of NaLS in D_2O	155
FIGURE 5.1. HIV-1 MPER Peptide.....	158

List of Tables

TABLE 2.1. Observed Rate Constant and Half-life for Dediazonation of 16-ArN ₂ ⁺ in 0.098 and 0.069 M C ₁₂ Sar and C ₁₂ Gly.....	31
TABLE 2.2. The Retention Time Ranges for Each Product from Chemical Imaging Experiments.....	34
TABLE 2.3. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10 ⁻⁴ M 16-ArN ₂ ⁺ in 0.098 M C ₁₂ Sar Aqueous Solution at 40 ± 0.1°C with a Reaction Time of 12 hours.....	35
TABLE 2.4. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10 ⁻⁴ M 16-ArN ₂ ⁺ with 0.069 M C ₁₂ Sar Aqueous Solution at 40 ± 0.1°C with a Reaction Time of 12 hours.....	36
TABLE 2.5. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10 ⁻⁴ M 16-ArN ₂ ⁺ with 0.098 M C ₁₂ Gly Aqueous Solution at 40 ± 0.1°C with a Reaction Time of 12 hours.....	37
TABLE 2.6. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10 ⁻⁴ M 16-ArN ₂ ⁺ with 0.069 M C ₁₂ Gly Aqueous Solution at 40 ± 0.1°C with a Reaction Time of 12 hours.....	38
TABLE 2.7. HPLC Normalized Yields for the Reaction of 8×10 ⁻⁴ M 16-ArN ₂ ⁺ with 0.1 M C ₁₂ Ala Aqueous Suspension at pH 6 and 40 ± 0.1°C with a Reaction Time of 12 hours.....	39
TABLE 2.8. Normalized Yields of the Key Products for the Reaction of 2×10 ⁻⁴ M 16-ArN ₂ ⁺ in Aqueous Solutions of 0.098 M C ₁₂ Sar and C ₁₂ Gly, and 0.100 M C ₁₂ Ala at 40 ± 0.1°C at pH 6-7 with Reaction Times of 12 hours.....	39
TABLE 2.9. Estimated Interfacial Local Concentrations of the Nucleophiles from Chemical imaging in Aqueous 0.098 M C ₁₂ Sar and C ₁₂ Gly, and 0.100 M C ₁₂ Ala.....	43
TABLE 2S2-A. Equations Used to Fit HPLC Calibration Curves for Dediazonation Products..	59
TABLE 2S4-A. Average Selectivities Determined from Dediazonation from 1-ArN ₂ ⁺ in the Presence of Aqueous Amides Two Different Water/amide, N _w /N _d , Molar Ratios at 40°C.....	68
TABLE 2S4-B. Normalized Yields for 1-ArN ₂ ⁺ Dediazonation in Aqueous Glycine and the Selectivity of Carboxylate Group at 40°C.....	69
TABLE 3.1. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 5.0 M Ac-Gly-Gly at pH 4.0 and 40 ± 0.1°C with a Reaction Time of 12 hours.....	77
TABLE 3.2. Observed Rate Constant and Half Life for Dediazonation of 16-ArN ₂ ⁺ in 0.1 M C12Ala-Glu at pH 6.2 and pH 4.8 at 40°C.....	79

TABLE 3.3. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.1 M C ₁₂ Gly-Glu Aqueous Solution at pH 6 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	81
TABLE 3.4. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.1 M C ₁₂ Gly-Glu Aqueous Solution at pH 4.5 (Gel) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	81
TABLE 3.5. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.1 M C ₁₂ Ala-Glu Aqueous Solution at pH 6 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	83
TABLE 3.6. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.1 M C ₁₂ Ala-Glu Aqueous Solution at pH 4.5 (Gel) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	83
TABLE 3.7. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.08 M C ₁₂ Ala-Phe Aqueous Solution at pH 6.8 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	85
TABLE 3.8. HPLC Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.05 M C ₁₂ Gly-Phe Aqueous Solution at pH 6.5 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	85
TABLE 3.9. HPLC Peak Area, Observed and Normalized Yields for the Reaction of 16-ArN ₂ ⁺ with 0.08 M C ₁₂ Ala-Phe-Glu-Glu-Glu Aqueous Solution at pH 6.0 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.....	86
TABLE 3.10. Estimated Interfacial Molarities of the Nucleophiles from Chemical Imaging in Aqueous 0.100 M of C ₁₂ Gly-Glu at pH 4.5 (Gel) and pH 6.0 (Solution), respectively.....	92
TABLE 3.11. Estimated Interfacial Molarities of the Nucleophiles from Chemical imaging in Aqueous 0.100 M of C ₁₂ Ala-Glu at pH 4.0 (Gel) and pH 6.0 (Solution), respectively.....	93
TABLE 3.12. Estimated Interfacial Molarities of the Nucleophiles from Chemical Imaging in Aqueous 0.080 M of C ₁₂ Ala-Phe at pH 5-6 (Gel) and pH 6.8 (Solution), respectively.....	93
TABLE 3.13. Estimated Interfacial Molarities of the Nucleophiles from Chemical Imaging in Aqueous 0.050 M of C ₁₂ Gly-Phe at pH 6.5.....	93
TABLE 3S4-1. Equations Used to Fit HPLC Calibration Curves of the Relevant Key Products.....	115
TABLE 3S6-A. Chemical Fingerprint of C ₁₂ Ala-Phe at a Concentration of 80 mM at pH 5 – 6 at 40°C (Gel).....	121
TABLE 3S6-B. Chemical Fingerprint of C ₁₂ Ala-Phe at a Concentration of 100 mM at pH below 6.5 (rt: Gel, 48 hrs, 10 µL injection).....	121

TABLE 3S6-C. Chemical Fingerprint of C ₁₂ Gly-Glu at a Concentration of 100 mM at pH 4.5 (rt, solution).....	121
TABLE 4.1. pH Values at Various Concentrations of NaLS at Ambient Temperature.....	134
TABLE 4.2. <i>Z/E</i> Rotamer Ratios of the Sarcosine Surfactant Below and Above the cmc.....	138
TABLE 4.3. Peak Area Ratios of the <i>Z</i> - and <i>E</i> -Rotamers at Low Surfactant Concentrations in Protonated (HLS) and Deprotonated (NaLS) Forms in D ₂ O and CD ₃ OD.....	140
TABLE 4.4. Chemical Shifts and the Differences in the Chemical Shifts of the -CH ₂ CO ₂ D Hydrogens of DLS and -CH ₂ CO ₂ ⁻ Hydrogens of NaLS Adjacent to a Negative Charge in D ₂ O and CD ₃ OD.....	144
TABLE 4S-1a. Melting Points of isolated and purified surfactants and the precipitates collected by cooling the cloudy suspension of NaLS from titration of 27.2 mM NaLS, initial pH 6, no buffer.....	156
TABLE 4S-1b. The volume of aqueous 0.01 M NaOH needed to titrate <i>N</i> -lauroylsarcosine obtained from a commercial sample and the sample collected by cooling the cloudy suspension of NaLS from titration of 27.2 mM NaLS, initial pH 6, no buffer.....	156

List of Schemes

SCHEME 1.1. Principle of Independent Reactivity in Bulk Aqueous Phase and Aggregate Pseudophase for a Unimolecular Reaction.....	13
SCHEME 1.2. Principle of Independent Reactivity in Aqueous Region and Micellar Region for a Bimolecular Reaction.....	14
SCHEME 1.3. Possible Reaction Pathways of Arenediazonium Ion.....	16
SCHEME 1.4. Representative Dediazonation Reactions.....	17
SCHEME 2.1. Reaction Products from Heterolytic Dediazonation and the Structures of <i>N</i> -Lauroylglycine, C ₁₂ Gly, <i>N</i> -Lauroylsarcosine, C ₁₂ Sar and <i>N</i> -Lauroylalanine, C ₁₂ Ala.....	24
SCHEME 2.2. Synthetic Routes for Preparations of 16-ArN ₂ BF ₄ and of All Reaction Products Formed During Dediazonation Reactions in C ₁₂ Ala, C ₁₂ Sar and C ₁₂ Gly Micelles.....	26
SCHEME 2.3. Tagging and Hydrolysis Pathways for the Chemical Imaging Reactions of 16-ArN ₂ ⁺ in C ₁₂ Sar, C ₁₂ Ala and C ₁₂ Gly Micelles.....	40
SCHEME 2.4. Proposed General Base Catalyzed Formation of 5- <i>n</i> -Hexadecyl-7-methyl-1H-indazole.....	41
SCHEME 3.1. Proposed Reaction Mechanism for Trapping of 16-ArN ₂ ⁺ by the Amide Oxygen.....	73
SCHEME 3S-1. Example Synthetic Route of Peptide Amphiphile, C ₁₂ AE.....	101
SCHEME 4.1. All Equilibria of the Protonated and Deprotonated Forms of <i>N</i> -Lauroylsarcosine, and Its <i>Z</i> and <i>E</i> Rotamers that are Present at the Air/Water Interface, in Bulk Solution, and in the Interfacial Region of the Micellar Pseudophase.....	127
SCHEME 4.2. Illustration of <i>N</i> -Lauroylsarcosine Distribution between Various Forms at the Air/Water and Micellar Core/Water Interfaces.....	128
SCHEME 4.3. Acid/Base and Ion Exchange Equilibria in Micellar Solutions of NaLS.....	129

Abbreviations

1-ArOAc = 4-methyl-2, 6-dimethylphenylacetate

1-ArOGAc = 4-methyl-2, 6-dimethylphenyl-*N*-acetyl glycinate

1-ArOGGAc = 4-methyl-2, 6-dimethylphenyl-*N*-acetyl glycylglycinate

16-ArBr = 4-*n*-hexadecyl-2, 6-dimethylbromobenzene

16-ArEC₁₂ = 4-*n*-hexadecyl-2, 6-dimethylphenyllaurate ester

16-ArF = 4-*n*-hexadecyl-2, 6-dimethylfluorobenzene

16-ArInd = 5-*n*-hexadecyl-7-methyl-1H-indazole

16-ArNHAc = 4-*n*-hexadecyl-2,6-dimethylphenyl acetamide

16-ArOH_h = 16-ArOH formed by hydrolysis and C-N bond cleavage of 16-ArOI

16-ArOH_w = 16-ArOH formed by trapping of 16-ArN₂⁺ by water

16-ArOI = Imido ester intermediate formed by amide oxygen trapping of 16-ArN₂⁺

16-ArEGC₁₂ = **16-ArGE** = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylglycinate ester

16-ArESC₁₂ = **16-ArSE** = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylsarcosinate ester

16-ArEAC₁₂ = **16-ArAE** = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alaninate ester

16-ArEFAC₁₂ = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alanine-*L*-phenylalaninate ester

16-ArEFGC₁₂ = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylglycine-*L*-phenylalaninate ester

16-ArEEAC₁₂ (alpha) = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alanine-*L*-glutamate- α -ester

16-ArEEAC₁₂ (gamma) = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alanine-*L*-glutamate- γ -ester

16-ArEEGC₁₂ (alpha) = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylglycine-*L*-glutamate- α -ester

16-ArEEGC₁₂ (gamma) = 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylglycine-*L*-glutamate- γ -ester

z-ArN₂BF₄ = 4-*n*-alkyl-2, 6-dimethylbenzenediazonium ion, z = 16, alkyl = hexadecyl, z = 1, alkyl = methyl

z-ArH = 4-*n*-alkyl-3, 5-dimethylbenzene, z = 16, alkyl = hexadecyl, z = 1, alkyl = methyl

z-ArNH₂ = 4-*n*-alkyl-2, 6-dimethylaniline, z = 16, alkyl = hexadecyl, z = 1, alkyl = methyl

z-ArOH = 4-*n*-alkyl-2, 6-dimethylphenol, z = 16, alkyl = hexadecyl, z = 1, alkyl = methyl

CTAB = Cetyltrimethylammonium bromide

NaLG = Sodium *N*-lauroylglycinate

NaLS = Sodium *N*-lauroylsarcosinate

NaLA = Sodium *N*-lauroyl-*L*-alaninate

HAS = *N*-acetylsarcosine

HLG = C₁₂Gly = *N*-lauroylglycine

HLS = C₁₂Sar = *N*-lauroylsarcosine

HLA = **C₁₂Ala** = *N*-lauroyl-*L*-alanine

C₁₂Ala-Glu = *N*-lauroyl *L*-alanyl-*L*-glutamic acid

C₁₂Ala-Phe = *N*-lauroyl *L*-alanyl-*L*-phenylalanine

C₁₂Gly-Glu = *N*-lauroylglycyl-*L*-glutamic acid

C₁₂Gly-Phe = *N*-lauroylglycyl-*L*-phenylalanine

C₁₂Ala-Phe-Glu-Glu-Glu = *N*-lauroyl-*L*-alanyl-*L*-phenylalanyl-*L*-glutamyl-*L*-glutamyl-*L*-glutamic acid

H₂O_m, H₂O_w, X_m, X_w = Interfacial and aqueous molarities of water and anionic nucleophiles, respectively.

CsLS = Cesium *N*-lauroylsarcosinate

LiLS = Lithium *N*-lauroylsarcosinate

MLS = Alkali metal *N*-lauroylsarcosinate

TMALS = Tetramethylammonium *N*-lauroylsarcosinate

cmc = Critical micelle concentration

CD = Circular dichroism

k_{obs} = Observed rate constant

IR = Infrared spectroscopy

LA = Lauric acid

MALDI-TOF = Matrix assisted laser desorption/ionization – time-of-flight mass spectrometry

DHAP = 2, 5-Dihydroxyacetophenone

PA = Peptide amphiphile

PDB = Protein data bank

S_w^x = Selectivity of the trapping reaction toward an anionic nucleophile, X, compared to water.

SPPS = Solid-phase peptide synthesis

TM = Transmembrane

UV = Ultraviolet

XRC = X-ray crystallography

FEG = Field emission gun

***i*-PrOH** = 2-propanol

MeOH = Methanol

ACN = Acetonitrile

PIE = Pseudophase ion-exchange model

HLB = Hydrophilic lipophilic balance

Chapter I. Introduction

Membrane protein structures are difficult to determine. This dissertation provides a chemical concept that is potentially useful in obtaining topological and structural information of membrane proteins and peptides at the biomimetic interfaces. Chapter I begins with a brief review on membrane protein structural studies in the past decades, followed by introducing surfactants, i.e., amphiphiles, and a chemical model that helps better understand chemical reactions at the interfaces of self-assembled aggregates of surfactants, models to mimic protein/peptide biomembrane interface. This chapter ends with introducing the rationale of the chemical concept that could help reveal proteins/peptides structural information at interfaces.

1.1 Membrane proteins and their structural elucidation

Elucidating molecular organization of biological molecules and their aggregates is a continuing problem that, once solved, could assist in materials research, drug design and the development of therapies.^{1,2} For example, the structural elucidation of membrane proteins (**Figure 1.1**) is of great importance, but only a few structures have been determined. Being a large bio-molecule family and encoded by an estimated 30-40% of the sequenced mammalian genome,^{1,2} membrane proteins play a vital role in a large variety of eukaryotic cellular processes, such as transport (e.g. nutrient/waste uptake and export), cell-molecule recognition, cell-cell communication and cell division.^{2,3} Any imbalance, mutation, or depletion in the cell membrane proteins may be related to a large number of human diseases such as diabetes, cystic fibrosis and some forms of cancer.³ Therefore, membrane proteins structural information is becoming more important in drug design and formulation, as well as the development of therapies. More than half of the drugs currently on the market or under development as active pharmaceutical ingredients (APIs), find membrane protein as a target. However, structural characterization of membrane proteins remains “far behind that of their water soluble counterparts”,⁴ and only contribute ca. 1%

of the currently solved protein structures deposited in the Protein Data Bank (PDB).² This is because membrane protein structures are not fixed by covalent bonds, but generally organized by multiple hydrophobic, electrostatic, hydration, and hydrogen bonding interaction at interfaces with or without the direct participation of membrane components.³ Therefore, many technical challenges exist, including protein expression, purification (*esp.* eukaryotic membrane proteins),²

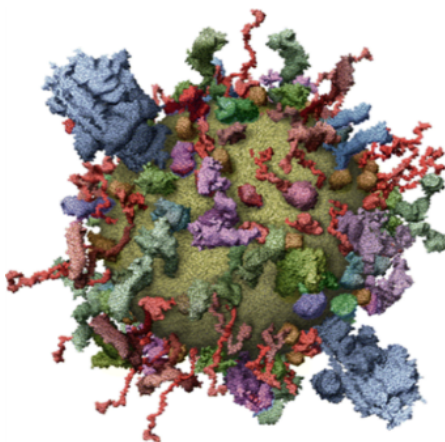


Figure 1.1. Molecular Model of the Cross Section of an Average Synaptic Vesicle Showing the Orientations of Different Types of Proteins and the Membrane Phospholipids.³⁸

solubilization, and selection of the appropriate membrane mimetic systems that will retain the stability and function of membrane proteins,^{5,6} Despite these difficulties, the number of solved membrane protein structure has grown considerably over the past two decades,² because of rapid progress in the classical techniques and the emergence of a number of new approaches, such as x-ray crystallography, circular dichroism, electron microscopy, infrared spectroscopy, nuclear magnetic resonance, computational simulations and a series of chemical modification methods, ranging across biology, chemistry, biophysics, and computational science.

X-ray crystallography (XRC) is a “direct” structural method that provides accurate structural information of the membrane proteins at the atomic level. XRC provides the precise arrangement of atoms within crystal structures by converting the diffraction pattern of X-rays to an electron density map (**Figure 1.2**),² followed by a series of mathematical treatments. Till recently,

however, only about 100 unique membrane protein structures have been determined via XRC.⁷ The slow step, however, is the isolation and purification of sufficient quantities of high quality crystals, which requires the search for proper crystallization conditions from various detergent solutions.⁶ This is a difficult step despite of a number of recent advances.² Another issue is the 3-D structure of the membrane “environment”.

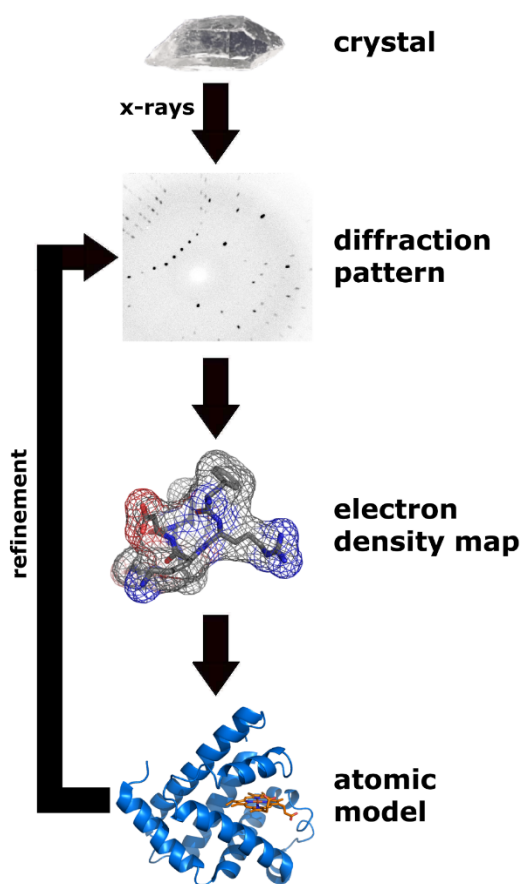


Figure 1.2. Flow Chart of Protein Structural Elucidation by X-ray Diffraction.⁷

A second important technique is circular dichroism (CD) which is used to estimate the secondary structure of proteins, often in the presence of membrane mimics such as liposomes and micelles.⁸ The CD spectrum for a given membrane protein is approximated by a linear combination of spectra of several known secondary structures of a protein (**Figure 1.3**).^{8,9} However, CD does not provide accurate and detailed secondary structural information, but only

an approximation of the percentage for each specific secondary structure type, *i.e.*, α -helix and β -sheet.^{8,10} To make matters worse, the presence of membrane mimics like liposomes may interrupt data interpretation of CD measurements by light scattering, making the determination of protein concentration less reliable and the data collection below 200 nm virtually impossible.⁶

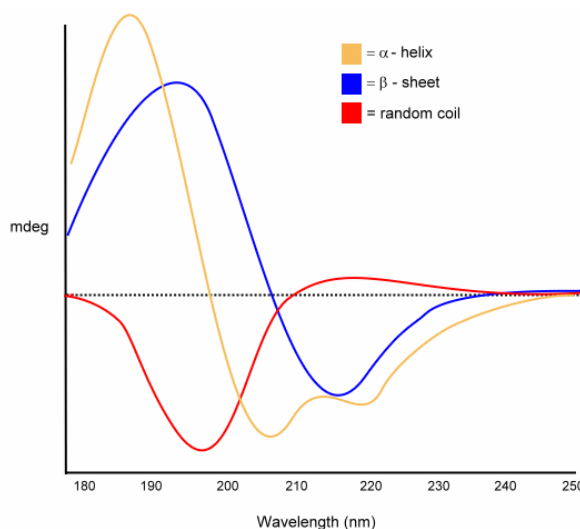


Figure 1.3. Examples of Different Pure Secondary Structures Detecting by Circular Dichroism.¹⁰

Electron microscopy (EM) is a relatively new promising technique. Measurements of membrane protein structures by EM are conducted in reconstituted lipid bilayers and even native membrane, instead of detergents, which is usually a prerequisite for structural studies by X-ray and solution NMR.² EM is capable of achieving as high resolution as X-ray diffraction.² However, to obtain highly precise resolution for samples with good quality, the electron source must be very coherent, which is achieved by using the field emission gun (FEG) as the electron source.² Second, the radiation damage to the sample remains a problem, but must be minimized via cryo-electron microscopy (cryo-EM), in which samples are prepared at a very low temperature under high vacuum.²

Infrared spectroscopy (IR) is a simple yet powerful tool for membrane protein structure determination. One advantage of IR is that the presence of lipid environment does not affect the

resolution of spectra,¹¹ such that IR analysis could report on protein-membrane interactions, e.g., conformational changes of membrane proteins induced by the membrane or other factors in the environment.¹² As with CD, interpretation of amide bands by IR spectra can provide useful information on the percentage of different secondary structure types.⁸ Time-resolved FT-IR, provides a tool that is capable of detecting weak absorbance signals.^{8,12} IR is usually co-used with isotopic labeling method in order to determine the exact location of amide bonds inside peptides.^{6,13}

Nuclear magnetic resonance (NMR) is becoming a powerful tool in determining membrane protein structures in membrane mimics over the last couple of decades (**Figure 1.4**),⁸ despite a number of limitations on the sensitivity,¹⁴ the size of the bio-molecules under investigation and so forth.¹⁵ With the careful selection of the detergent and the membrane mimetic systems being

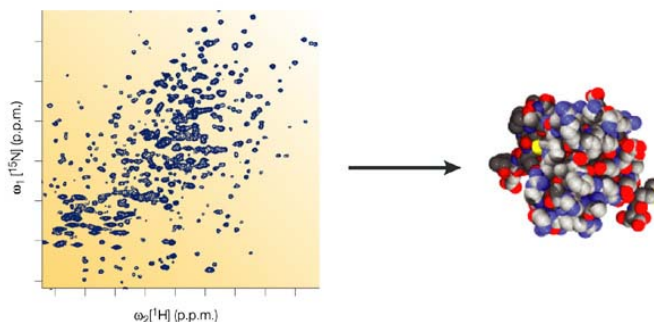


Figure 1.4. NMR Spectroscopy and Protein Structure Determination. An example of the complex pattern obtained from NMR spectroscopy (left), which correlates the nitrogen atoms (¹⁵N; vertical axis) of amide groups with the hydrogen atoms (¹H; horizontal axis) in a protein. A complex series of measurements that correlate other nuclei (carbon-hydrogen, hydrogen-hydrogen), calculations and deductions allows the structure of the protein to be derived (right).³⁹

selected properly, including cationic/zwitterionic micelles^{6,16} and bicelles,^{6,15,16,17} solution and solid-state NMR techniques currently undergoes rapid development and have helped elucidate a number of membrane protein structures.^{2,5,6,16,18,19,20,21,22} Especially, solid-state NMR technique, which targets membrane protein entities adsorbed or bound with lipid bilayers bearing solid

carriers, has demonstrated a high capacity to determine high-resolution structure and topology of membrane proteins as well as their interactions with lipids.^{23,5,6,15,18,15,19}

Molecular dynamics (MD) simulations are an effective approach to aid the understanding of membrane protein structure-function relationships and are a complement to the existing experimental approaches.^{24,25} Computational analysis of the membrane protein structures has helped in many categories of membrane protein studies, for example, exploring the mechanism of the manifold process of the proteins associated with the membranes,²⁶ and so forth.^{27, 28, 29, 30} They are achieved by applying the “force field” to evaluate the empirical energy from the coordinates of the atoms.²⁴

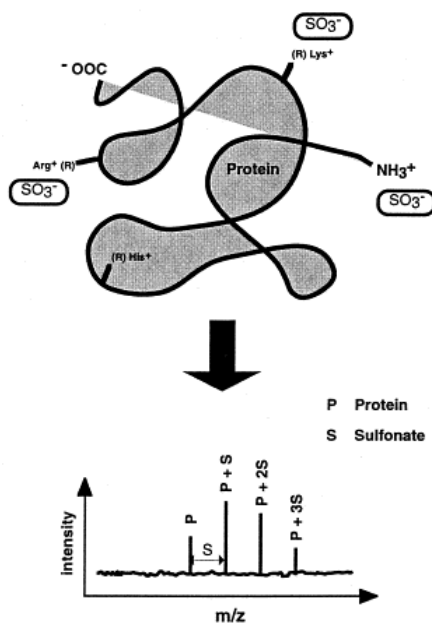


Figure 1.5. Principle of Titration of Basic Amino Acid Residues by Sulfonates.³²

In the recent years, a large number of chemical modification approaches have been developed, mostly depending either upon enzyme, isotopic labeling, proteolytic cleavage, cross-linking, or chemical labeling methods.^{4,31} For example, naphthalene-sulfonic acid derivatives bind specifically to arginine residues and provide information on the exposed surface of the proteins in MALDI-TOF mass spectrometry experiments (**Figure 1.5**).³² Another example is cysteine

scanning mutagenesis, in which the Cys residues in the membrane protein surface are accessible to the labeling reagent. The extent of labeling can be monitored by a number of detection methods such as fluorescence spectroscopy and provide meaningful information on the location of loops and other exposed trans-membrane regions.³⁰ In another recent study, a chemically oxidative labeling approach was used in conjunction with Tandem MS to obtain structural information on membrane proteins. The premise of this approach is that the oxidative labeling of hydroxyl radicals works at solvent-exposed side chains of membrane proteins, bearing one or more specific amino acids, instead of “sterically protected” moieties (**Figure 1.6**).^{33,34} Although these approaches aid in structural elucidation of membrane proteins with moderate to excellent effectiveness, they have limitations.^{11,13,14} An unique approach that will be presented in some detail below, chemical imaging reveals new information on the contribution of ion-pairing and hydration to the balance of forces controlling aggregate morphology.³⁵ In principle, chemical imaging provides experimentally measured estimates of the interfacial concentrations of all the weakly basic nucleophiles located in the interfacial regions of a variety of self-assembled aggregates. We believe that this approach is a potentially valuable for obtaining topological and structural information on membrane proteins. To test this viability, we applied the chemical imaging approach to estimate the interfacial concentrations of a set of weakly basic nucleophiles, including amide bonds that are located at the interfacial region of association colloids prepared from a variety of amino acid and peptide amphiphiles, and are simple mimics of the protein-biomembrane systems.

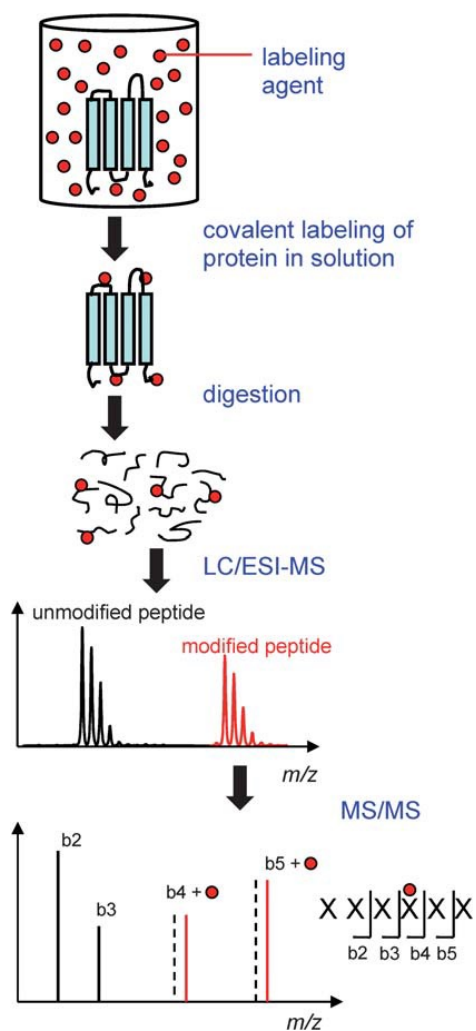


Figure 1.6. General Workflow for Covalent Labeling of Proteins with Subsequent Analysis by Limited Proteolysis (Usually tryptic digestion), Liquid Chromatography (LC) and Electrospray Mass Spectrometry (ESI-MS).^{33,34}

1.2. Introduction to Surfactants

Amino acid/peptide amphiphiles belong to one special class of molecules called surfactants. Surfactants are amphiphilic molecules with distinguishable polar and nonpolar segments that are generally water soluble as both monomers and self-assembled aggregates. The name “surfactant” are surface active agents capable of substantially reducing surface tension at oil-water and air-water interfaces. Not only are surfactants widely used in a variety of scientific areas such as polyacrylamide gel electrophoresis, membrane protein solubilization and crystallization, but also used as membrane mimetics in *in vitro* studies and as carriers for substance delivery such as small drug molecules and macromolecules.⁴⁰

Surfactants can be categorized into “stick-like” and “bean-like” sub-classes. A large variety of surfactants belong to the first class, meaning that they both have a “water-loving” component, i.e., a hydrophilic headgroup, and a “water-loathing” component that is usually one or more hydrocarbon tails. More commonly, however, surfactants are categorized into non-ionic and ionic sub-classes, depending on the charge of their headgroups. Ionic surfactants are further categorized into cationic, anionic and zwitterionic surfactant (**Figure 1.7**).⁴¹ The amino acid/peptide surfactants/amphiphiles used in this study are anionic surfactants. A common property of all surfactants is that they self-assemble at or above a specific concentration, called the critical micelle (aggregation) concentration or cmc (cac), to form water-soluble aggregates with various shapes with dimensions ranging across ca. five to hundreds of nanometers. Examples are spherical, rod-like micelles, vesicles, nanofibers, lamellar bilayers, hexagonal reversed phase, *etc.* (**Figure 1.8**).⁴² The primary driving force of self-aggregation is the hydrophobic interactions that contribute to minimize the Gibbs free energy of the colloidal solution. The strength of hydrophobic interactions arise from the thermodynamics of surfactant dispersed in the aqueous medium, driven by entropy.⁴³ The balance of forces is provided by a number of weak, noncovalent interactions, such

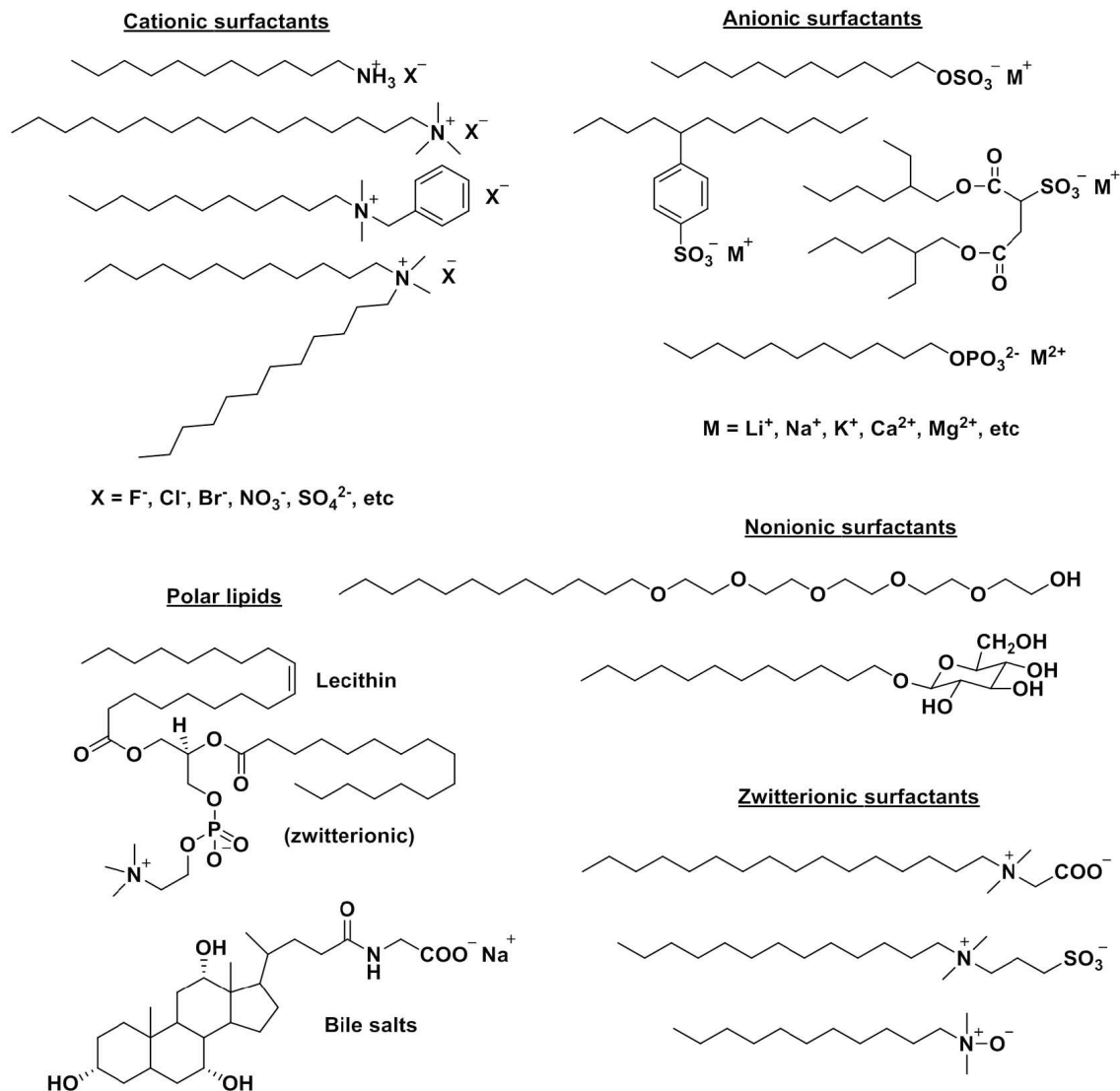


Figure 1.7. Representative Surfactants and Their Structures.⁴¹

as electrostatic and Van der Waals interactions, of surfactant headgroups, which maintain association colloidal system thermodynamically and kinetically stable. Therefore, physical properties of surfactants and their self-assembled aggregates in aqueous solutions, such as shape, aggregation number, Krafft temperature, cloud point and probably most important, the critical aggregation concentration, are highly variable and mainly depend on the hydrophobic and hydrophilic components of their structures, as well as the medium conditions such as temperature, pH, electrolyte concentrations and ionic strength.

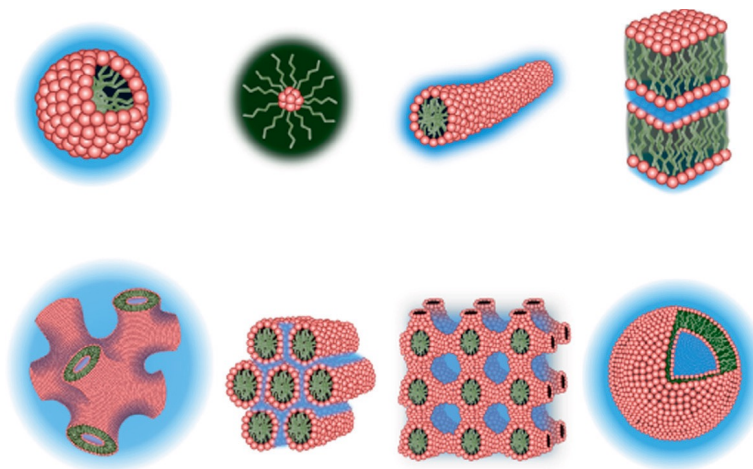


Figure 1.8. Representative Self-assembled Aggregates in Association Colloids.⁴² Top (from Left to Right): Spherical Micelle, Reversed Micelle, Elongated Micelle (Worm or Rod-like) and Lamellar Sheets. Bottom (from Left to Right): Bicontinuous (Disordered Sponge), Hexagonally Packed Rod Micelle, Cubic and Vesicle.

1.3. Chemical Reactivity in Association Colloids

Biomembranes are made of lipid bilayers whose organization is driven by hydrophobic effects and stabilized by a number of non-covalent interactions. A large number of mimetic systems have been used to model the biomembrane environment, such as micelles, bicelles, vesicles and nanodiscs.^{44,45} These systems belong to a broad class of soft materials, named association colloids. Below is a brief introduction to association colloid properties that are used to characterize molecular organization, interactions and chemical reactivity of biomembrane.

Association colloids are homogenous, thermodynamically stable solutions in which surfactant molecules self-assemble into mesophases.⁴⁶ Micelles, for example, the simplest association colloid composed of surfactant and water, are dynamic aggregates in which the monomer is in constant, rapid exchange between micelles and water. Surfactants and other small molecules move at or near their diffusion-controlled limits in and out of the micelle, such that the distribution of these components, i.e., surfactants, counter-ions and water, are in dynamic equilibrium that is similar to that of components of non-micellar solutions. In 2015, Romsted *et al.* proposed that surfactant-based association colloids can be treated as either discrete structures or

separate regions, depending on methods used for measurement of their properties.⁴⁷ Physical observations of colloidal solutions, such as light scattering and microscopies, provide structural snapshots that are captured within the limited time. Other methods such as chemical reactions, however, sense the colloidal solutions as a whole, because the reactions are orders of magnitude slower than diffusion of the reactants, but are in equilibrium distribution between the molecular assemblies present. The reactions sense the properties and total volume of the different regions, but not their sizes and shapes.

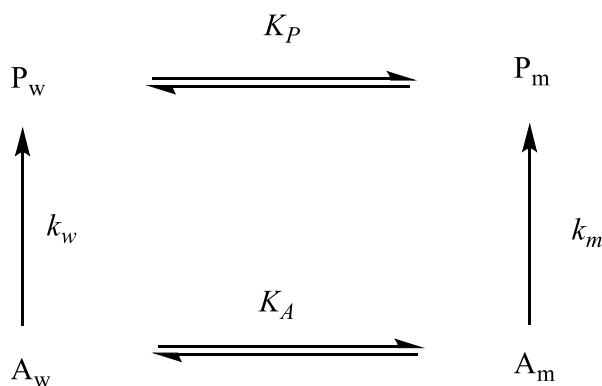
Supra-molecular assemblies such as micelles, vesicles, nano-fibers, even poly-electrolytes like proteins and DNA, share a common structural feature: an interfacial region with moderate polarity that is similar to that of short chain alcohol solutions, adjacent to an aqueous region of high polarity.⁴⁸ Properties in bulk region like pH, ionic strength, degree of ionization (α) and counterion binding (β), composition and conformations of small/macro-molecules. *etc.*, need not be the same within the interfacial region. However, various interfacial parameters, such as organization of the local headgroups, counter-ion concentrations and ion binding affinities contribute to a number of bulk properties such as cmc, size, shape and aggregation number. For these reasons, studies of interfacial regions of association colloids provide important information on their molecular organization and the interactions that contribute to their properties. A number of physical methods, such as conductivity, NMR, UV/visible and sum frequency generation spectroscopy,⁴⁷ are used to explain interfacial properties. Most of these methods, however, are only amenable to monitor one component or property at a time.⁵³ Chemical reactions can also be used to probe the interfacial region of association colloids. Models that interpret the effects of aggregation on reactivity in association colloids were developed in the recent decades. One important example is the pseudophase ion-exchange model (PIE).

The pseudophase ion-exchange model provides the conceptual basis for using chemical reaction as methods to probe molecular organization and composition of associate colloids. In the

PIE model, association colloidal compartments are treated as an independent phase that is separated from the bulk solution. The concentration of surfactant present in self-assembled aggregates is expressed as **Equation 1.1**, in which $[S_T]$ and $[S_m]$ represent the total surfactant concentration and the surfactant concentration in self-assembled aggregates:

$$[S_m] = [S_T] - cmc \quad (\text{Eq. 1.1})$$

For a spontaneous reaction, the rate of the reaction is generally a first-order. Reactant, A, is assumed to partition between the colloidal pseudophase and the bulk region, with the binding constant, K_A , used to describe its distribution in two phases.



Scheme 1.1. Principle of Independent Reactivity in Bulk Aqueous Phase (*w*) and Aggregate Pseudophase (*m*) for a Unimolecular Reaction.

$$K_A = \frac{[A_m]}{[A_w][m]} \quad (\text{Eq. 1.2})$$

As previously mentioned, because the rate of reactions in association colloids are essentially orders of magnitude slower than that of molecular diffusivities in and out of self-assembled aggregates, the overall rate of the reaction is the sum of the rates in water and in the aggregates:

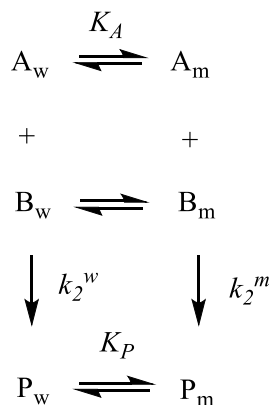
$$\text{Rate} = k_{obs}[A_T] = k_w[A_w] + k_m[A_m] \quad (\text{Eq. 1.3}), \quad [A_T] = [A_w] + [A_m] \quad (\text{Eq. 1.4})$$

Therefore, the observed reaction rate depends on rate constants and the concentrations of reactants both in the aggregate pseudophase and in bulk region. The observed rate constant for a unimolecular reaction is then represented by the equation below:

$$K_{obs} = (k_w + k_m K_A [S_m]) / (1 + K_A [S_m]) \quad (\text{Eq. 1.5})$$

Note that when $[S_T] \leq \text{cmc}$, k_{obs} is equal to k_w , and when $[S_T] \gg \text{cmc}$, k_{obs} approaches k_m infinitely. Therefore, **Equation 1.5** shows that the observed rate constant in the solution is dependent on the rate constants both in the aggregate phase and the bulk phase, as well as the distribution of the reactant in the two phases.

For a bimolecular reaction, the rate of the reaction is a second-order. In association colloid solutions, the reactant, A, is assumed to partition between the colloidal pseudophase and the bulk region, with the binding constant, K_A , used to describe its distribution between two regions.



Scheme 1.2. Principle of Independent Reactivity in Aqueous Region (*w*) and Micellar Region (*m*) for a Bimolecular Reaction.

As one example, in dediazonation, arenediazonium ion (A) reacts with a nucleophile (B) and the reaction is bimolecular. Both reactants are in rapid, dynamic equilibrium between the aggregate and bulk phases. By experimentally setting the concentration of nucleophile (B) in large excess over that of arenediazonium ion (A), the overall reaction follows a pseudo first order.

$$\text{Rate} = k_{obs}[A_T] = k_2[A_T][B_T] = k_2^w[A_w][B_w] + k_2^m[A_m][B_m] \quad (\text{Eq. 1.6})$$

Note: (B_m) stands for the estimated local concentration of the nucleophile B.

Therefore, the observed rate constant for a bimolecular reaction is represented by the equation below:

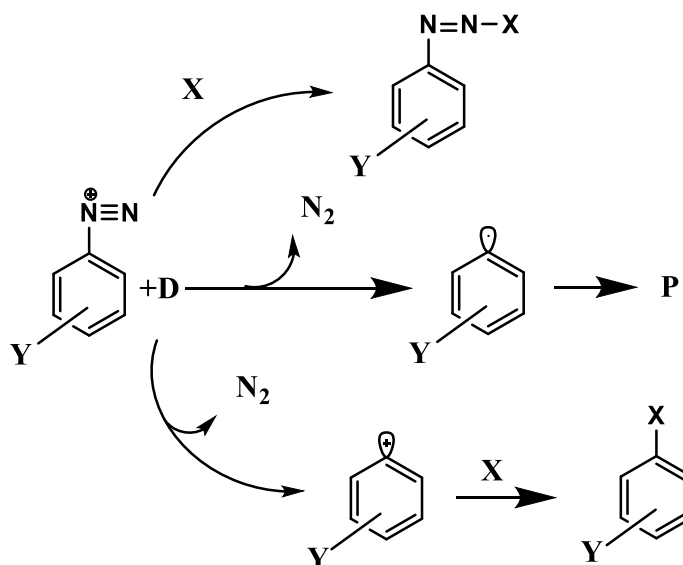
$$K_2 = \frac{K_{\text{obs}}}{[B_T]} = \frac{k_2^w}{(1 + K_A[S_m])(1 + K_B[S_m])} + \frac{K_A K_B [S_m] k_2^m / V_m}{(1 + K_A[S_m])(1 + K_B[S_m])} \quad (\text{Eq. 1.7})$$

Note that when $[S_T] \leq \text{cmc}$, k_2 is equal to k_2^w , and when $[S_T] \gg \text{cmc}$, k_2 is equal to $(k_2^m / V_m) / ([S_m])$. Using an estimate for the interfacial volume hypothesized (V_m is usually assumed to be the Stern layer volume), k_2^m is obtained.

Equations 1.5 and 1.7 show that chemical reactions can be used to probe the compositions (e.g. local concentrations) and properties (e.g., rate constants) in the interfacial region of association colloids. However, chemical methods that provide information on the composition of multiple charged and uncharged molecules simultaneously are very rare. The above interpretation provides the basis for the chemical imaging method in which an amphiphilic, aggregate bound arenediazonium ion probe reacts with a variety of weakly basic nucleophiles whose concentrations are in large excess (see the next section).

1.4. The Rationale of the Present Work

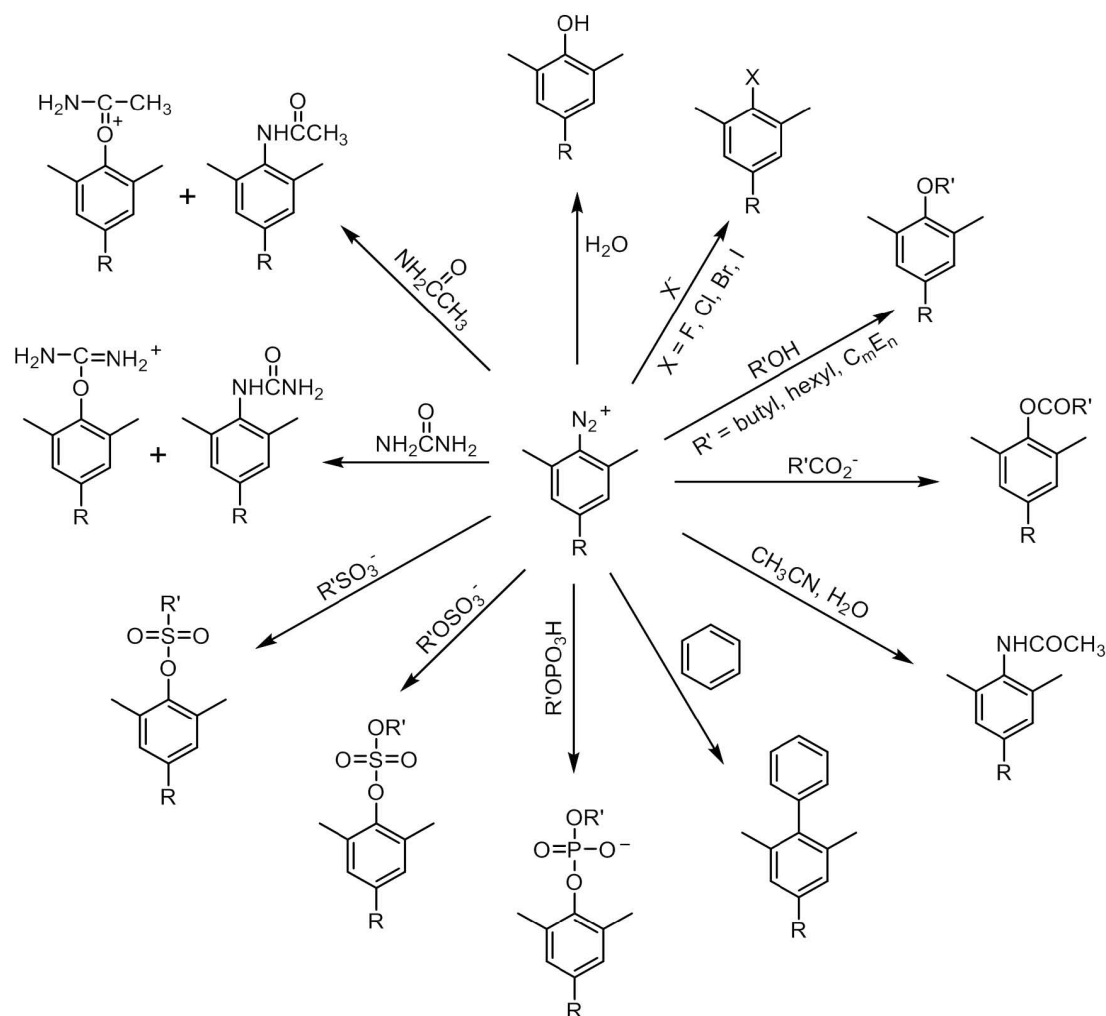
The rationale of this chemical approach is grounded in the heterolytic dediazonation chemistry that has been well studied for decades.^{36,37} Arenediazonium ion has a complex chemistry and their reactions depend on substituents type, solvent type and reactants such as nucleophiles, *etc.* Arenediazonium ions generally undergo three types of reactions (**Scheme 1.3**):⁴⁹ 1) addition to the terminal nitrogen by strongly basic nucleophiles such as OH⁻ and N₃⁻;



Scheme 1.3. Possible Reaction Pathways of Arenediazonium Ion. Y = Substituents, D = Electron Donors, X = Nucleophiles, P = Products.⁴⁹

2) homolytic displacement of the diazo group by electron transfer from an electron donor; and 3) heterolytic loss of N_2 in the reactions with weakly basic nucleophiles. The rate-determining step of the third pathway at neutral to acidic solution pHs in the absence of UV light and reducing reagents is loss of N_2 , followed by a fast and competitive reaction with a large variety of weakly basic nucleophiles including water, halide ions, urea, and those commonly found in biologically-related material, such as amide carbonyl oxygen and carboxylate groups (**Scheme 1.4**).⁴⁹ The observed rate constant for dediazonation is highly insensitive to solvent polarity.^{36,49} In solutions containing two or more nucleophiles, whose concentrations are in large excess over that of the arenediazonium ion, product yields are proportional to the concentrations and their selectivities, S_w^X , are generally small, on the order of 1 to 20.

In the early 1980s, Loughlin and Romsted developed amphiphilic arenediazonium ion and demonstrated that this micellar bound probe provided meaningful information on the local concentrations of counterions that bind competitively within the micellar interfacial region. One



Scheme 1.4. Representative Dediazonation Reactions.⁴⁹

disadvantage of this probe, however, is that solution pH needed to be adjusted at ca. 1 to minimize side reaction that gives rise to a yellow product, probably due to the formation of a dye molecule.⁵¹ To overcome this difficulty and to expand the use of the molecular probe, in the early 1990s Chaudhuri and Romsted prepared a new arenediazonium probe that has both the ortho and para positions protected by the alkyl groups. Results demonstrated that this new class of arenediazonium probe showed no significant side reaction at a wide pH range up to 8-9 in a non-aggregate aqueous solution, and 6 - 7 in aggregate solutions.^{51,52}

Therefore, an amphiphilic-like arenediazonium ion is able to function as a molecular probe at the interfacial region of amphiphile (same as surfactant hereafter) aggregates and bilayers. By carefully optimize its concentration in amphiphile solutions, which is normally $\sim 1\%$ of that of amphiphile molecules, arenediazonium ion should react competitively and reproducibly with interfacial H_2O , oxygen of amide bonds, as well as reactive sidechains in a peptide of known sequence that is also in the interfacial region, but not those in the core or in the surrounding aqueous region (**Figure 1.9**). Products formed in reaction solutions are separated, characterized and quantified by a number of chromatographic, spectroscopic analytical methods, such as HPLC, MALDI-TOF mass spectrometry and HPLC spiking experiments with calibration curves that are

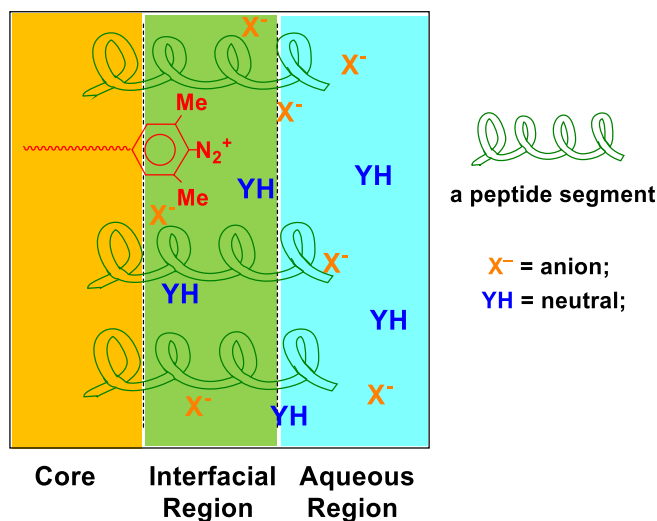


Figure 1.9. Illustration of a small section of the immediate vicinity of the interfacial region of membrane mimic in aqueous solution, in which an hydrophobic arenediazonium ion is trapped by weakly basic nucleophiles, H_2O (not shown), X^- (anionic), YH (neutral) and those on peptide segment (helices in green). Different colors are used to distinct micelle core (orange), interfacial region (light green) and the bulk aqueous region (blue).

developed from a series of independently synthesized compounds. Analyses of various tagging and fragmentation patterns, i.e., the chemical “fingerprints”, provide information on the locations and local concentrations of peptide backbone and reactive sidechains within the interfacial region. To estimate these local concentrations, we assume that the selectivity of a set of two competitive dediazonation reactions with one weakly basic nucleophile versus another at the association

colloidal interfaces, as determined by the amphiphilic molecular probe, are essentially the same as that in bulk solution in the absence of self-assembled aggregates, determined by the short chain analogue of the amphiphilic probe. This assumption is very important throughout the thesis and it is represented by the equation below:

$$\frac{\%(1 - \text{ArX})}{\%(1 - \text{ArOH})} \frac{[\text{X}_w]}{[\text{H}_2\text{O}_w]} = \frac{\%(16 - \text{ArX})}{\%(16 - \text{ArOH})} \frac{\text{X}_m}{\text{H}_2\text{O}_w}$$

Eq. 1.8

The validity of this assumption is based on the fact that dediazonation reactions are generally insensitive of the solution composition, solvent polarity and the medium viscosity.⁴⁹ The degree of cleavage and tagging of the peptide models depends on the distribution of each specific peptide (amide) bond and sidechain in the interfacial region. Therefore, the results provide “fingerprints” of themselves in the interfacial region. This approach should also provide new insight on the

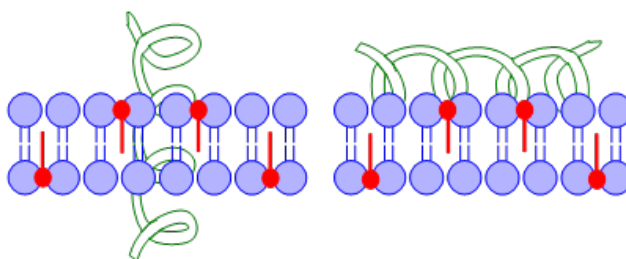


Figure 1.10. An α -helix (green) peptide segment oriented either perpendicular (left) or parallel (right) to a small section of membrane bilayer (purple) that contains molecular probes with reactive functional groups located in the interfacial region (circled).⁵⁰ It is assumed that only peptide segments at the interfacial region react with the molecular probe. Those buried in the core and extended in the bulk region do not react.

orientations and secondary structures of peptide at interface, *e.g.*, distinguishing between α -helix, β -sheet, and perhaps random coil structures (**Figure 1.10**), which complement results obtained by a number of traditional approaches such as crystallography, EM, IR, CD, NMR, *etc.* Membrane proteins are biopolymers made of up to 20 types of natural *L*-amino acids, associated via amide

covalent bonds, with certain amino acid sequences, and stabilized by a certain 3-D conformations under some circumstance. Undoubtedly, its inherently polar backbone must be shielded from the hydrophobic component of the lipid bilayer via the adoption of secondary structures, like α -helices, β -sheets and β -turns, in order to maintain its stability at the biomembrane system. Consequently, some or all of the amide bonds and side chains/headgroups might be located in the interfacial region. Being good weakly basic nucleophiles, amide oxygen, nitrogen, and reactive side chains/headgroups should conveniently trap the long chain arenediazonium ions, which should then lead to the cleavage of peptide backbones, and the formation of a series of products that can be analyzed both qualitatively and quantitatively.

Bibliography

- (1) Shai, Y. *Biopolymers* **2002**, *66*, 236.
- (2) King, G.; Dixon, A. M. Membrane protein structure: Techniques. In *Wiley Encyclopedia of Chemical Biology*; Begley, T. P., Ed.; John Wiley & Sons: New York, 2008; Vol. 2, p 761.
- (3) Benedetti, E. *Biopolymers* **2000**, *55*, 357.
- (4) Bechinger, B.; Resende, J. M.; Aisenbrey, C. *Biophys Chem* **2011**, *153*, 115.
- (5) Arora, A.; Tamm, L. K. *Curr Opin Struc Biol* **2001**, *11*, 540.
- (6) Speers, A.; Wu, C. *Compr Anal Chem* **2008**, *52*, 213.
- (7) Zhang, H.; Cramer, W. *J Struct Funct Genomics* **2005**, *6*, 219–223.
- (8) White, S. H. Membrane Protein Structure. *The American Physiological Society Methods in Physiology Series*. Oxford University Press. New York. **1994**.
- (9) Sreerama, N.; Woody, R. W. *Protein Science* **2004**, *13*, 100.
- (10) Taylor, J. W.; Osapay, G. *Accounts Chem Res* **1990**, *23*, 338.
- (11) Arrondo, J. L. R.; Goni, F. M. *Prog Biophys Mol Bio* **1999**, *72*, 367.
- (12) Radu, I.; Schleege, M.; Nack, M.; Heberle, J. *Aust J Chem* **2011**, *64*, 9.
- (13) Sackett, K.; Shai, Y. *J Mol Biol* **2005**, *351*, 1169.
- (14) Torres, J.; Stevens, T. J.; Samso, M. *Trends in Biochemical Sciences* **2003**, *28*, 137.
- (15) Marcotte, I.; Auger, M. *Concept Magn Reson A* **2005**, *24A*, 17.
- (16) Kang, C. B.; Li, Q. X. *Curr Opin Chem Biol* **2011**, *15*, 560.
- (17) Morrison, E.; Henzler-Wildman, K. *Bba-Biomembranes* **2012**, *1818*, 814.
- (18) Montaville, P.; Jamin, N.; Lacapere, J.-J. *Methods Mol. Biol.* **2010**, *654*, 261.
- (19) Hong, M.; Su, Y. C. *Protein Science* **2011**, *20*, 641.
- (20) Zerbe, O. *Angew Chem Int Edit* **2012**, *51*, 860.
- (21) Nevzorov, A. A.; Mesleh, M. F.; Opella, S. J. *Magn Reson Chem* **2004**, *42*, 162.
- (22) Jacso, T.; Franks, W. T.; Rose, H.; Fink, U.; Broecker, J.; Keller, S.; Oschkinat, H.; Reif, B. *Angew Chem Int Edit* **2012**, *51*, 432.
- (23) Watts, A. *Bba-Rev Biomembranes* **1998**, *1376*, 297.
- (24) Fuchs, P. F. J.; Lacapère, J.-J. *Molecular Dynamics of Membrane Peptides and Proteins: Principles and Comparison to Experimental Data*; Humana Press: Clifton NJ, 2010; Vol. 654.
- (25) Khalid, S.; Holyoake, J.; Sansom, M. S. P. Molecular dynamics studies of membrane proteins: Outer membrane proteins and transporters. In *Biophysical Analysis of Membrane Proteins: Investigating Structure and Function*; Pebay-Peyroula, E., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008; Chapter 7, p 159.
- (26) Gumbart, J.; Wang, Y.; Aksimentiev, A.; Tajkhorshid, E.; Schulten, K. *Curr Opin Struc Biol* **2005**, *15*, 423.
- (27) Ash, W. L.; Zlomislic, M. R.; Oloo, E. O.; Tieleman, D. P. *Bba-Biomembranes* **2004**, *1666*, 158.
- (28) Arinaminpathy, Y.; Khurana, E.; Engelman, D. M.; Gerstein, M. B. *Drug Discov Today* **2009**, *14*, 1130.
- (29) Friemann, R.; Larsson, D. S. D.; Wang, Y. F.; van der Spoel, D. *J Am Chem Soc* **2009**, *131*, 16606.
- (30) Liang, J.; Naveed, H.; Jimenez-Morales, D.; Adamian, L.; Lin, M. *Bba-Biomembranes* **2012**, *4*, 942.
- (31) Vachet, R. W.; Mendoza, V. L. *Mass Spectrometry Reviews* **2009**, *28*, 785.
- (32) Friess, S. D.; Zenobi, R. *J Am Soc Mass Spectr* **2001**, *12*, 810.
- (33) Pan, Y.; Konermann, L. *Analyst* **2010**, *135*, 1191.

- (34) Konermann, L.; Stocks, B. B.; Pan, Y.; Tong, X. *Mass Spectrometry Reviews* **2010**, 29, 651.
- (35) *Interfacial Compositions of Amphiphile Assemblies by Chemical Trapping with Arenediazonium Ions: Method and Its Applications*; Romsted, L., Ed.; Marcel Dekker: New York, 2001; Vol. 10.
- (36) *Diazo Chemistry I: Aromatic and Heteroaromatic Compounds*; Zollinger, H., Ed.; VCH Publishers: Weinheim, 1994.
- (37) *The chemistry of the Diazonium and Diazo Groups, Part 2*; Hegarty, A. F., Ed.; John Wiley & Sons: New York, 1978.
- (38) de Jongh H. H. J.; Goormaghtigh, E.; Killian J. A. *Biochemistry* **1994**, 48, 14521-14528.
- (39) Pietzsch, J. Protein folding technology.
- (40) Detergents and Their Uses. Affymetrix, Anatrace Products Catalog.
- (41) K. Holmberg, B. Jönsson, B. Kronberg, and B. Lindman, *Surfactants and Polymers in Aqueous Solution*. John Wiley & Sons, 2002.
- (42) Eastoe, J.; Tabor, R. *Surfactants and Nanoscience. Colloidal Foundations of Nanoscience*. Berti, D. & Palazzo, G.; Ed.; Elsevier, 2014.
- (43) *Biophysical Chemistry*. Allen, J.; Ed.; Wiley-Blackwell, A John Wiley & Sons, Ltd., Publication, 2008.
- (44) Marcotte, I.; Auger, M. *Concept Magn Reson A* **2005**, 1, 17-37.
- (45) Serebryany, E.; Zhu, G.; Yan, E. *Biochim Biophys Acta* **2012**, 1818, 225-233.
- (46) Romsted, L. S. Introduction to surfactant self-assembly. In *Supramolecular Chemistry: From Molecules to Nanomaterials*, 1st ed.; Gale, P. A., Steed, J. W., Eds.; Wiley-Blackwell: Chichester, U.K., 2012; Vol. 1, p 181.
- (47) Bravo-Díaz, C.; Romsted, L. S. *et. al. Langmuir* **2015**, 31, 8961-8979.
- (48) Romsted, L. S.; Yoon, C.-O. *J Am Chem Soc* **1993**, 115, 989-994.
- (49) Romsted, L. S. Interfacial compositions of surfactant assemblies by chemical trapping with arenediazonium ions: Method and applications. In *Reactions and Synthesis in Surfactant Systems*; Texter, J., Ed.; Marcel Dekker: New York, 2001; Vol. 10, p 265.
- (50) Zhang, Y.; Romsted, L. S.; Zhuang, L.; de Jong, S. *Langmuir* **2012**, 29, 534-544.
- (51) Chaudhuri, A. Simultaneous Determination of Interfacial Compositions of Three Component Microemulsions Using Dediazonation Reactions. Ph.D. Thesis. Rutgers, The State University of New Jersey, 1991.
- (52) Zhuang, L. Mechanism of Reaction of Arenediazonium Ions with the Amide Bond. Ph.D. Thesis. Rutgers, The State University of Jersey, 1998.
- (53) Chaudhuri, A.; Loughlin, J.; Romsted, L. S. *J Am Chem Soc* **1993**, 115, 8351-8361.

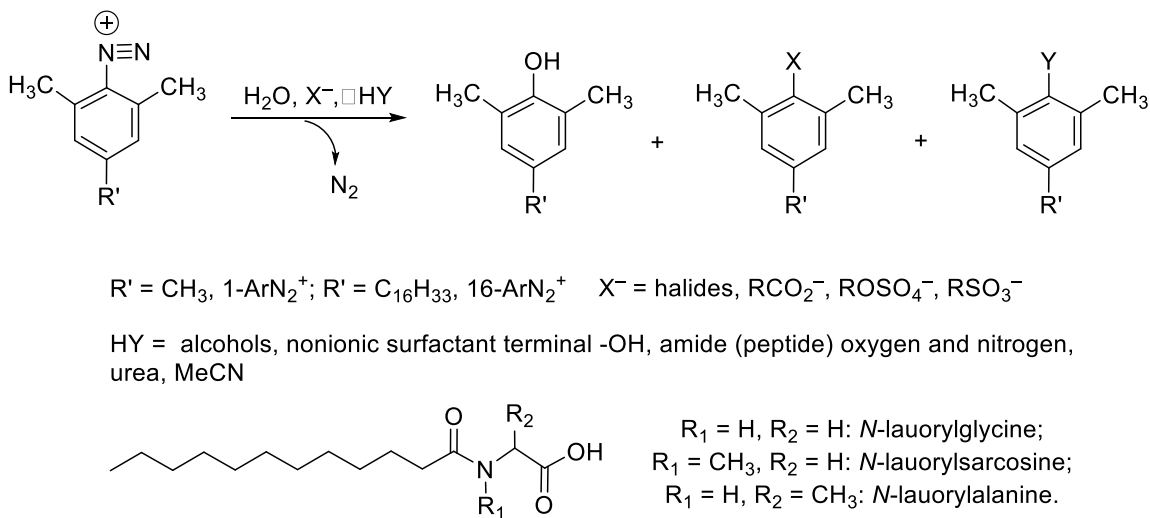
Chapter II. Simultaneous Determination of Local Concentrations of Amide Bonds, Carboxylate Groups, and Water in Micelles of Amino Acid Amphiphiles Containing Peptide Bond Models.

In this chapter, we summarize a series of studies to demonstrate the potential of the chemical concept, i.e., the chemical imaging approach, in tagging and fragmenting amino acid side chain and amide bonds of amino acid-based amphiphiles. Amino acids are unique small molecular building blocks with variable structures, polarity, charge and lipophilicity, which provide many possibilities in the design and application of biomaterial at the nanoscale.^{1,2} As one important class of surfactants that is widely used in cosmetics and food industry,³ amino acid surfactants (amphiphiles) undergo spontaneous self-aggregation in aqueous solution above their cmcs. Their amino acid headgroups are oriented within the interfacial region of association colloids, which make them simplest models of peptides at biomembrane interfaces. Chapter II begins with introducing the rationale of the amino acid amphiphile study along with a series of possible reaction pathways taking place at the interfacial region of amino acid amphiphile-based self-assembled aggregates. After a detailed experimental section, we provide a series of detailed results that shed light on the location and local concentration of amino acid headgroups, as is summarized in discussion toward the end.

2.1. Introduction

Arenediazonium chemistry has been extensively investigated during the past decades.^{4,5} Previous research conducted in our group has demonstrated that arenediazonium⁴ cations of 4-hexadecyl-2,6-dimethylbenzenediazonium tetrafluoroborate (16-ArN_2^+), and its short chain analog, 2,4,6-trimethylbenzenediazonium tetrafluoroborate (1-ArN_2^+), are trapped competitively and quantitatively by a large variety of weakly basic nucleophiles in association colloids and aqueous solution, respectively, **Scheme 2.1**. This scheme also shows the structures of the three

amphiphiles, C₁₂Gly, C₁₂Ala and C₁₂Sar, the arenediazonium ions used in these experiments, and products formed from reaction with anionic or neutral weakly basic nucleophiles in micellar solutions.



Scheme 2.1. Top: reaction products formed from heterolytic dediazonation with short and long chain arenediazonium ions in competitive reactions between water, anions, X^- , and neutral, Y , nucleophiles in aqueous and micellar solutions. **Bottom:** the structures of *N*-lauorylglycine, C₁₂Gly, *N*-lauorylsarcosine, C₁₂Sar and *N*-lauorylalanine, C₁₂Ala.

The logic for estimating the interfacial molarities of weakly basic nucleophiles in C₁₂Gly, C₁₂Ala and C₁₂Sar micelles using the chemical imaging approach is based on the pseudophase model for chemical reactivity (see also *Chapter I*). **Figure 2.1** illustrates the basic approach. **Box A** shows a small cross section of a micellar interface containing the long chain probe, which works at the interfacial region, and nucleophiles X and H (water not shown). In micelles, X is the carboxylate group, Y is the amide carbonyl. **Box B** shows an aqueous reference solution containing the same reactive components in which the stoichiometric concentrations of the nucleophiles are determined by measurement. We assume that product formation in dediazonation reactions, is competitive. When the product yields determined by HPLC from

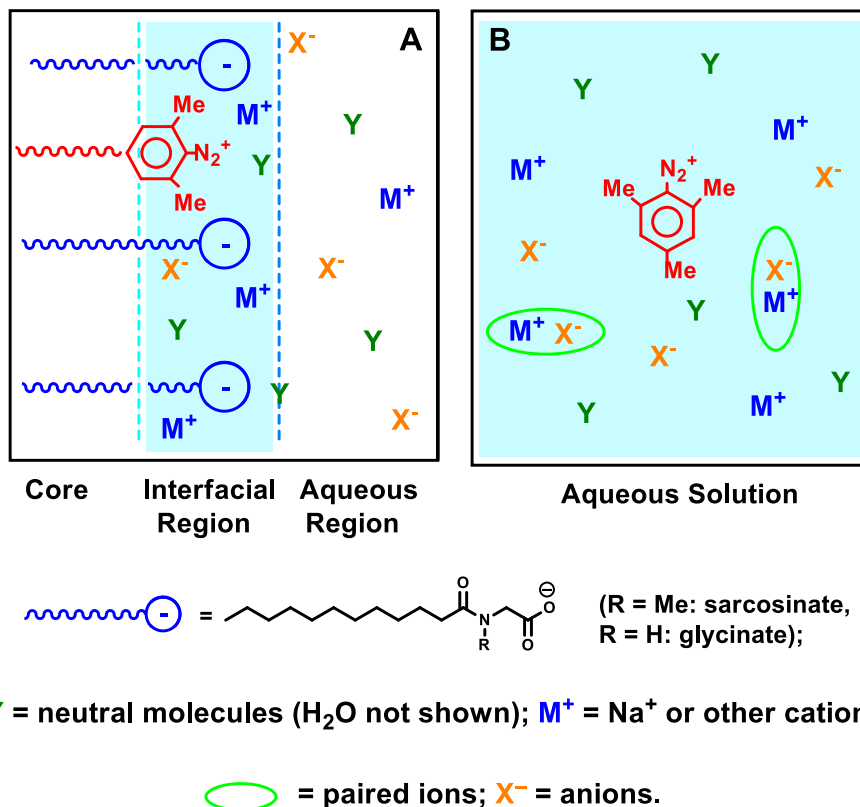


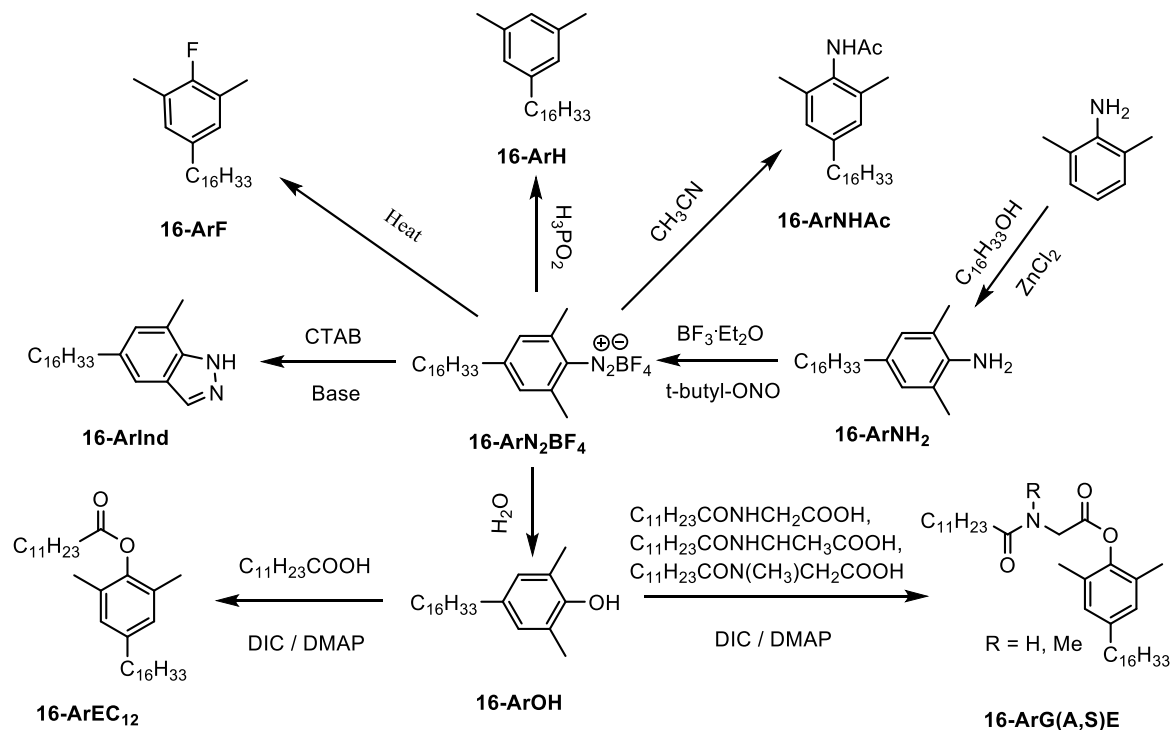
Figure 2.1. **Box A** illustrates a small section of the immediate vicinity of the interfacial region of an anionic micelle in aqueous solution in which 16- ArN_2^+ is trapped by weakly basic nucleophiles, H_2O (not shown), X^- , and **Y**. **Box B** shows an aqueous reference solution in which 1- ArN_2^+ is trapped by the same nucleophiles in the absence of micelles. Nucleophile molarities in the blue interfacial region (**Box A**) are assumed to be the same as those in the blue aqueous solution (**Box B**) when the product yields in the interfacial region and bulk aqueous solution are determined to be the same.

competitive reaction of 16- ArN_2^+ with **X**, **Y**, and H_2O in the micellar interface are the same as the yields from competitive reaction of 1- ArN_2^+ with the same (or structurally similar) nucleophiles in the reference aqueous solution, then the molarities of the functional groups in the micellar interface will be the same as their counterparts or themselves in the reference aqueous solution. We use this assumption to estimate the local concentrations of amide bonds and carboxylate sidechains of the three candidate molecules in their self-assembled aggregates.

2.2. Experimental Section

2.2.1. Materials.

All commercial reagents were purchased from Sigma-Aldrich and used as received, except for sodium *N*-lauroyl sarcosinate ($C_{12}\text{Sar}$), which was recrystallized from MeOH/Et₂O 3 times before use. Davisil silica gel from Sigma-Aldrich (grade 634, 100-200 mesh, 60Å / grade 10180, 70-230 mesh, 40Å) was used in column chromatography. HPLC grade solvents, *i*-PrOH, MeOH, and MeCN were filtered (Whatman Nuclepore Polycarbonate Track-Etched Membranes, pore size: 0.2 μm , diameter: 47 mm) before use to ensure their purity. HPLC grade ethyl acetate, hexanes were used as received. All solutions used for chemical reactions were



Scheme 2.2. Synthetic Routes for Preparations of 16-ArN₂BF₄ and of All Reaction Products Formed during Dediazonation Reactions in C₁₂Ala, C₁₂Sar and C₁₂Gly Micelles.³⁶⁻⁴³

prepared in water that was distilled, passed over a charcoal filter and a deionization resin, and redistilled using a Corning water purifier, LD-5 α . The preparations of 16-ArN₂⁺ and of all dediazonation products are summarized in **Scheme 2.2**, including references to published results. Three products, 16-ArNHAc, 16-ArF and 16-ArBr, were prepared previously in our laboratory

and used directly. Synthetic routes for some of the compounds, 16-ArNH₂, 16-ArN₂BF₄, 16-ArOH, 16-ArNHAc, 16-ArF and 16-ArInd, are published. Only new preparations of reaction products and also of sodium *N*-lauroyl glycinate (C₁₂Gly) and sodium *N*-lauroylalaninate (C₁₂Ala) are described. The details on the syntheses of new compounds are listed in the *Appendix* Section S1 at the end of this chapter.

2.2.2. Methods.

¹H-NMR spectra of amino acid amphiphiles and dediazonation products were recorded on Varian VNMRs 300 MHz or 400 MHz spectrometers, using either CD₃OD, CDCl₃ or D₂O as solvents. High resolution mass spectra were obtained on a ThermoQuest Finnigan LCQ-DUO mass spectrometer. Surface tensions of C₁₂Sar and C₁₂Gly solutions were determined using a *du Noüy* ring tensiometer, a Fisher Surface Tensiomat (Model 21). Kinetic (UV) measurements were performed on a Perkin-Elmer Lambda 45 UV/VIS spectrophotometer equipped with a Peltier Temperature Programmer 6 operated with UV WinLab 6.0.3 software. HPLC measurements were performed on a Perkin-Elmer Series 200 controlled by TotalChrom Navigator 6.2.1 software. Separations were carried out using Varian Microsorb MV C₁₈ columns, 250 mm length, 5 μm particle size, 100Å pore size, with *i*-PrOH/MeOH mixtures as eluents. All pH values were measured on a two-buffer standardized Fisher Accumant pH meter.

2.2.3. Calibration Curves for Dediazonation Reaction Products Formed in C₁₂Ala, C₁₂Sar and C₁₂Gly Micelles.

Calibration curves were created for converting measured peak areas of reaction products into percent yields. A MeOH stock solution of known concentration of each compound was prepared and diluted serially to give a range of concentrations of each compound that spanned the concentration ranges in the dediazonation experiments, e.g., ca. 1×10⁻⁴ M. The eluting solvent was 45%/55% v/v *i*-PrOH/MeOH (except for 16-ArInd for which the solvent ratio was 40%/60%). Injection volumes were 100 μL (except for 16-ArInd for which the injection volume was 50 μL)

and the detector was set at $\lambda = 220$ nm. Each peak area was measured in triplicate and an average value was taken with the variation within $\sim 5\%$. Each calibration curve was obtained from a plot of the average peak area of each compound against its concentration. Values for the slopes and intercepts were obtained from least squares fits. The plots are shown in *Appendix Figures 2S2-A-2S2-I*. The correlation coefficients are excellent (0.9945 to 1.0000) and the slopes used to calculate product yields are numerically close, probably because the primary chromophore in the reaction products is the aromatic ring contributed by the arenediazonium ion.

2.2.4. Critical Micelle Concentrations of C₁₂Sar & C₁₂Gly.

The cmc values of C₁₂Sar and C₁₂Gly were measured by surface tension at room temperature to determine the minimum amphiphile concentrations required for micelle formation and to check for the presence of hydrophobic impurities. Stock solutions of C₁₂Sar and C₁₂Gly were prepared at concentrations that were about twice the cmc. A 10 mL aliquot of the stock solution was transferred into a carefully cleaned, glass, crystallizing dish (rinsed serially by HPLC grade hexanes, MeOH, and distilled water and shaken dry) by using a volumetric glass pipette. The platinum ring was rinsed serially by HPLC grade hexanes, MeOH, and distilled water and dried in the air and passed through a flame to burn off impurities. The methods for calibrating the tensiometer and carrying out surface tension measurements are in the online Tensiomat 21 manual. The surface tension was measured repeatedly at each amphiphile concentration until the variation of three consecutive measurements was $\leq 0.1 \text{ mN}\cdot\text{m}^{-1}$. The amphiphile solution was diluted with an aliquot of water, typically 2 - 4 mL, with the measurement procedure repeated. The pH of the solution in the dish was measured with a calibrated pH meter before the first surface tension measurement and after the last. Plots of surface tension versus $\ln [\text{amphiphile}]$ were used to obtain the cmc from the intersection plots above and below the cmc (see *results*).

2.2.5. Dediazonation Kinetics in C₁₂Sar and C₁₂Gly Micelles.

The observed rate constant, k_{obs} , for dediazonation of 16-ArN₂⁺ in solutions of 0.069 and 0.098 M (5 - 6 times the cmc) C₁₂Sar, and of C₁₂Gly, were determined by UV spectrometry at 40°C. Values of k_{obs} were calculated from the slopes of the integrated rate equation for first order reactions and the values of k_{obs} are average values of duplicated runs. The pH of the amphiphile solutions were adjusted to 6-7 by adding aliquots dropwise of 2 M HCl. An aliquot of 10⁻² M cold and freshly prepared stock solution of 16-ArN₂⁺ (in HPLC grade MeCN) was added to the amino acid amphiphile solution in a glass cuvette giving a final concentration of 2×10⁻⁴ M 16-ArN₂⁺. The disappearance of 16-ArN₂⁺ was monitored at $\lambda = 285.5$ nm over 7 hours. Each kinetics experiment was run in duplicate.

2.2.6. Chemical Imaging in C₁₂Ala, C₁₂Sar & C₁₂Gly Micelles.

Solutions of C₁₂Ala, C₁₂Sar and C₁₂Gly at several different concentrations were prepared with the pH of each solution adjusted 5-7 by careful addition of small aliquots of HCl (see *Dediazonation Kinetics*). An aliquot of 1×10⁻² M 16-ArN₂⁺ (in cold MeCN) was added to initiate reaction (final concentration: 2×10⁻⁴ M) and mixed thoroughly, placed in a 40°C constant temperature bath for 12 hours (>10 half-lives, half-life: 44.7 min [see *Results*]). Chromatograms of 100 μ L of each solution were obtained in triplicate by HPLC at $\lambda = 220$ nm, eluting solvent: 35%/65% v/v, *i*-PrOH/MeOH, flow rate: 0.40 mL/min.

2.3. Results

Several pieces of experimental information must be gathered to determine the local concentrations (interfacial molarities) of the functional groups on amino acid headgroups in self-assembled micelles. (a) Ensure the amphiphile's purity including measuring its cmc. (b) Demonstrate that the mechanism of dediazonation is the same in C₁₂Ala, C₁₂Sar and C₁₂Gly micelles as in other amphiphile micelles.⁶ (c) Identify the products from reaction with each nucleophile, which requires independent synthesis of each product, and authenticate each reaction

product peak in the HPLC chromatograms, typically by a spiking experiment. (d) Prepare a calibration curve for each product to convert HPLC peak areas into product yields. (e) Determine the selectivity of each nucleophilic functional group toward 1-ArN_2^+ relative to water in aqueous solution in the absence of micelles to convert product yields from dediazonation of 16-ArN_2^+ into interfacial molarity of each nucleophile.

2.3.1. Cmc's of C_{12}Sar and C_{12}Gly .

The cmc value of C_{12}Gly at ambient temperature is 12.2 mM (**Figure 2.2**, no buffer added). The initial and final pH values of the solution are 8.2 and 7.0, respectively. The pH probably decreases because the terminal carboxylate group of C_{12}Gly is a weak acid (see *discussion*). No minimum was found, indicating that C_{12}Gly is free of surface active

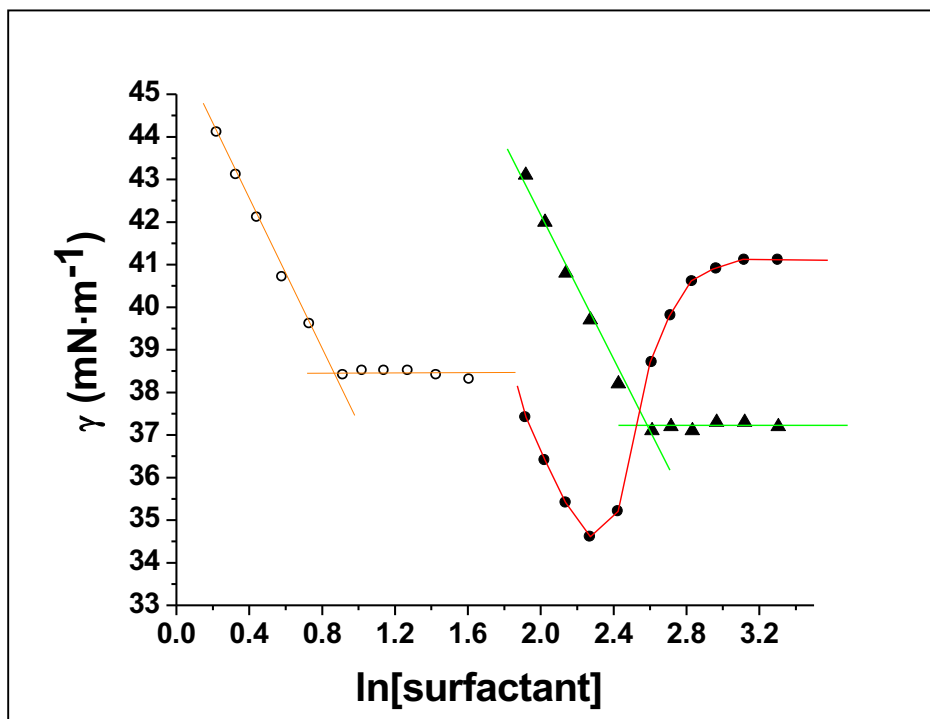


Figure 2.2. Surface Tension versus the Natural Logarithm of C_{12}Sar and C_{12}Gly Concentrations (mM) at Ambient Temperature, ca. 23°C : \blacktriangle C_{12}Gly at ambient pH; \bullet C_{12}Sar at ambient pH; and \circ , C_{12}Sar in aqueous $\text{NaHCO}_3/\text{NaOH}$ buffer (pH=11.0).

impurities and the cmc value is consistent with those in the literature.¹⁴ The surface tension results for C_{12}Sar , however, are dramatically different, **Figure 2.2**. The initial and final pH values

in the absence of buffer are 7.6 and 6.0, respectively, and a large minimum is present with the lowest point at about 9.5 mM. The minimum did not disappear on carrying out a new surface tension experiment after recrystallizing the C₁₂Sar 3 times. (see *Appendix* Section 2S1-I). The minima might indicate formation of a surface active pH dependent agent such as protonated, neutral form of C₁₂Sar (HLS). Consistent with this possibility, the surface tension plot of C₁₂Sar in carbonate buffer at pH 11.0 (0.1 M NaHCO₃ titrated with 1 M NaOH to pH 11.0) is without a minimum, **Figure 2.2**. If the minimum at pH 6.0 was caused by a hydrophobic impurity, it should probably be observed at pH 11.0. However, if the minimum is caused by a complex between the protonated and deprotonated forms of C₁₂Sar, then it should disappear because C₁₂Sar should be completely deprotonated at pH 11.0. A detailed interpretation is provided in *Chapter IV*.

2.3.2. Determination of k_{obs} for Dediazonation in C₁₂Sar and C₁₂Gly Micelles.

Values for k_{obs} and half-lives for dediazonation were determined in both aqueous C₁₂Sar and C₁₂Gly micelles at two different amphiphile concentrations to ensure that heterolytic dediazonation mechanism was not changed with the amphiphile concentration. The pH values of

Table 2.1. Observed Rate Constant, k_{obs} , and Half-life, $t_{1/2}$, for Dediazonation of 16-ArN₂⁺ (2×10^{-4} M) in 0.098 and 0.069 M C₁₂Sar and C₁₂Gly (pH = 6.8, at 40°C).

Amphiphile (M)	$10^4 k_{\text{obs}}$ (s ⁻¹) ^a	R ²	$t_{1/2}$ (min) ^{a,b}
C ₁₂ Sar (0.098)	2.72	1.0000	42.5
C ₁₂ Gly (0.098)	2.68	0.9981	43.0
C ₁₂ Sar (0.069)	2.62	0.9998	44.1
C ₁₂ Gly (0.069)	2.37	1.0000	48.9

a. Average value of $k_{\text{obs}} = (2.64 \pm 0.1) \times 10^{-4} \text{ s}^{-1}$, Average deviation: $\pm 3.8\%$. Average half-life: 44.7 min.

b. Number of $t_{1/2}$ followed: 3 – 5.

the C₁₂Sar and C₁₂Gly solutions at 2 different concentrations were adjusted, measured and confirmed to be constant. **Table 2.1** summarizes the results. Values of k_{obs} were obtained from the decrease in the absorbance of the arenediazonium ion at $40 \pm 0.1^\circ\text{C}$, $\lambda = 285.5 \text{ nm}$ (λ_{max} for 16-ArN₂⁺) with time and by plotting the data by using the integrated first order rate law. 40°C was selected to minimize amphiphile solubility problems that occur at room temperature. The procedure for obtaining k_{obs} and $t_{1/2}$ are in *Appendix Section S3*, including the kinetic data in **Figures 2S3-A-2S3-F**.

At 25°C , C₁₂Gly does not form a totally homogenous solution at either 0.098 M or 0.069 M, but contains a tiny amount of white precipitate. After warming the solution to 45°C , the precipitate dissolves slowly to give a homogenous solution. The precipitate reappears slowly on cooling to 40°C or below. However, because only miniscule amount of precipitate is formed at 40°C by the end of the kinetic experiment, the values of k_{obs} and $t_{1/2}$ in C₁₂Gly micelles should be accurate.

2.3.3. Chemical Imaging Experiments in C₁₂Ala, C₁₂Sar and C₁₂Gly.

Figures 2.3-2.5 and **Table 2.2** show the HPLC chromatogram results for the products formed from the chemical imaging reaction in C₁₂Ala, C₁₂Sar and C₁₂Gly solutions at 40°C . All the labeled peaks in the chromatograms were identified by spiking experiments using independently synthesized products, **Scheme 2.2**. Peak areas of the products were converted to percent yields by using the calibration curves listed in *Appendix Table 2S2-A*. The measured and normalized yields of each product in C₁₂Ala, C₁₂Sar and C₁₂Gly were determined and summarized in **Tables 2.3-2.9**. The products from the reaction of the terminal carboxylate groups of the three amphiphilies, 16-ArAE, 16-ArSE and 16-ArGE, were obtained directly from the chemical imaging reactions. They have similar normalized percent yields ranging from 2% to 4%. However, the yields of 16-ArEC₁₂, the products formed by trapping the amide oxygen of C₁₂Ala, C₁₂Sar and C₁₂Gly followed by hydrolysis, are significantly different. This difference is important

in the interpretation of the results (See *discussion*). The total measured percent yields from the chemical imaging in aqueous C₁₂Ala, C₁₂Sar & C₁₂Gly micelles account for at least 90+% and as high as 95% of the amount of added 16-ArN₂⁺.

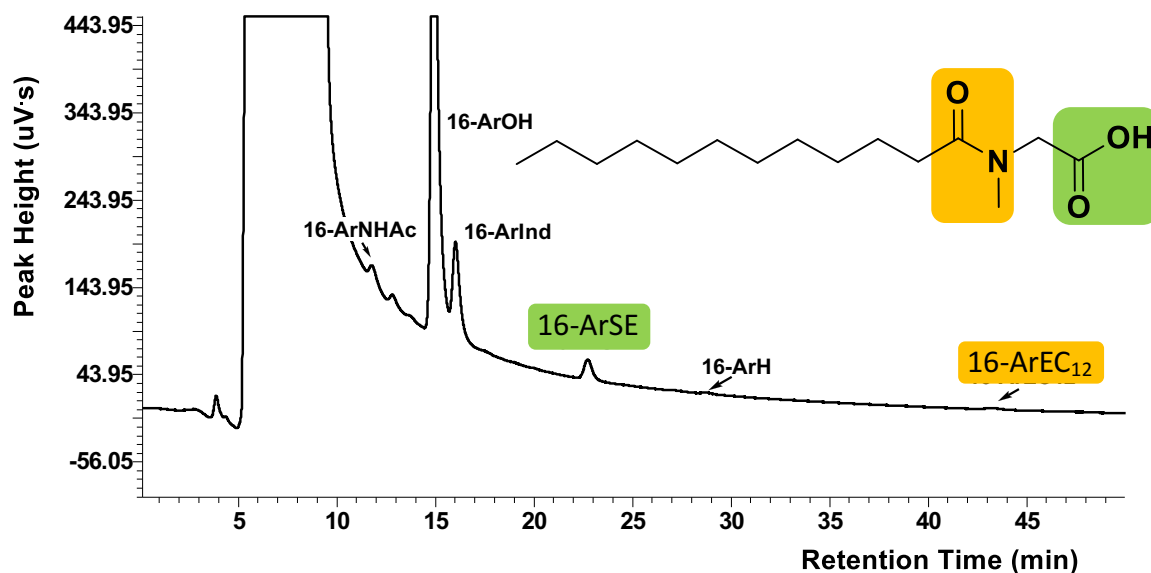


Figure 2.3. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.098M C₁₂Sar Aqueous Solution.

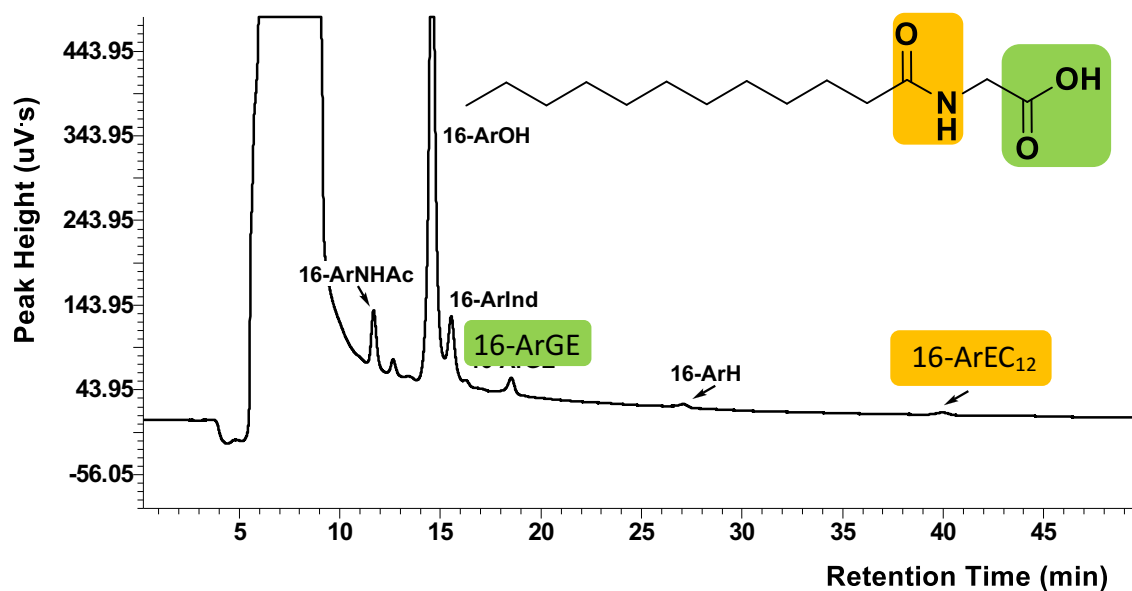


Figure 2.4. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.098M C₁₂Gly Aqueous Solution.

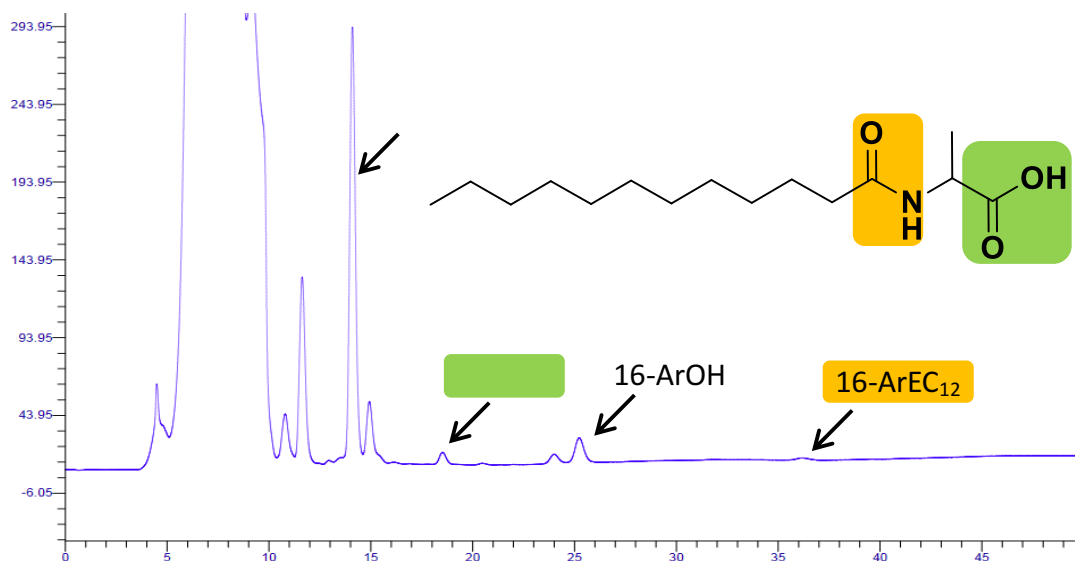


Figure 2.5. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.098M C₁₂Ala Aqueous Solution.

Table 2.2. The Retention Time Ranges for Each Product from Chemical Imaging Experiments.^a

Reaction Product	Retention Time ^b (min)
16-ArNHAc	12-13
16-ArOH	14-15
16-ArInd	15-16
16-ArGE	18-19
16-ArAE	19-20
16-ArSE	21-22
16-ArH	25-27
16-ArEC ₁₂	36-43

a. HPLC chromatograms of 100 μ L of samples were obtained in triplicate at $\lambda = 220$ nm, eluent 36%/64% v/v, *i*-PrOH/MeOH, flow rate: 0.40 mL/min.

b. Retention times obtained from chemical imaging result.

Table 2.3. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10^{-4} M 16-ArN₂⁺ in 0.098 M C₁₂Sar Aqueous Solution at $40 \pm 0.1^\circ\text{C}$ with a Reaction Time of 12 hours.

Reaction Product	10 ⁴ Peak Area (μV•s)	Observed Yields ^a (%)	Normalized Yields ^b (%)
16-ArNHAc	115.89	3.97	4.19
16-ArOH	1508.19	75.40	79.64
16-ArInd	358.47	12.77	13.49
16-ArSE	48.37	2.15	2.27
16-ArH	3.44	0.17	0.18
16-ArEC ₁₂	4.32	0.22	0.23
Total		94.68	100

a. Calibration curves are in **Table 2S2-A**.

b. Normalized Yields: %16-ArX^b = (Observed Yield %16-ArX^a / Total Observed Yield) × 100% (X = NHAc, OH, Ind, SE, H, EC₁₂).

Table 2.4. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10^{-4} M 16-ArN_2^+ with 0.069 M C_{12}Sar Aqueous Solution at $40 \pm 0.1^\circ\text{C}$ with a Reaction Time of 12 hours.

Reaction Product	10^4 Peak Area ($\mu\text{V}\cdot\text{s}$)	Observed Yields ^a (%)	Normalized Yields ^b (%)
16-ArNHAc	89.79	3.07	3.28
16-ArOH	1560.77	78.05	83.29
16-ArInd	270.49	9.64	10.29
16-ArSE	57.06	2.54	2.71
16-ArH	4.13	0.21	0.22
16-ArEC ₁₂	3.99	0.20	0.21
Total		93.71	100

a. Calibration curves are in **Table 2S2-A**.

b. Normalized Yield: $\%16\text{-ArX}^b = (\text{Observed Yield } \%16\text{-ArX}^a / \text{Total Observed Yield}) \times 100\%$ (X = NHAc, OH, Ind, SE, H, EC₁₂).

Table 2.5. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10^{-4} M 16-ArN₂⁺ with 0.098 M C₁₂Gly Aqueous Solution at $40 \pm 0.1^\circ\text{C}$ with a Reaction Time of 12 hours.

Reaction Product	10 ⁴ Peak Area ($\mu\text{V}\cdot\text{s}$)	Observed Yields ^a (%)	Normalized Yields ^b (%)
16-ArNHAc	128.79	4.41	4.88
16-ArOH	1498.70	74.95	82.85
16-ArInd	216.41	7.71	8.52
16-ArGE	49.66	2.28	2.52
16-ArH	6.21	0.31	0.34
16-ArEC ₁₂	15.93	0.80	0.88
Total		90.46	100

a. Calibration curves are in **Table 2S2-A**.

b. Normalized Yield: $\%16\text{-ArX}^b = (\%16\text{-ArX}^a / \text{Total Observed Yield}) \times 100\%$ (X = NHAc, OH, Ind, GE, H, EC₁₂).

Table 2.6. HPLC Peak Areas and Observed and Normalized Yields for the Reaction of 2×10^{-4} M 16-ArN_2^+ with 0.069 M C_{12}Gly Aqueous Solution at $40 \pm 0.1^\circ\text{C}$ with a Reaction Time of 12 hours.

Reaction Product	10^4 Peak Area ($\mu\text{V}\cdot\text{s}$)	Observed Yields ^a (%)	Normalized Yields ^b (%)
16-ArNHAc	101.20	3.46	3.76
16-ArOH	1599.06	80.00	86.96
16-ArInd	135.53	4.83	5.25
16-ArGE	51.62	2.37	2.58
16-ArH	12.58	0.64	0.70
16-ArEC ₁₂	13.71	0.69	0.75
Total		92.00	100

a. Calibration curves are in **Table 2S2-A**.

b. Normalized Yield: $\%16\text{-ArX}^b = (\%16\text{-ArX}^a / \text{Total Observed Yield}) \times 100\%$ (X = NHAc, OH, Ind, GE, H, EC₁₂).

Table 2.7. HPLC Normalized Yields for the Reaction of 8×10^{-4} M 16-ArN₂⁺ with 0.1 M C₁₂Ala Aqueous Suspension at pH 6 and $40 \pm 0.1^\circ\text{C}$ with a Reaction Time of 12 hours.

Reaction Product	Normalized Yields ^b (%)
16-ArOH	73.8
16-ArAE	2.63
16-ArEC₁₂	0.42

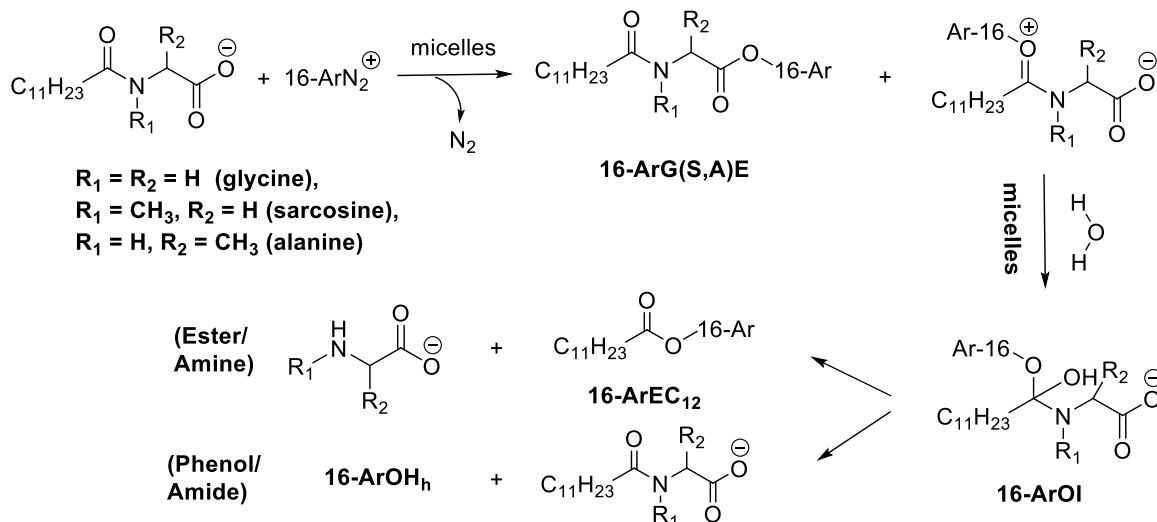
a. Calibration curves are in **Table 2S2-A**.

b. Normalized Yield: %16-ArX^b = (Observed Yield %16-ArX / Total Observed Yield) $\times 100\%$ (X = OH, AE, EC₁₂). Total observed yield = 96.2%.

Table 2.8. Normalized Yields of the Key Products for the Reaction of 2×10^{-4} M 16-ArN₂⁺ in Aqueous Solutions of 0.098 M C₁₂Sar and C₁₂Gly, and 0.100 M C₁₂Ala at $40 \pm 0.1^\circ\text{C}$ at pH 6-7 with Reaction Times of 12 hours.

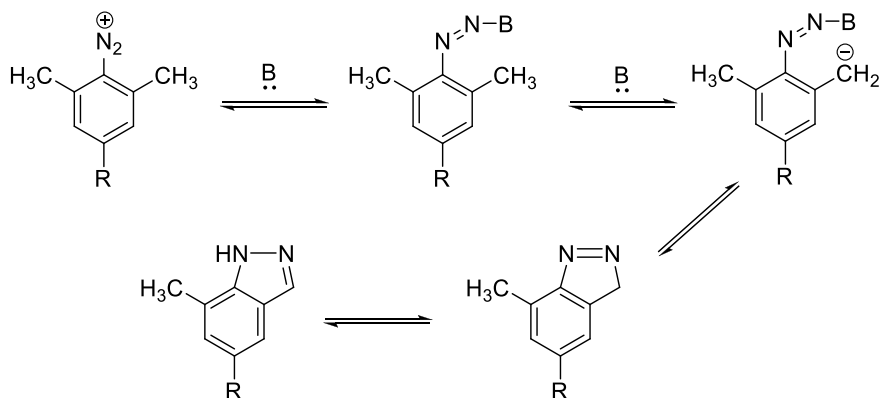
Amphiphile (M)	16-ArOH _w	16-ArS(G,A)E	16-ArEC ₁₂ + 16-ArOH _h	Total
C₁₂Sar (0.098)	96.37	2.76	0.87	100
C₁₂Gly (0.098)	92.84	2.92	4.24	100
C₁₂Ala (0.100)	93.77	3.42	1.71	100

Normalized Yield: %16-ArX = (%16-ArX / %16-ArOH_w + %16-ArAE, SE or GE + %16-ArEC₁₂ + OH_h) $\times 100\%$ (X = OH_w, AE, SE or GE, EC₁₂ + OH_h). Note that %16-ArOH_h = %16-ArOH - %16-ArOH_w.



Scheme 2.3. Tagging and Hydrolysis Pathways for the Chemical Imaging Reactions of 16-ArN_2^+ in C_{12}Sar , C_{12}Ala and C_{12}Gly Micelles. Parallel reactions occur in aqueous solutions with 1-ArN_2^+ in the absence of micelles. Competing reactions of 16-ArN_2^+ with H_2O yields 16-ArOH_w are not shown.

Scheme 2.3 summarizes the proposed mechanisms for the formation of the main products based on published results.^{7,10} Three important nucleophiles give significant product yields: interfacial water, the terminal carboxylate groups, and the amide oxygens of C_{12}Ala , C_{12}Sar and C_{12}Gly .¹⁰ 16-ArN_2^+ also reacts with amide nitrogens of primary and secondary, but not tertiary amides such as the amide nitrogen on C_{12}Sar .⁷ Note that the phenol product, 16-ArOH , has 2 different sources, i.e., trapping with water (16-ArOH_w) and with hydrolysis of 16-ArOI , the imido ester intermediates (16-ArOH_h). Also, the yield of 16-ArEC_{12} , one of the two products formed by trapping with amide oxygen and break down of 16-ArOI . The other ester product, 16-ArA(S,G)E , is formed directly by trapping with terminal carboxylate groups, **Scheme 2.3**. Note that the reduced product, 16-ArH , is formed by reaction of unreacted 16-ArN_2^+ with the phenol product, 16-ArOH ,¹³ but the yield is small fraction of the total 16-ArOH yield (**Tables 2.3-2.8**), showing that it is not an important pathway in these solutions and we ignore it. 16-ArF , the fluoro product formed by the Schiemann reaction of 16-ArN_2^+ with BF_4^- ,¹¹ was not observed.



Scheme 2.4. Proposed General Base, **B:**, Catalyzed Formation of 5-*n*-Hexadecyl-7-methyl-1H-indazole (16-ArInd). The General Base is Assumed to Be the Carboxylate Groups of C₁₂Sar, C₁₂Ala and C₁₂Gly.

The yield of imido ester product, 16-ArOI, formed by trapping with amide O, cannot be obtained directly from the yield of the ester product, 16-ArEC₁₂, because 16-ArOI is not stable and hydrolyses to both 16-ArEC₁₂ and 16-ArOH_h (**Scheme 2.3**). To obtain the yields of these two products, we used the same procedure as in the preparation of 1-ArEC₁₂ + 1-ArOH_h from acetamides.⁴⁴ In brief, the dediazonation reaction with 1-ArN₂⁺ was run in concentrated aqueous acetamide (*N*-methylacetamide, a model compound for the hydrolysis of the imido ester from C₁₂Gly, and *N,N*-dimethylacetamide, a model compound for the hydrolysis of the imido ester from C₁₂Sar, at 2.0 M and 4.0 M, respectively) in the presence and absence of 43.84% H₂¹⁸O. Note that the labeled 1-ArOH is obtained only during direct reaction of H₂¹⁸O with 1-ArN₂⁺, i.e., the hydrolysis of the imido ester intermediate gives only unlabeled 1-ArOH (**Scheme 2.3**). This difference was used to determine the yield of 1-ArOH from reaction of 1-ArN₂⁺ with water (1-ArOH_w) and the yield of 1-ArOH by hydrolysis (1-ArOH_h). The total yields of the imido ester intermediate in the reactions are: 1-ArOI = 1-ArOAc + 1-ArOH_h, where 1-ArOH_h is corrected for the amount of 1-ArOH produced during the reaction of 1-ArN₂⁺ with water (1-ArOH_w).

The yield of 16-ArOH_h was obtained by assuming that the 16-ArOH_h/16-ArEC₁₂ yield ratio in amino acid amphiphile micelles is the same as the 1-ArOH_h/16-ArEC₁₂ yield ratio, obtained

from aqueous acetamides. Then product yields were further modified by assuming that the total yield of the competitively formed products, 16-ArOH, 16-ArSE or 16-ArGE, and 16-ArEC₁₂ is 100%, where %16-ArOI = %16-ArEC₁₂ + %16-ArOH_h, and %16-ArOH = %16-ArOH_w + %16-ArOH_h (see **Table 2.7-2.9**).

2.3.4. Estimates of the Interfacial Molarities of Water and Amide and Carboxylate Groups in Amino Acid Amphiphile Micelles.

The selectivity of a competitive reaction toward any nucleophile, **X**, compared to water, H₂O, is given by **Equation 2.1**. The subscripts **w** and **m** stand for aqueous and micellar reaction media. The two important nucleophiles in the interfacial regions of amino acid amphiphile micelles are the oxygen of the amide carbonyl (X = -CON-)

$$S_w^X = \frac{[H_2O_x](\%1-ArX)}{[X_w][(\%1-ArOH)]} = \frac{[H_2O_m](\%16-ArX)}{[X_m][(\%16-ArOH)]} \quad (\text{Eq. 2.1})$$

and the terminal carboxylate group (X = CO₂⁻). The selectivities of these two groups were determined previously⁷ in concentrated aqueous solutions of water soluble model amide bonds, *N*-methylacetamide for C₁₂Gly and C₁₂Ala, and *N,N*-dimethylacetamide for C₁₂Sar, and the carboxylate group of glycine for the terminal carboxylate groups of all three amphiphiles. The results showed clearly that the selectivity of the reaction toward the amide carbonyl oxygen versus water depends on the degree of methylation of the N, where: S_w^O follows the order: acetamide (ca. 0.86) > *N*-methylacetamide (0.63) ≈ *N,N*-dimethylacetamide (0.63), **Appendix Table S4a**,⁷ and that of carboxylate group is about 1.00, **Table S4b**. Note that the selectivities are small and near one, indicating that there is little preference of the reaction toward the functional groups or water and the yields of the products are almost directly proportional to their concentrations.

The yield of %1-ArGE as a function of the concentration of glycine in the reaction of 1-ArN₂⁺ with glycine in water was determined previously (**Equation 2.2**).⁹ The slope and intercept

were used to obtain the molarity of carboxylate groups in the interfacial regions of the amino acid amphiphile micelles (**Equation 2.3**), from the product yields from reaction of 16-ArN₂⁺ with the terminal carboxylate groups:

$$\%1\text{-ArG(S, A)E} = 1.796 [\text{CO}_2^-]_w + 0.130 \quad (\text{Eq. 2.2})$$

$$\%16\text{-ArG(S, A)E} = 1.796 [\text{CO}_2^-]_m + 0.130 \quad (\text{Eq. 2.3})$$

By taking both **Equation 2.3** and the selectivity numbers into consideration, the interfacial molarities of the amide groups and the carboxylate groups are obtained and summarized in **Table 2.9**.

Table 2.9. Estimated Interfacial Local Concentrations (in molar) of the Nucleophiles from Chemical imaging in Aqueous 0.098 M C₁₂Sar and C₁₂Gly, and 0.100 M C₁₂Ala.

Amphiphile (M)	[COO ⁻] _m	[H ₂ O] _m	[CON] _m
C ₁₂ Sar (0.098)	1.46	50.98	0.73
C ₁₂ Gly (0.098)	1.55	49.27	3.57
C ₁₂ Ala (0.100)	1.83	50.23	1.45

2.4. Discussion

One unique feature of the chemical imaging method is that it provides experimentally measured estimates of the interfacial molarities of all the weakly basic nucleophiles present in the interfacial region of surfactant aggregates and also is extraordinarily insensitive being extraordinarily insensitive to properties of interfaces and reaction media.⁸ Indeed, chemical imaging lends a useful tool that is grounded in the use of heterolytic chemistry of arenediazonium ions, the application of pseudo-phase models and the assumption that when the yields of the products from competitive dediazonation reaction of 16-ArN₂⁺ with the nucleophiles in the interfacial region are the same as those of 1-ArN₂⁺ in the aqueous reference solution containing the same or similar nucleophiles, the concentrations of the nucleophiles located at the interfacial

region will be the same as in the reference solution at those molarities. Preliminary studies carried out with 1-ArN_2^+ in glycine, *N*-methylacetamide and *N,N*-dimethylacetamide aqueous solution, respectively,⁷ provides the following information:

- 1) A linear relationship between the percent yield of 1-ArGE, the ester product formed from trapping with carboxylate group of glycine, and the concentration of glycine in aqueous solution (**Equation 2.2**).
- 2) Selectivity of carboxylate group versus water determined from the chemical imaging of 1-ArN_2^+ in the same aqueous solution ($S_W^{\text{COO}^-} = 1.0$, see *Appendix S4*).
- 3) Selectivity of amide oxygen versus water determined from the chemical imaging of 1-ArN_2^+ in *N*-methylacetamide and *N,N*-methylacetamide aqueous solution ($S_W^{\text{O}} = 0.63$, see *Appendix S4*).

Results show that there is a small difference in interfacial concentrations of water and carboxylate groups in micelles (1.6 M on average, **Table 2.9**). However, the concentration of the amide bonds for C_{12}Sar micelles is 2 - 5 times less (ca. 0.7 M) than that of C_{12}Ala (~ 1.6 M) and C_{12}Gly micelles (~ 3.0 M). This is probably because the methyl group is more hydrophobic than a hydrogen and its presence in C_{12}Sar and C_{12}Ala leads to a higher fraction of the amide bonds located in hydrophobic region of the micelles, where 16-ArN_2^+ is less likely to be located, than the C_{12}Gly . The location of C_{12}Ala is probably between C_{12}Sar and C_{12}Gly .

The mechanism for the formation of the three key products has been elucidated explicitly. Other side reaction products include: 16-ArNHAc from reaction of 16-ArN_2^+ with MeCN, the solution for the arenediazonium ion stock solution,¹¹ and 16-ArInd formed by a general base-induced cyclization, **Scheme 2.4**.¹² Both pathways reduce the total product yields, but do not alter the relative yields from the competitive reactions of 16-ArN_2^+ with the amphiphiles. 16-ArH is formed from reaction of unreacted 16-ArN_2^+ with the phenol product, 16-ArOH .¹³ But the yield is

only a small fraction of the total 16-ArOH yield, and it is not an important pathway in these solutions. 16-ArF, the fluoro product formed by the Schiemann reaction of 16-ArN₂⁺ with BF₄⁻.¹¹

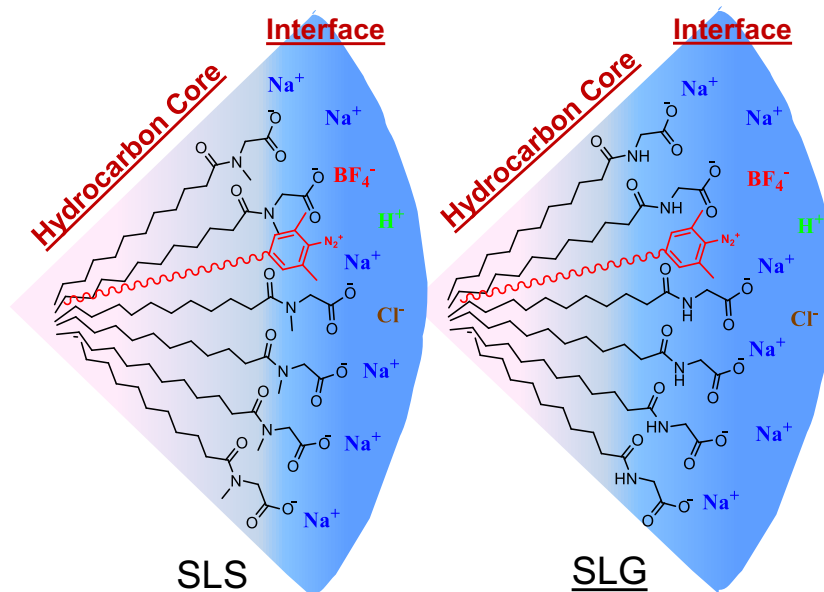


Figure 2.6. Cartoons of a Small Cross Section of C₁₂Sar and C₁₂Gly Micelles Containing the Chemical Imaging Probe 16-ArN₂⁺.¹⁰ The color transition from the aqueous interfacial region (blue) to hydrocarbon core (white) illustrates that water region hydrates more of the second amide group of C₁₂Gly than the tertiary amide group of C₁₂Sar. Bulk aqueous region not shown.

A number of uncertainties could pose challenges upon the interpretation of data obtained from chemical imaging. One case in particular is that whether the selectivities toward COO⁻ and COOH, respectively, versus water, would remain constant. At the ambient pH, protonated and deprotonated forms of the amphiphiles should coexist at the micellar interface, therefore, it may be impossible to apply the selectivity determined from glycine aqueous solution into the micellar environment (The interfacial pHs of the amino acid amphiphile micelles is unknown and hard to determine, but are estimated to be lower than the bulk pH, considering their anionic nature). Unpublished results demonstrate the selectivities toward acetic acid, CH₃COOH, and its deprotonated form, CH₃COO⁻, compared with water, are essentially the same,⁸ which validates the feasibility of data interpretation obtained from chemical imaging results.

A second question is that if the presence of the rotamers for C₁₂Sar amphiphile would affect the chemical imaging at any level. Results demonstrate that the estimated interfacial molarities of the key nucleophiles, COO⁻ (COOH) and CON obtained by conducting the chemical imaging in C₁₂Sar micellar solution at 2 difference concentrations, 0.098 M and 0.069 M, respectively, are in agreement in general (**Table 2.9**), which indicates that the difference between the rotamer ratios at 2 difference concentrations (0.098 M and 0.069 M) for C₁₂Sar micellar solutions will not pose a threat on the stability of product distribution and the comparison on the estimated interfacial molarities for each key nucleophile, respectively, in micellar solutions.

2.5. Summary

The potential of the chemical imaging approach to identifying differences in functional-group concentrations and locations within the interfacial region is illustrated by the significantly greater interfacial molarity of the secondary amide of C₁₂Gly and C₁₂Ala compared to the tertiary amide of C₁₂Sar. The estimated local concentrations in **Table 2.9** suggest that the micellar interfacial regions of amino acid amphiphile micelles are filled with water and that the concentrations of groups within the interfacial region are on the order of 1-3 M and that of water is on the order of 50 M. Chemical imaging could potentially aid in developing a better understanding of the interfacial properties of peptide amphiphiles for various applications.¹⁵⁻¹⁷ To date, no other chemical method has the potential to demonstrate the location and determine the local concentrations of chemical bonds in an interfacial region.

Bibliography

- (1) Ulijn, R. Peptide-Based Materials via Molecular Self-Assembly. In: Castillo, J, Sasso L, Svendsen, WE, editors. *Self-Assembled Peptide Nanostructures: Advances and Applications in Nanobiotechnology*. Pan Stanford Publishing; 2012. p. 67-91.
- (2) Altunbas, A.; Pochan, D. Peptide-Based and Polypeptide-Based Hydrogels for Drug Delivery and Tissue Engineering. In *Peptide-Based Materials: Topics in Current Chemistry (Book 310)*. Deming, T. volume editor. Springer; 2012.
- (3) Infante, M. R.; Perez, L.; Pinazo, A.; Clapes, P.; Moran, M. D. C. In *Novle Surfactants: Preparation, Applications, And Biodegradability*; 2nd ed.; Holmberg, K., Ed.; Taylor & Francis: New York, 2005, p 623.
- (4) *Diazo Chemistry I: Aromatic and Heteroaromatic Compounds*; Zollinger, H., Ed.; VCH Publishers: Weinheim, 1994.
- (5) *The Chemistry of the Diazonium and Diazo Groups, Part 2*; Hegarty, A. F., Ed.; John Wiley & Sons: New York, 1978.
- (6) Romsted, L. S. Do Amphiphile Aggregate Morphologies and Interfacial Compositions Depend Primarily on Interfacial Hydration and Ion Specific Interactions? The Evidence from Chemical Trapping. *Langmuir* **2007**, *23*, 414–424.
- (7) Romsted, L.; Zhang, J.; Zhuang, L. Mechanism of Reaction of an Arenediazonium Ion in Aqueous Solutions of Acetamide, *N*-Methylacetamide, and *N, N*-Dimethylacetamide. A Potential Method for Chemically Tagging Peptide Bonds at Aggregate Interfaces. *J Am Chem Soc* **1998**, *120*, 10046-10054.
- (8) Romsted, L. S. Interfacial Compositions of Surfactant Assemblies by Chemical Trapping with Arenediazonium Ions: Method and Applications. In *Reactions and Synthesis in Surfactant Systems*; Texter, J., Ed.; Marcel Dekker: New York, 2001; Vol. 10, p 265.
- (9) Zhuang, L. Mechanism of Reaction of Arenediazonium Ions with the Amide Bond. Ph.D. Thesis. Rutgers, The State University of Jersey, 1998.
- (10) Zhang, Y.; Romsted, L. S.; Zhuang, L.; de Jong, S. Simultaneous Determination of Interfacial Molarities of Amide Bonds, Carboxylate Groups, and Water by Chemical Trapping in Micelles of Amphiphiles Containing Peptide Bond Models. *Langmuir* **2012**, *29*, 534–544.
- (11) Romsted, L. S.; Yao, J. Arenediazonium salts: New Probes of the Interfacial Compositions of Association Colloids. 4. Estimating Hydration Numbers of Aqueous Hexaethylene Glycol Mono Dodecyl Ether, C₁₂E₆, Micelles by Chemical Trapping. *Langmuir* **1996**, *12*, 2425–2432.
- (12) Banerjee, R.; Das, P. K.; Chaudhuri, A. Interfacial Indazolization: Novel Chemical Evidence for Remarkably High Exo-surface pH of Cationic Liposomes Used in Gene Transfection. *Biochim Biophys Acta* **1998**, *1373*, 299–308.
- (13) Geng, Y.; Romsted, L. S.; Menger, F. M. Specific Ion-pairing and Interfacial Hydration as Controlling Factors in Gemini Micelle Morphology. Chemical Trapping Studies. *J Am Chem Soc* **2006**, *128*, 492–501.
- (14) Miyagishi, S.; Ishibai, Y.; Asakawa, T.; Nishida, M. Critical Micelle Concentration in Mixtures of *N*-acyl Amino Acid Surfactants. *J Colloid Interf Sci* **1985**, *103*, 164-169.
- (15) Brito, R. O.; Silva, S. G.; Fernandes, R. M. F.; Marques, E. F.; Enrique-Borges, J.; do Vale, M. L. C. Enhanced Interfacial Properties of Novel Amino Acid-derived Surfactants: Effects of Headgroup Chemistry and of Alkyl Chain Length and Unsaturation. *Colloids Surf. B* **2011**, *86*, 65–70.
- (16) Menger, F. M.; Zhang, H. L. Peptoids, a Group of Amphiphilic Long-chain Triamides. *Langmuir* **2005**, *21*, 10428–10438.

- (17) Mohanty, A.; Dey, J. Enantioselectivity of Vesicle-Forming Chiral Surfactants in Capillary Electrophoresis. Role of the Surfactant Headgroup Structure. *J. Chromatogr. A* **2006**, *1128*, 259–266.

Appendix

Section S1. Synthesis and/or purification procedures for the preparation of arenediazonium ion, dediazonation products, amino acid amphiphiles and other related products, and their ^1H NMR spectra, respectively.

a. Synthetic routes for some of the products.

16-ArNH₂, 16-ArN₂⁺, 16-ArOH, 16-ArNHAc, 16-ArF, 16-ArBr and 16-ArInd, have been published. For detailed information, please refer to the literature cited in **Scheme 2.2**.

b. 1-*n*-Hexadecyl-3, 5-dimethylbenzene, 16-ArH.

Using a slightly modified method published by Kornblum, *et al.*^{1,2} 200 mg of 16-ArN₂⁺ was dissolved in 5.0 mL THF. The resulting solution was added dropwise to a solution of 951 mg of hypophosphorous acid (H₃PO₂, 50% w/w) dissolved in 5.0 mL distilled water in an ice bath. After 5 min the ice bath was removed and the mixture was stirred overnight. Excess THF was evaporated and 10 mL×3 hexane (HPLC grade) aliquots were added to extract the remaining reaction mixture. The combined aliquots of hexane solution were dried with Na₂SO₄, filtered and evaporated to give a white solid. The solid was further purified by silica gel column chromatography eluted by hexanes and 86 mg (54%) of white crystals were obtained. *M_w* (calc.): 330.6 g mol⁻¹. ^1H NMR (300 MHz, CDCl₃): δ ppm 0.86 (3H, t, *J* = 8.0 Hz), 1.28 (26H, br s), 1.57 (2H, br), 2.24 (6H, s), 2.55 (2H, t, *J* = 8.0 Hz), 6.80 (2H, s), 7.26 (1H, s, overlapped by CHCl₃ at 7.26). See ^1H NMR spectrum, **Figure 2S1-B**.

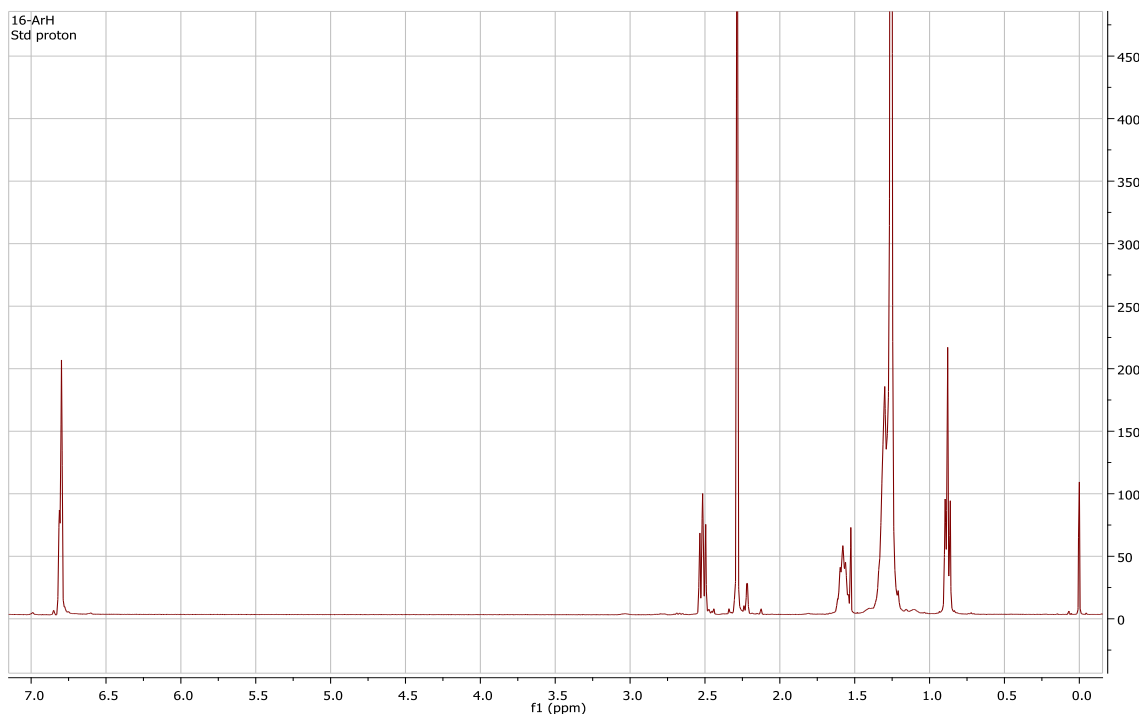


Figure 2S1-B. ^1H -NMR Spectrum for 16-ArH.

c. 4-*n*-Hexadecyl-2,6-dimethylphenyllaurate, 16-ArEC₁₂.

N, N'- Diisopropylcarbodiimide (DIC, 55.9 μL , 0.357 mmol, 1.5 eq.) and then DMAP (14.5 mg, 0.5 eq.) were added to a solution of 16-ArOH (82.6 mg, 0.238 mmol) and dodecanoic acid (95.5 mg, 2 eq.) in dichloromethane (5.0 mL). The mixture was stirred at r.t. overnight. The solid product was removed by filtration and the filtrate was evaporated to give a white solid. This solid was dissolved in EtOAc, which was washed successively with saturated Na_2CO_3 , NH_4Cl and NaCl solution dried over Na_2SO_4 and the EtOAc evaporated. The product was purified via column chromatography using 5% ethyl acetate/hexanes and recrystallized from methanol. 84 mg (67%) of pure compound was obtained. M_w (calc.): 528.49 g mol^{-1} . ^1H NMR (300 MHz, CDCl_3): δ ppm 0.87 (6H, t, $J = 6.5$ Hz), 1.36 (b, s, 42H), 1.54 (2H, m), 1.77 (2H, m), 2.11 (6H, s), 2.54 (2H, t, $J = 7.6$ Hz), 2.58 (2H, t, $J = 7.6$ Hz), 6.86 (s, 2H). **Figure 2S1-C.**

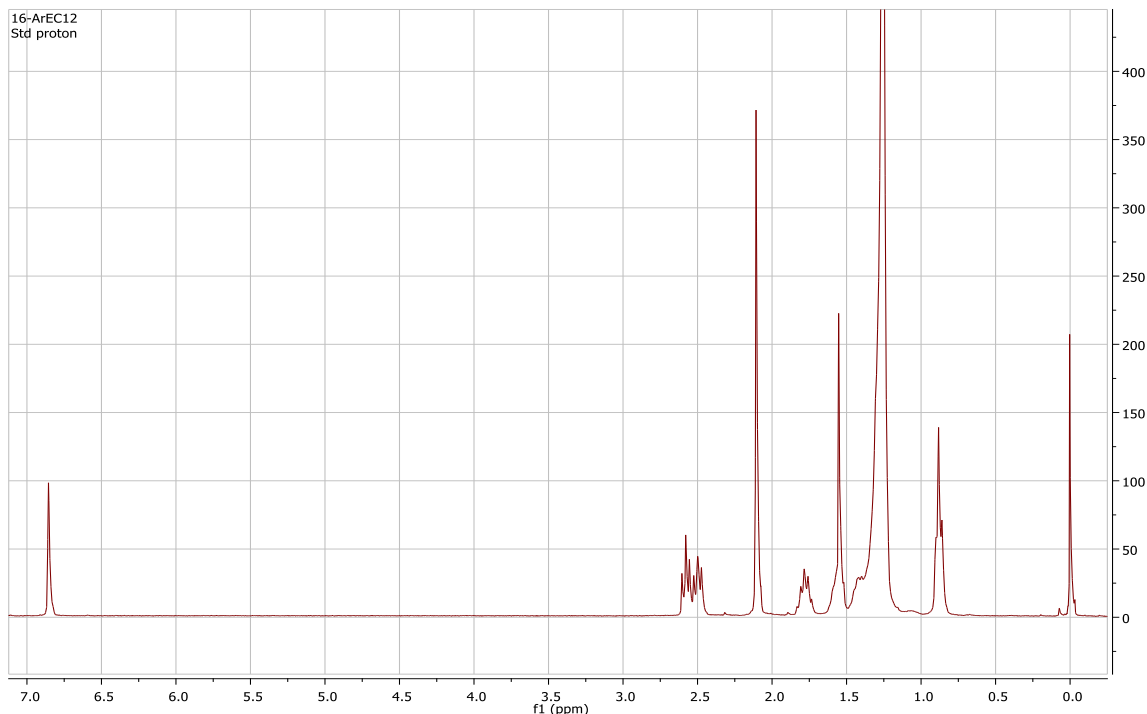


Figure 2S1-C. ^1H -NMR Spectrum for 16-ArEC₁₂.

d. 4-*n*-Hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl sarcosinate, 16-ArSE.

DIC (67.8 μL , 0.433 mmol) and then DMAP (17.62 mg, 0.5 eq.) were added to a solution of 16-ArOH (100 mg, 0.289 mmol) and *N*-lauroylsarcosine (156 mg, 2 eq.) in dichloromethane (5.0 mL). The mixture was stirred at r.t. overnight, the precipitate filtered out and the filtrate evaporated to give a white solid. The solid was dissolved in EtOAc, which was washed successively with saturated Na_2CO_3 , NH_4Cl and NaCl solution, and dried over Na_2SO_4 . Evaporation of the solvent yielded 120 mg (69%) of white solid that was purified by column chromatography using 15% ethyl acetate/hexanes, and recrystallization from methanol. M_w (calc.): 599.53 g mol^{-1} . ^1H NMR (300 MHz, CDCl_3): δ ppm 0.88 (6H, t, $J = 6.5$ Hz), 1.28 (b, s, 42H), 1.5-1.7 (4H, m), 2.09 (6H, s), 2.39 (2H, t, $J = 7.6$ Hz), 2.50 (2H, t, $J = 7.9$ Hz), 3.12 (s, 3H), 4.35 (s, 2H), 6.86 (s, 2H). HRMS (ESI) m/z (found): 600.22 g mol^{-1} ($[\text{M}+\text{H}]^+$). **Figure 2S1-D.**

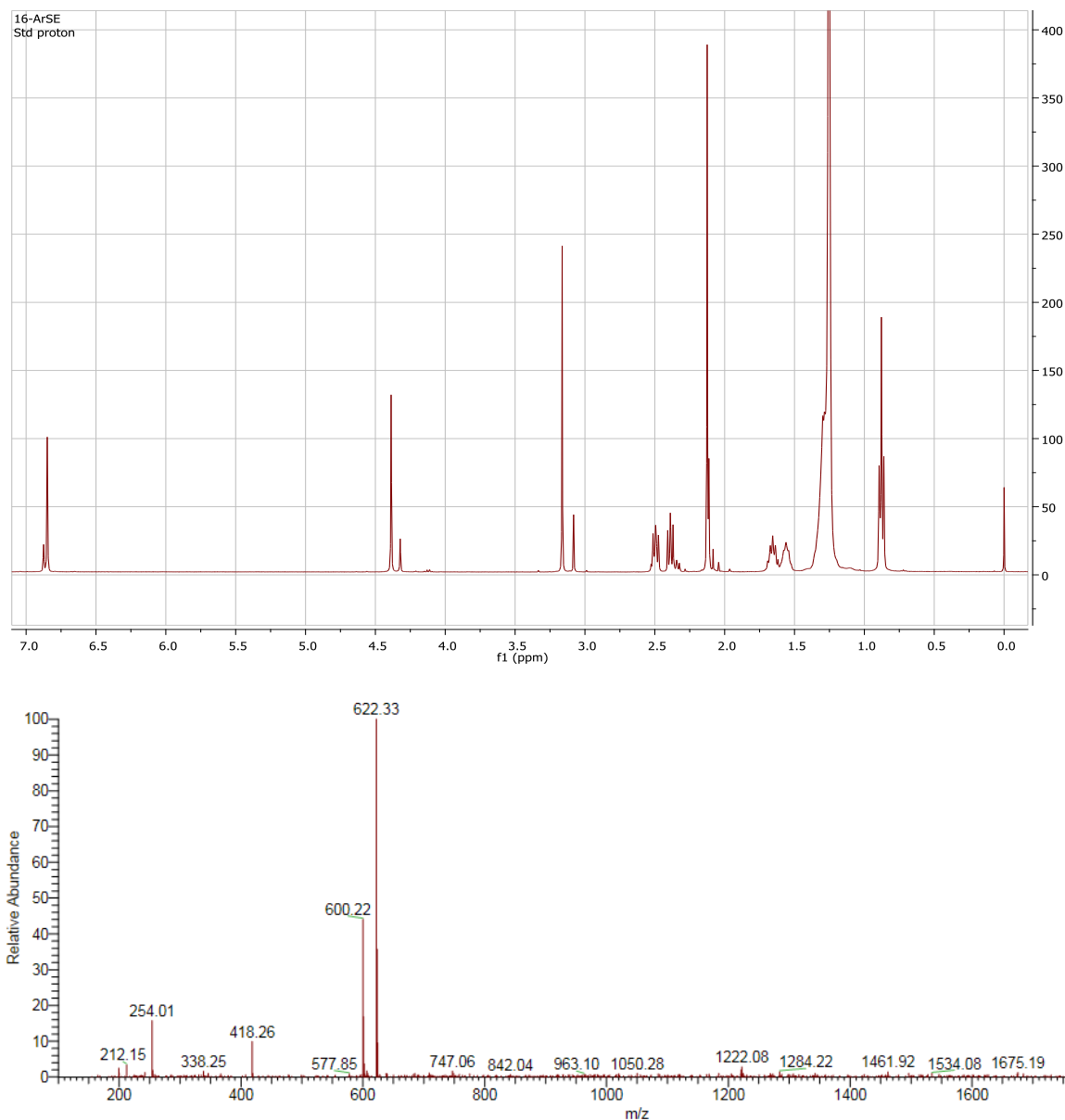


Figure 2S1-D. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for 16-ArSE.

e. 4-*n*-Hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl glycinate, 16-ArGE.

DIC (67.8 μL , 0.433 mmol) and DMAP (17.62 mg, 0.5 eq.) were added to a solution of 16-ArOH (100 mg, 0.289 mmol) and *N*-lauroylglycine (156 mg, 2 eq.) in dichloromethane (DCM, 5.0 mL). The mixture was stirred at r.t. overnight. Solid was removed by filtration and the filtrate evaporated to give a white solid. The solid was dissolved in EtOAc, which was washed successively with saturated Na_2CO_3 , NH_4Cl and NaCl solution, dried over Na_2SO_4 and the EtOAc

evaporated. 30 mg (17%) of white solid was obtained after column chromatography using 10% ethyl acetate/hexanes. M_w (calc.): $585.51 \text{ g mol}^{-1}$. ^1H NMR (300 MHz, CDCl_3): δ ppm 0.88 (6H, t, $J=6.7 \text{ Hz}$), 1.27 (b, s, 42H), 1.56 (2H, m), 1.64 (2H, m), 2.12 (6H, s), 2.28 (2H, t, $J=7.5 \text{ Hz}$), 2.50 (2H, t, $J=8.0 \text{ Hz}$), 4.35 (t, 2H, $J=5.1 \text{ Hz}$), 5.99 (1H, b), 6.87 (s, 2H). HRMS (ESI) m/z (found): $586.59 \text{ g mol}^{-1}$ ($[\text{M}+\text{H}]^+$). **Figure 2S1-E.**

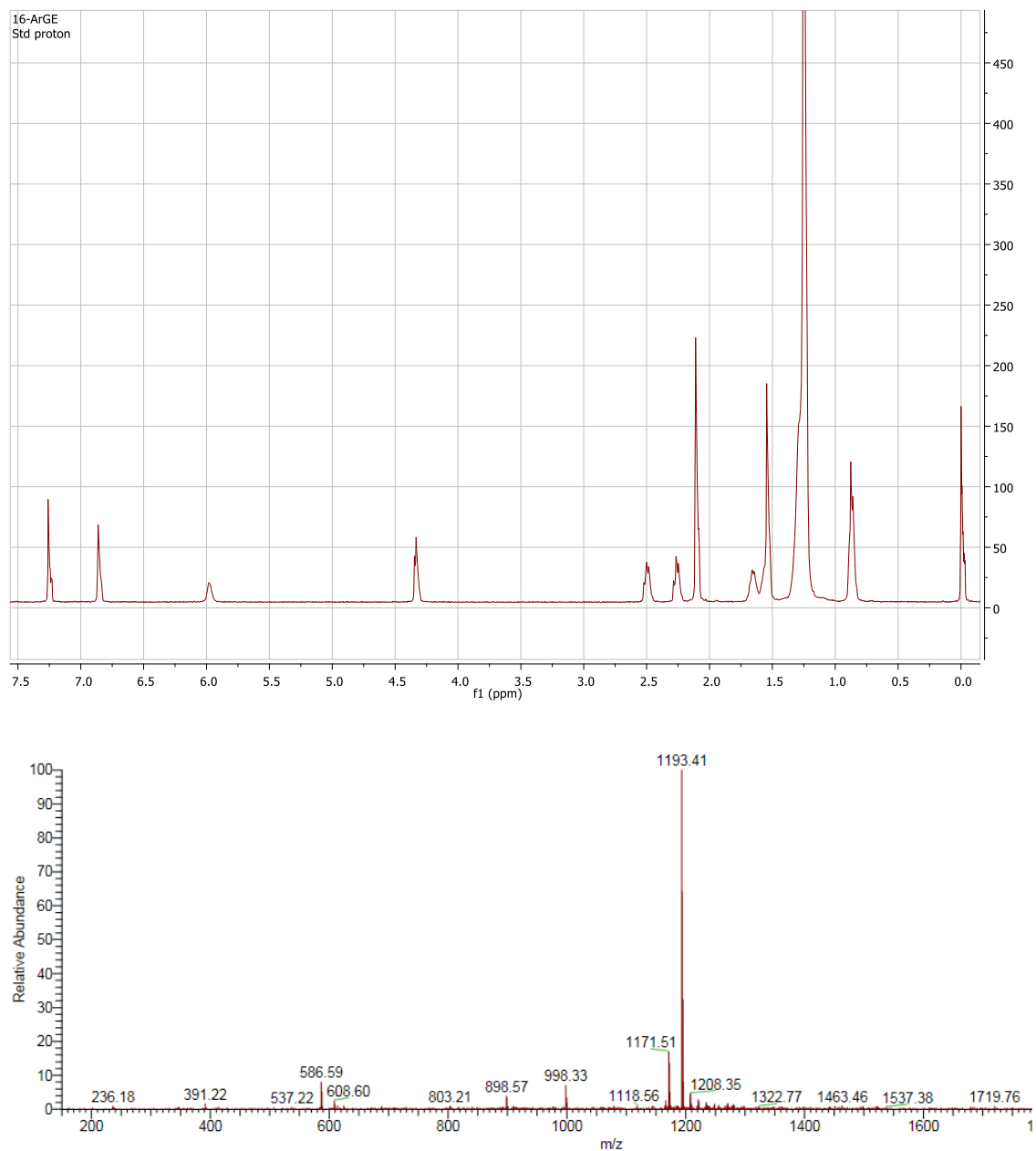
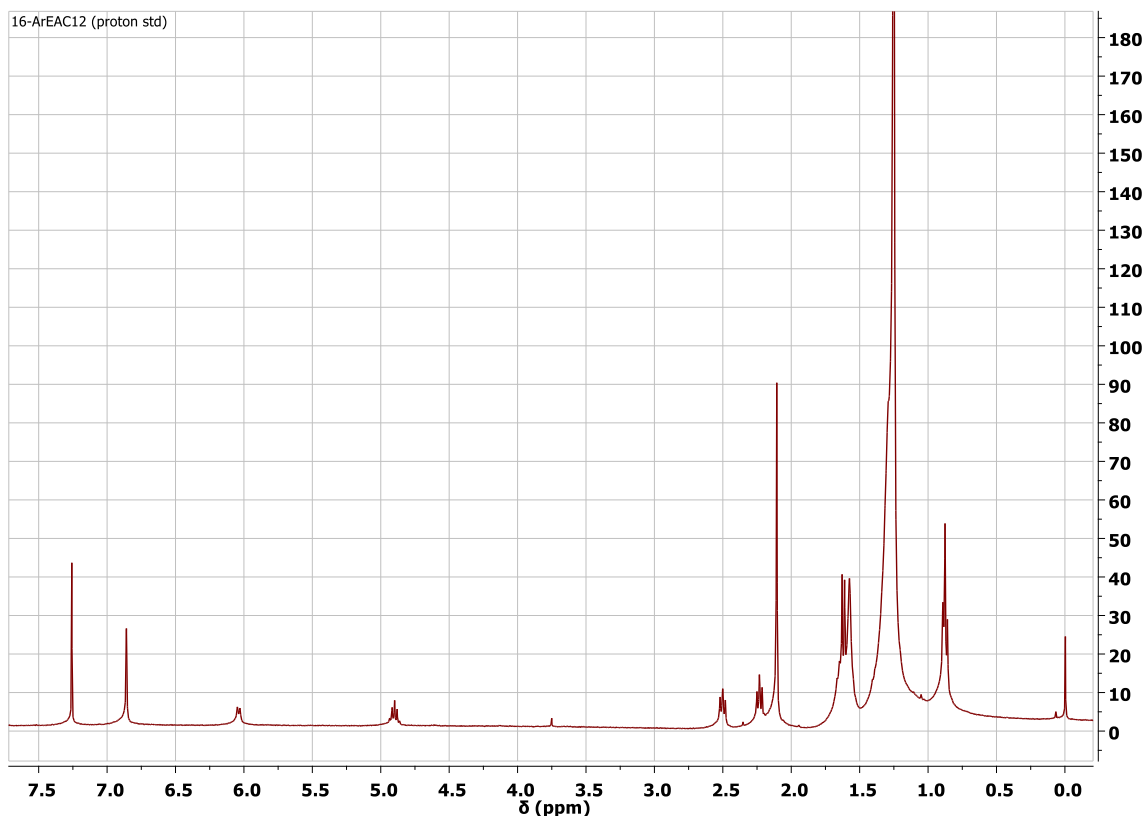


Figure 2S1-E. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for 16-ArGE.

f. 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylalaninate ester, 16-ArAE.

DIC (40.2 μ L, 1.5 eq.) and DMAP (10.6 mg, 0.5 eq.) were added to a solution of 16-ArOH (60 mg, 0.173 mmol) and *N*-lauroylalanine (94 mg, 2 eq.) in dichloromethane (DCM) (5.0 mL). The mixture was stirred at r.t. overnight, solid removed by filtration and the filtrate evaporated to give a white solid. The solid was dissolved in EtOAc, which was washed successively with saturated Na_2CO_3 , NH_4Cl and NaCl solution, dried over Na_2SO_4 and the EtOAc evaporated. 70 mg (67.4%) of white solid was obtained after column chromatography using 12% ethyl acetate/hexanes. M_w (calc.): 599.53 $\text{g}\cdot\text{mol}^{-1}$. ^1H NMR (300 MHz, CDCl_3): δ ppm 0.88 (6H, t), 1.27 (42H, bs), 1.54 (2H, m), 1.64 (2H, m), 2.12 (6H, s), 2.24 (2H, t), 2.51 (2H, t), 4.93 (2H, t), 6.06 (1H, b), 6.92 (2H, s). HRMS (ESI) m/z (found): 600.23 $\text{g}\cdot\text{mol}^{-1}$ ($[\text{M}+\text{H}]^+$). **Figure 2S1-F.**



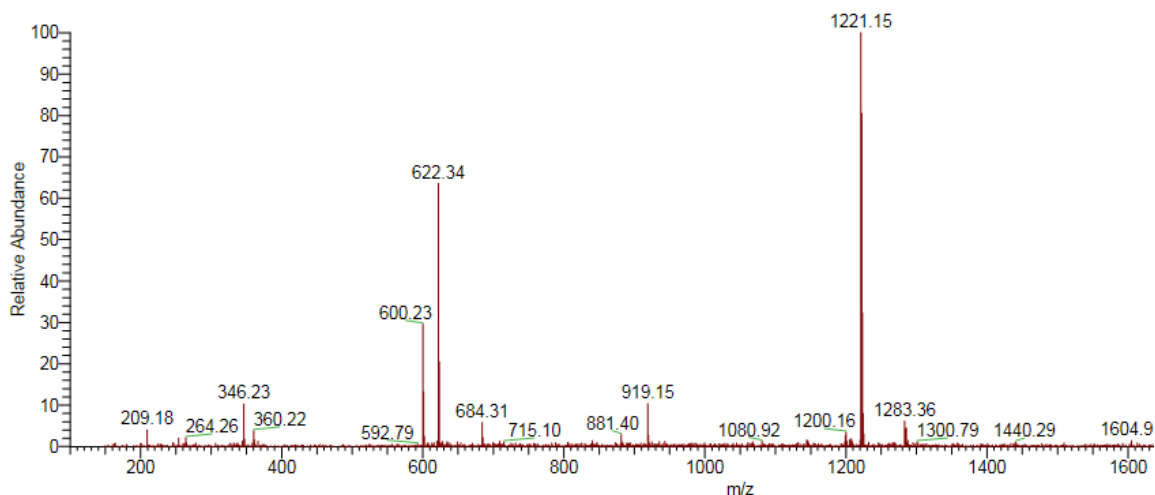


Figure 2S1-F. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for 16-ArAE.

g. Sodium *N*-lauroyl glycinate, NaLG.

Lauroyl chloride (5.00 mL, 21.19 mmol) was added to a suspension of methyl glycinate·HCl (2.93 g, 23.33 mmol) in DCM (70 mL). Dropwise addition of triethylamine (13.0 mL, 57.2 mmol) gave a thick suspension that was filtered after 30 min and the solid product was washed with DCM. The DCM filtrate was washed with 1.0 M HCl (30 mL) and brine, dried, and concentrated to afford a white solid. The solid was dissolved in DCM and precipitated with hexane, filtered and dried to give white lauroyl glycinate (LG-ester) (2.54 g, 44.2%). NaOH, (1 M, 18 mL) was mixed with a solution of LG-ester (2.51 g, 9.24 mmol) in ethanol (90 mL) and stirred at room temp overnight. The ethanol was removed by evaporation and the remaining solution was acidified with aqueous 1.0 M HCl to give white precipitate that was filtered, washed with water, and air dried. Recrystallization from MeCN gave white *N*-lauroyl glycinic acid (LG-acid) (2.1 g, 88.3%). A suspension of LG-acid (1.007 g, 3.91 mmol) in water (10 mL) was titrated with 1.0 M NaOH aqueous solution (3.91 mL) to form a clear solution. The solution was freeze-dried to afford a white solid. Recrystallization of the solid with MeOH gave white C_{12}Gly (0.85 g, 76.7%). M_w (calc.): 279.18 $\text{g}\cdot\text{mol}^{-1}$. ^1H NMR (300 MHz, D_2O): δ ppm 0.73 (3H, t, $J=7.0$ Hz), 1.15 (16H,

bs), 1.46 (2H, m), 2.17 (2H, t, $J=7.7$ Hz), 3.62 (2H, s). The signal for the proton on NH was not observed. HRMS (ESI) m/z (found): 256.38 $\text{g mol}^{-1} (\text{M}^-)$. **Figure 2S1-G.**

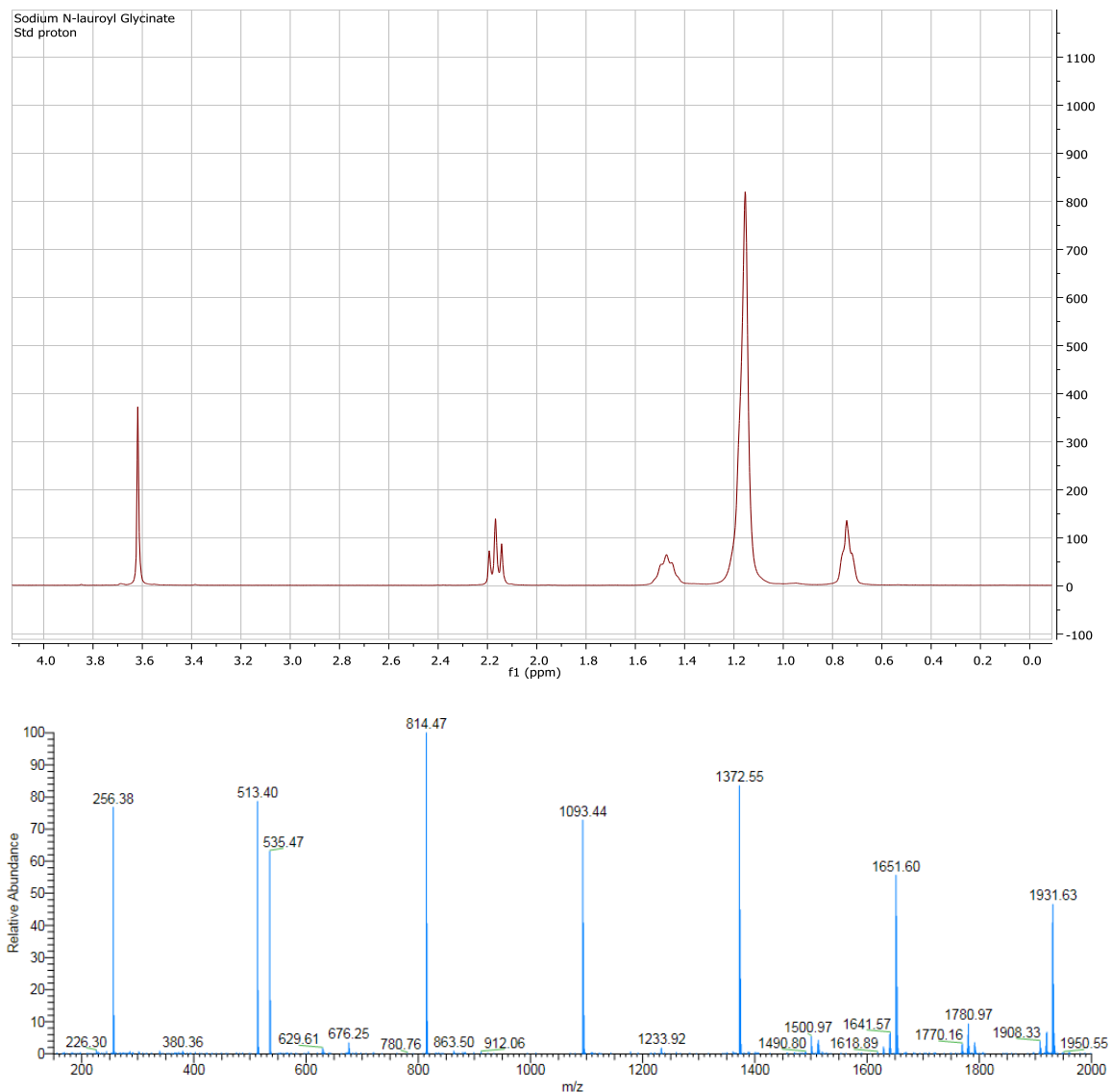


Figure 2S1-G. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for Sodium *N*-Lauroyl Glycinate (NaLG).

h. *N*-lauroyl-*L*-alanine, C_{12}Ala .

C_{12}Ala was synthesized by a resin-based solid phase peptide synthesis protocol. For details on synthesis, please refer to the “**Appendix**” in **Chapter III**. The crude product was dissolved in MeOH and an aliquot of 1 molar HCl was added to make the pH acidic at ca. 1 - 1.5. The solvent was evaporated and resulting solid was re-dissolved in ethyl acetate and washed with distilled

water twice. Ethyl acetate was evaporated and then the remaining product was rinsed with aliquots of hexanes three to five times to give white product. M_w (calc.): 271.21 g mol^{-1} . ^1H NMR (300 MHz, CDCl_3): δ ppm 0.88 (3H, t), 1.26 (16H, bs), 1.44 (2H, m), 1.63 (2H, t), 2.24 (2H, t), 4.58 (1H, m), 6.14 (2H, d). HRMS (ESI) m/z (found): 270.28 g mol^{-1} (M^-). **Figure 2S1-H.**

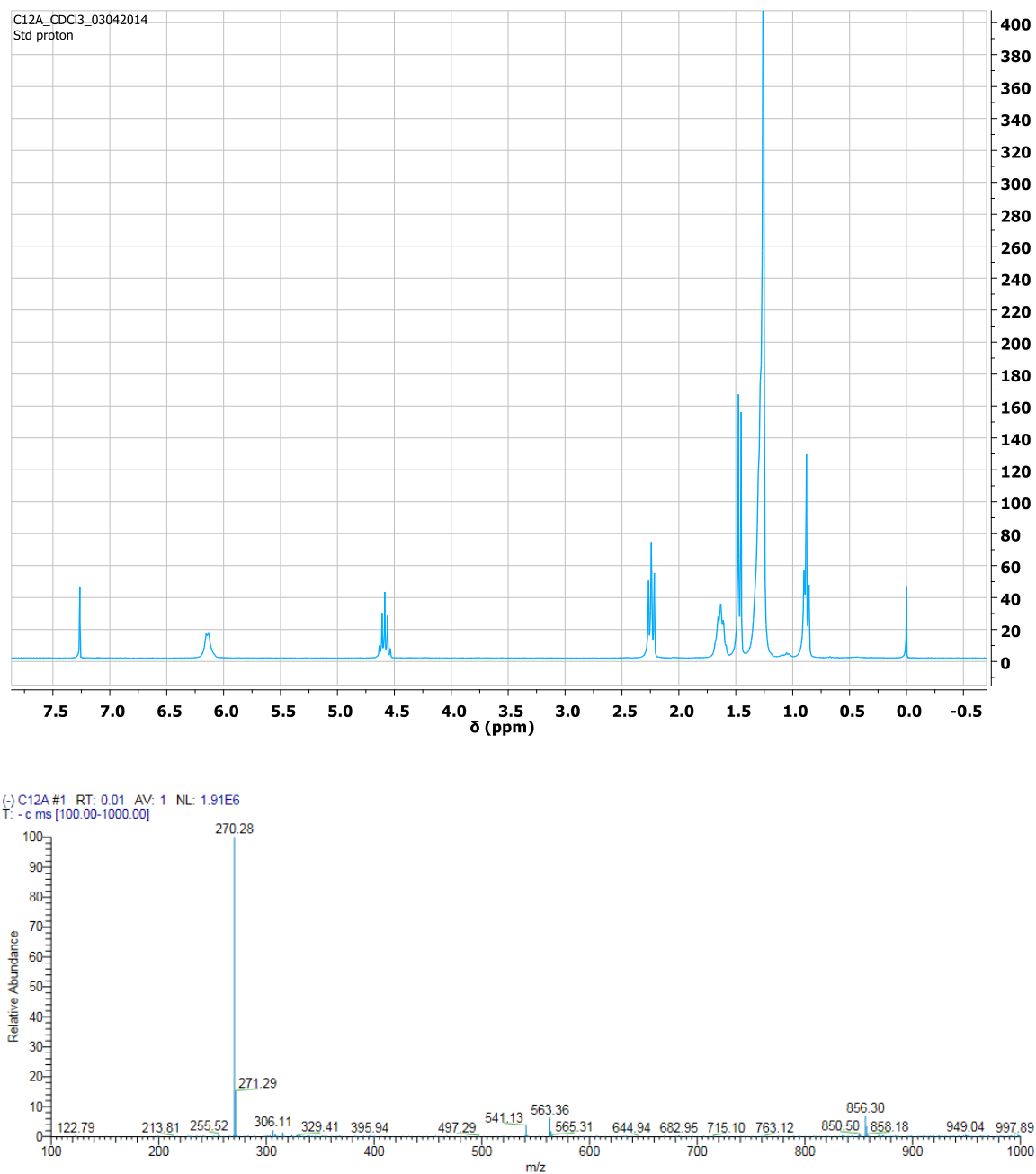


Figure 2S1-H. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for *N*-lauroylalanine (C_{12}A).

i. Purification of NaLS, its ^1H -NMR and its cmc determination in buffer.

Sodium *N*-lauroyl sarcosinate (Aldrich) was dissolved in HPLC grade hot MeOH. Undissolved solid was removed by filtration on a Büchner funnel, the filtrate cooled to room temperature, and then placed into an ice bath for 15 minutes. The white crystals were collected on a Büchner funnel, washed with small amounts of cold Et₂O and dried under oven. This process was repeated three times. M_w (calc.): 271.21 g mol⁻¹. ^1H NMR (300 MHz, D₂O): δ ppm 0.70 (3H, t, $J = 7.0$ Hz), 1.16 (16H, bs), 1.45 (2H, m), 2.14 and 2.29 (2 sets, 2H, t, $J = 7.7$ Hz), 2.77 and 2.93 (2 sets, 2H, s), 3.80 (2 sets, overlapped, 2H, s). HRMS (ESI) m/z (found): 270.28 g mol⁻¹ (M⁻). **Figure 2S1-G-1.**

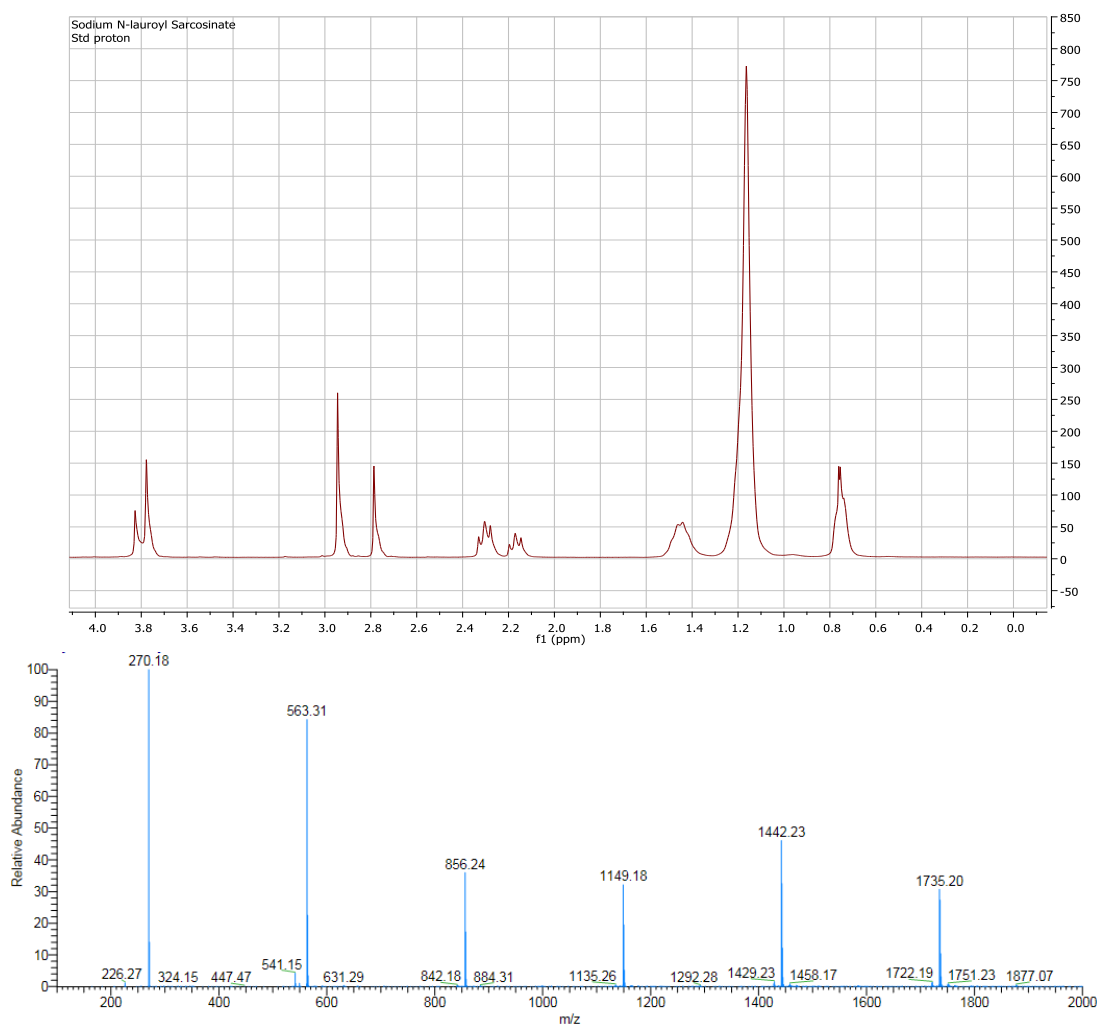


Figure 2S1-I. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for Sodium *N*-lauroyl Sarcosinate (NaLS).

Section S2. Calibration curves for reaction products.

Table 2S2-A. Equations Used to Fit HPLC Calibration Curves for Dediazonation Products.^a

Reaction Product	Calibration Equation ^b	R ²
16-ArOH	$y = 9.998 \times 10^{10} x$	0.9998
16-ArNHAc	$y = 1.416 \times 10^{11} x$	1.0000
16-ArH	$y = 9.883 \times 10^{10} x$	1.0000
16-ArF	$y = 6.955 \times 10^{10} x$	1.0000
16-ArSE	$y = 1.124 \times 10^{11} x$	0.9999
16-ArGE	$y = 1.090 \times 10^{11} x$	1.0000
16-ArAE	$y = 1.199 \times 10^{11} x$	1.0000
16-ArEC₁₂	$y = 9.997 \times 10^{10} x$	0.9945
16-ArInd	$y = 1.404 \times 10^{11} x$	0.9980

a. HPLC Eluting solvent: *i*-PrOH/MeOH, 35%/65% (v/v), or 45%/55% (v/v). Flow rate: 0.40 mL/min.

b. Units: **y**-peak area (μ Vs), **x**-concentration (molarity), and **R²** (correlation coefficient). The **y** intercept values are very small and not used in the calculations.

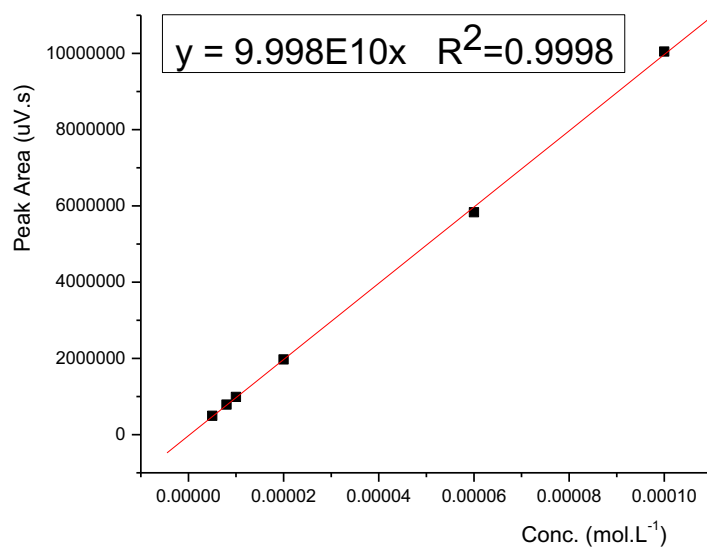


Figure 2S2-A. Calibration Curve for 16-ArOH.

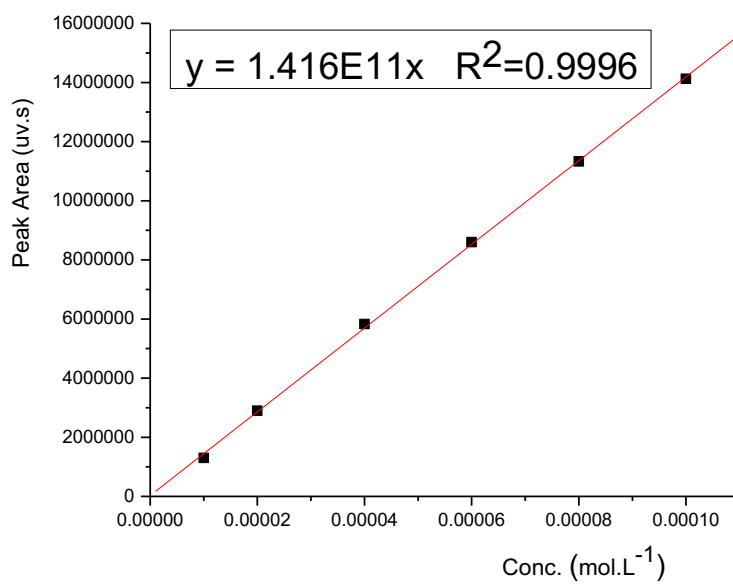


Figure 2S2-B. Calibration Curve for 16-ArNHAc.

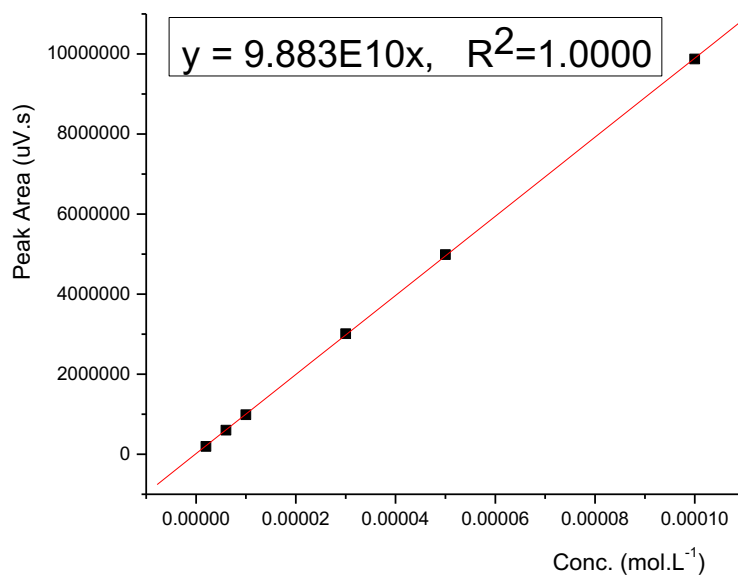


Figure 2S2-C. Calibration Curve for 16-ArH.

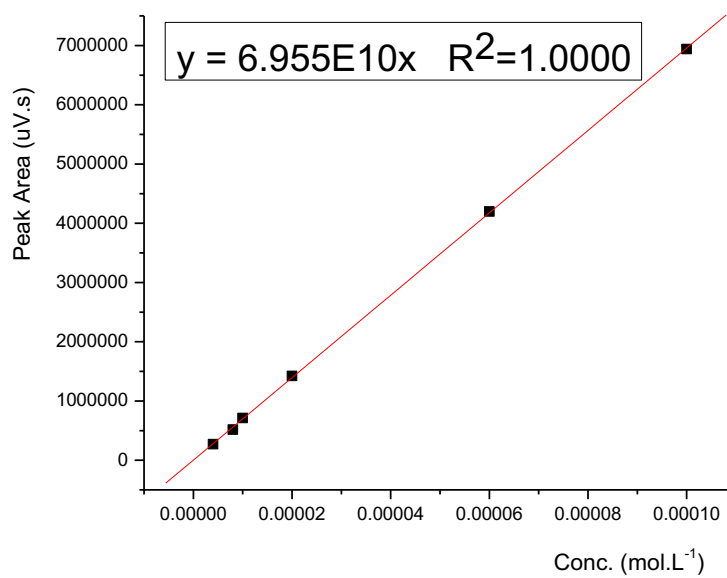


Figure 2S2-D. Calibration Curve for 16-ArF.

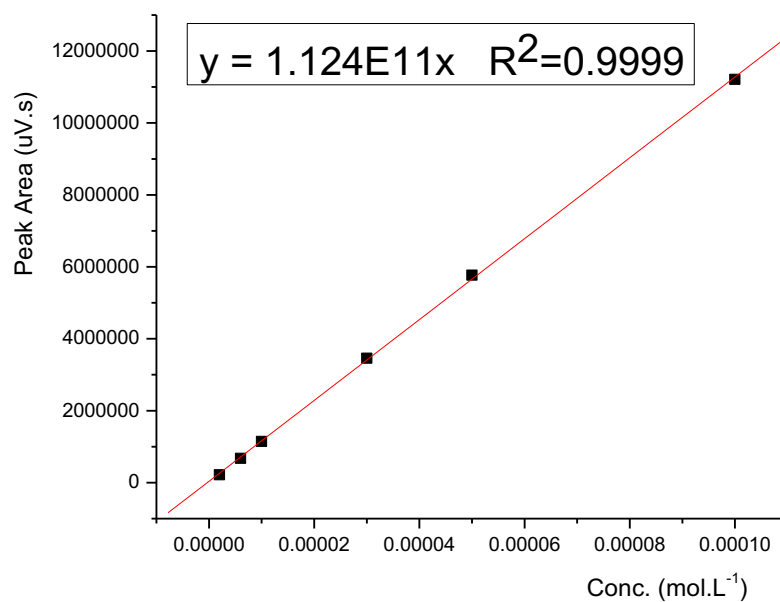


Figure 2S2-E. Calibration Curve for 16-ArSE.

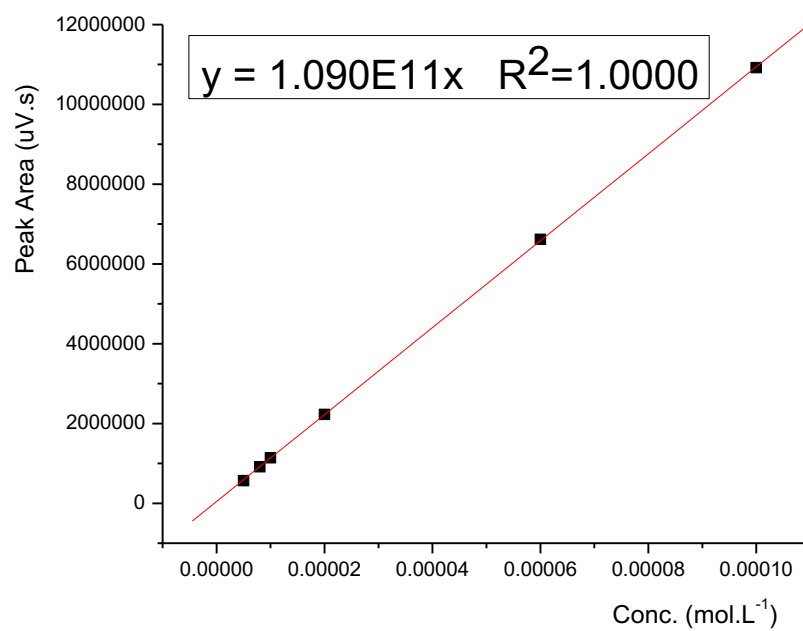


Figure 2S2-F. Calibration Curve for 16-ArGE.

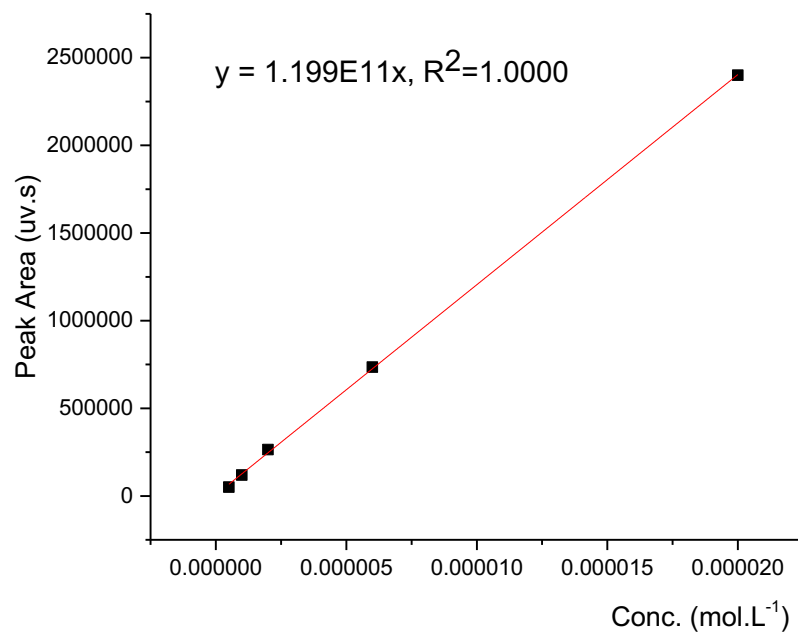


Figure 2S2-G. Calibration Curve for 16-ArAE.

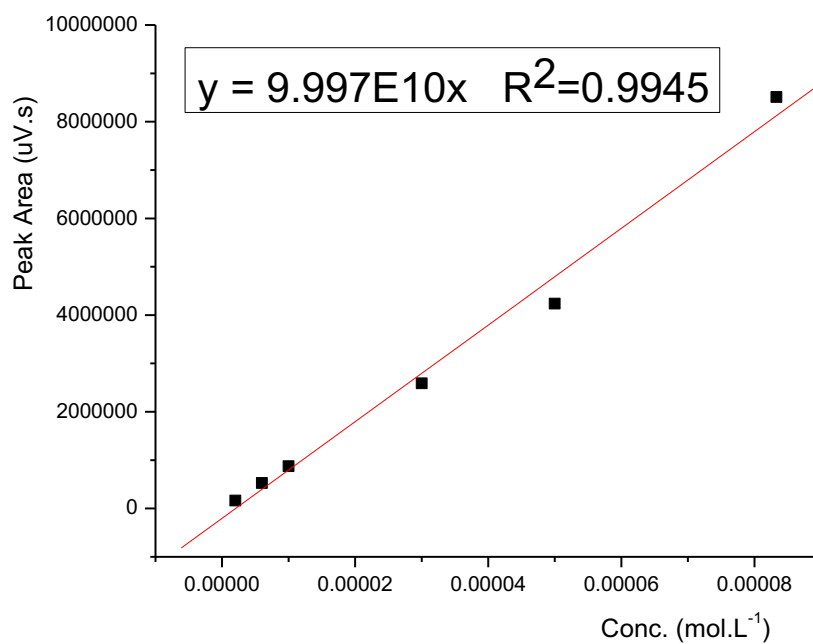


Figure 2S2-H. Calibration Curve for 16-ArEC₁₂.

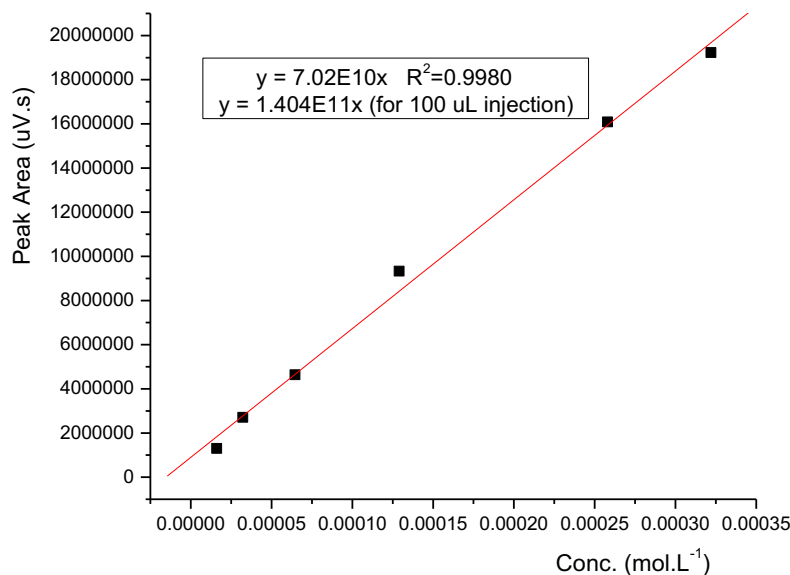


Figure 2S2-I. Calibration Curve for 16-ArInd.

Section S3. Analysis of dediazonation kinetic data.

Figures 2S3-A and **2S3-B** below show the absorbance versus time plots for the dediazonation of 16-ArN₂⁺ in C₁₂Sar and C₁₂Gly micelles at 40°C. The interfacial local concentrations of the nucleophiles, H₂O, amide oxygen and carboxylate oxygen, *etc.*, are in large excess, i.e. 16-ArN₂⁺ is the limiting reagent. Therefore, the reaction should be in the pseudo-first order and the rate of the reaction should only depend on the concentration of 16-ArN₂⁺. Values of k_{obs} were obtained by plotting the change in absorbance versus time, **Figures 2S3-C-2S3-F**, using the integrated first order rate law, **Equation S3-1**,^{3,4} where A_t is the absorbance at any time t , A_{∞} is the absorbance at infinite time, here ≥ 10 half-lives. The absorbance of 16-ArN₂⁺, A , at any time, t , is expressed as the difference in absorbance at time t (A_t), and B is the constant of integration.

$$-\ln (A_t - A_{\infty}) = k_{\text{obs}}t - B \quad (\text{Eq. S3-1})$$

The half-life of the reaction was calculated from:

$$t_{1/2} = \frac{\ln 2}{k_{obs}} \quad (\text{Eq. S3-2})$$

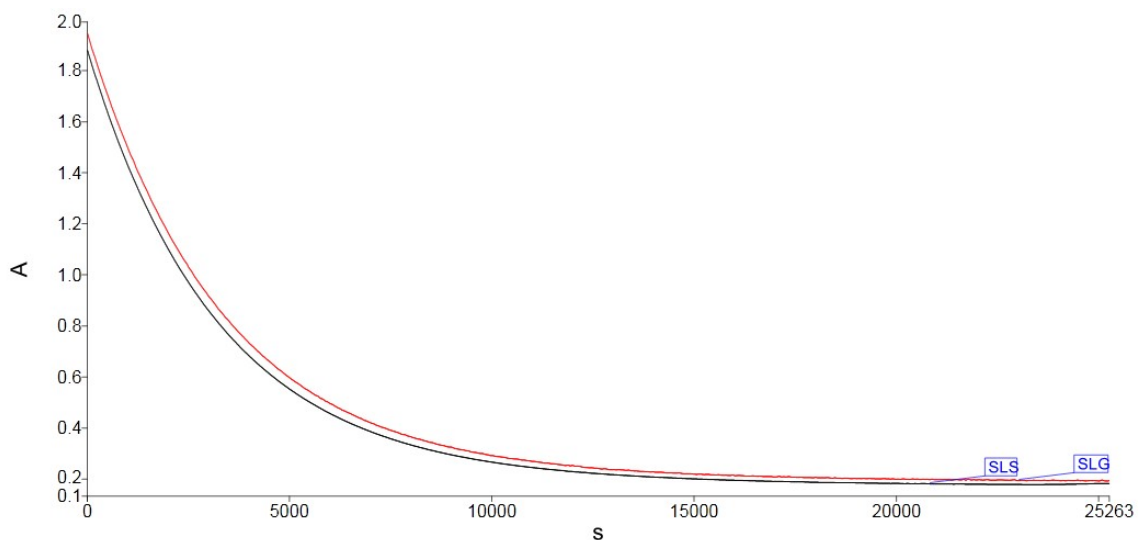


Figure 2S3-A. Decrease in Absorbance for Dediazonation of 2×10^{-4} M 16-ArN_2^+ in 0.098 M C_{12}Sar (black line) and C_{12}Gly (red line) Micelles at 285.5 nm at 40°C over 7 hours ($\geq 10 t_{1/2}$). Note: the x-axis label, S, stands for time (s) in seconds and the y-axis label, A, stands for absorbance.

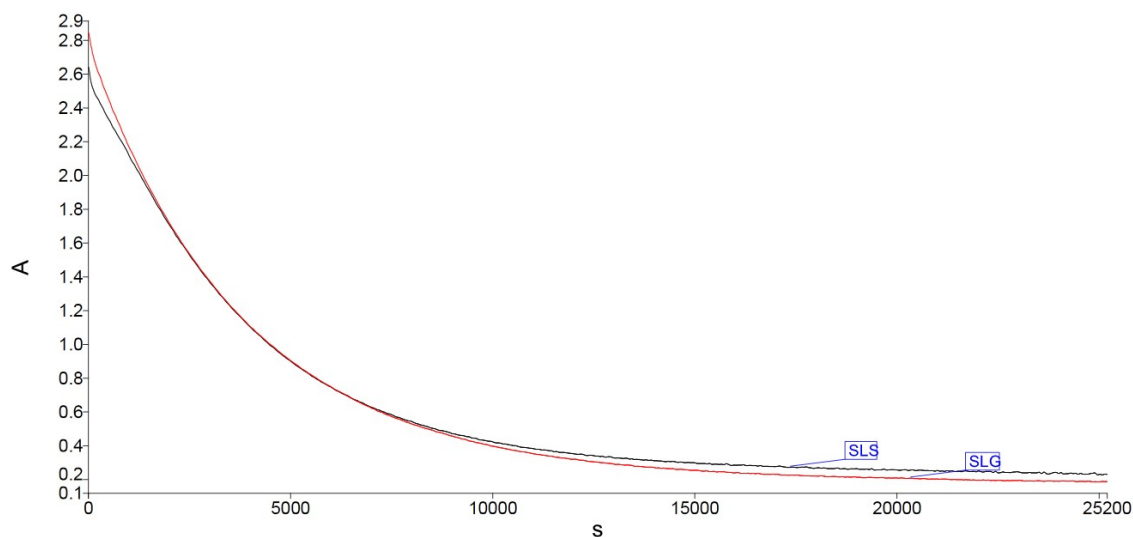


Figure 2S3-B. Decrease in Absorbance for Dediazonation of 2×10^{-4} M 16-ArN_2^+ in 0.069 M C_{12}Sar (black line) and C_{12}Gly (red line) Micelles at 285.5 nm at 40°C over 7 hours ($\geq 10 t_{1/2}$). Note: the x-axis label, S, stands for time (s) in seconds and the y-axis label, A, stands for absorbance.

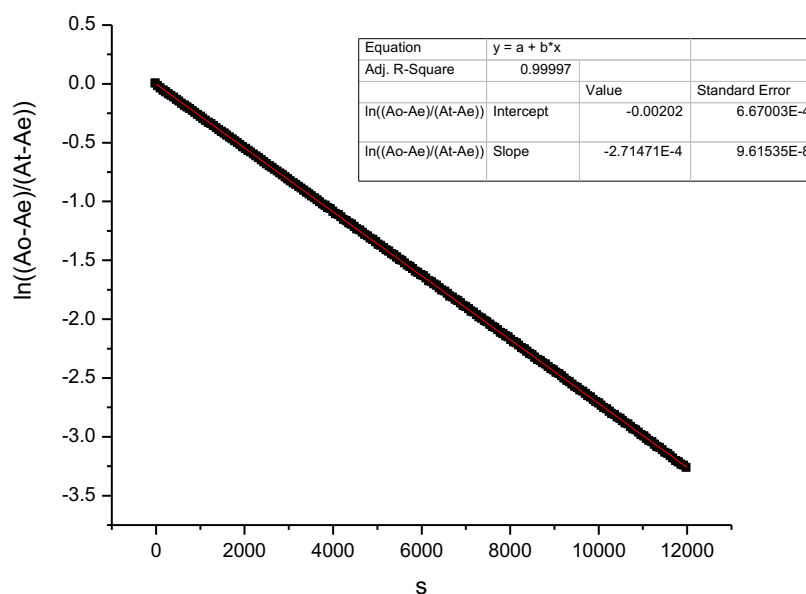


Figure 2S3-C. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN₂⁺ in 0.098 M C₁₂Sar. S stands for time in seconds.

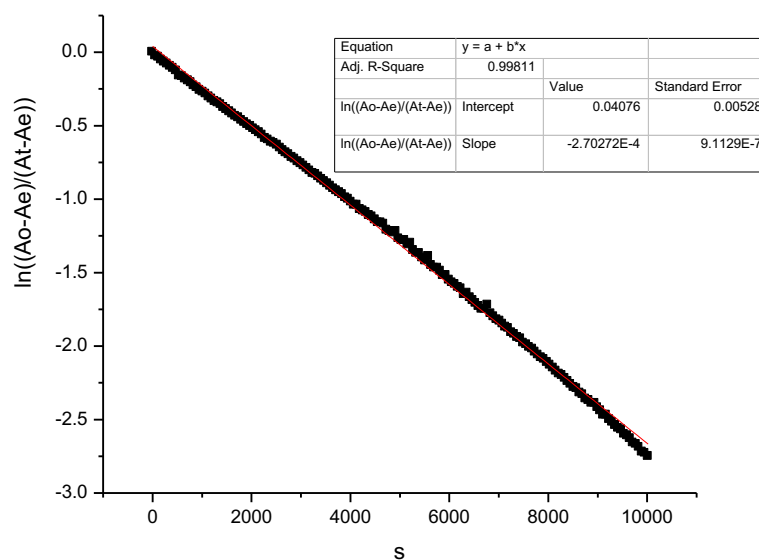


Figure 2S3-D. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN₂⁺ in 0.098 M C₁₂Gly. S stands for time in seconds.

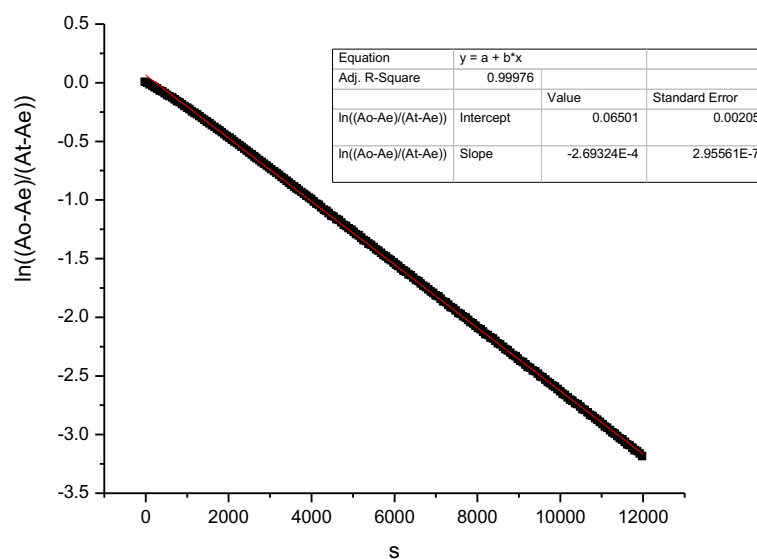


Figure 2S3-E. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN₂⁺ in 0.069 M C₁₂Sar. S stands for time in seconds.

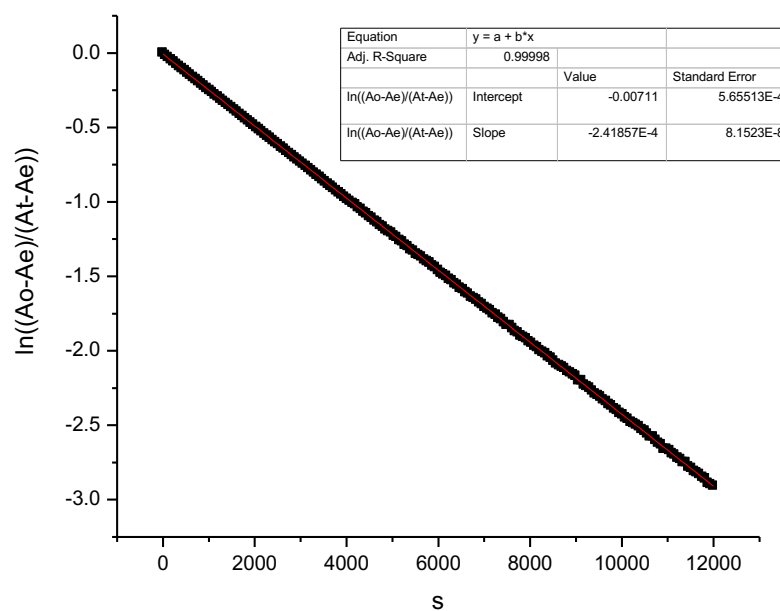


Figure 2S3-F. Ln Plot of UV Absorbance for the Dediazonation of 2×10^{-4} M 16-ArN₂⁺ in 0.069 M C₁₂Gly. S stands for time in seconds.

Section S4. Selectivity of the dediazonation reaction toward the amide oxygen and the carboxylate group.

The selectivity of the dediazonation reaction toward the amide O, compared to water (w) was determined previously.⁵ The percent yield of imido ester intermediate (16-ArOI), the immediate product from trapping by the amide nitrogen, was impossible to measure due to the rapid break down of the tetrahedral intermediate. The phenol product, 16-ArOH, from hydrolysis of the intermediate, is also not easy to measure because its corresponding peak in the HPLC chromatograms cannot be separated from the one for phenol formed by trapping by interfacial H₂O, **Scheme 2.3**. Published studies described the application of H₂O¹⁸ isotopic method, conjunct with GC/MS and HPLC, to determine the percent yield of short chain phenol (1-ArOH, the short chain analogue of the long chain phenol, 16-ArOH) from the hydrolysis.⁴⁵ The selectivity was determined in aqueous *N*-methylacetamide and *N,N*-methylacetamide solutions, and at 2 different amide concentrations, respectively. The results are summarized in **Table S4a**. Note that an average value of the estimated selectivity was used to calculate the interfacial molarities. Selectivity toward the carboxylate O, compared to water (w), was determined in aqueous glycine solutions, **Table S4b**. Note that an average value of the estimated selectivity was used to calculate interfacial molarities.

Table 2S4-A. Average Selectivities Determined from Dediazonation from 1-ArN₂⁺ in the Presence of Aqueous Amides Two Different Water/amide, N_w/N_A, Molar Ratios at 40°C.⁵

Amide	N _w /N _A	S _w ^O ^a
<i>N</i> -methylacetamide	2	0.63
	4	0.64
<i>N,N</i> -dimethylacetamide	2	0.62
	4	0.63

a. Average Value of S_w^O = 0.63

Table 2S4-B. Normalized Yields for 1-ArN₂⁺ Dediazoniation in Aqueous Glycine and the Selectivity of Carboxylate Group at 40°C.⁶

[Gly] (M)	1-ArOG(%)	1-ArOH(%)	$S_W^{(COO^-)b}$
2.70	4.85	95.15	0.87
1.48	3.00	97.00	1.05
0.78	1.58	98.42	1.09

b. Average value of $S_W^O = 1.00$

Section S5. Analysis of 16-ArEC12 and 16-ArOHh yields from hydrolysis of 16-ArOI from acetamides based on results published previously.⁵

The dediazoniation reaction with 1-ArN₂⁺ was run in concentrated aqueous acetamide *N*-methylacetamide solutions, a model compound for the hydrolysis of the imido ester from C₁₂Gly, and in *N,N*-dimethylacetamide solutions, a model compound for the formation and hydrolysis of the imido ester from C₁₂Sar, at 2.0 M and 4.0 M, respectively in the presence and absence of 43.84% H₂¹⁸O. Note that the labeled 1-ArOH is obtained only during direct reaction of H₂¹⁸O with 1-ArN₂⁺, i.e., the hydrolysis of the imido ester intermediate gives only unlabeled 1-ArOH, **Scheme 2.3**. This yield difference was analyzed by GC/MS (M and M+2 peaks) to determine the yield of 1-ArOH from reaction of 1-ArN₂⁺ with water (1-ArOH_w) and the yield of 1-ArOH by hydrolysis (1-ArOH_h). The total yields of the imido ester intermediate in the reactions are: 1-ArOI = 1-ArOAc + 1-ArOH_h, where 1-ArOH_h is corrected for the amount of 1-ArOH produced during the reaction of 1-ArN₂⁺ with water (1-ArOH_w).

On the other hand, the yield of 16-ArOH_h was obtained by assuming that the 16-ArOH_h/16-ArEC₁₂ yield ratio in C₁₂Gly, C₁₂Ala and C₁₂Sar micelles is the same as the 1-ArOH_h/1-ArOAc yield ratio from aqueous acetamides. The product yields were further modified by assuming that

the total yield of the competitively formed products, 16-ArOH, 16-ArSE (or 16-ArAE and 16-ArGE), and 16-ArEC₁₂ is 100%, where %16-ArOH = %16-ArOH_w + %16-ArOH_h and %16-ArOI = %16-ArEC₁₂ + %16-ArOH_h (see **Table 2.7-2.9**. Note that %16-ArOH_w stands for the percent yield of the phenol product by trapping with water, i.e., by subtracting the amount of phenol formed by hydrolysis of the imido ester intermediate %16-ArOH_h from the total).

Bibliography

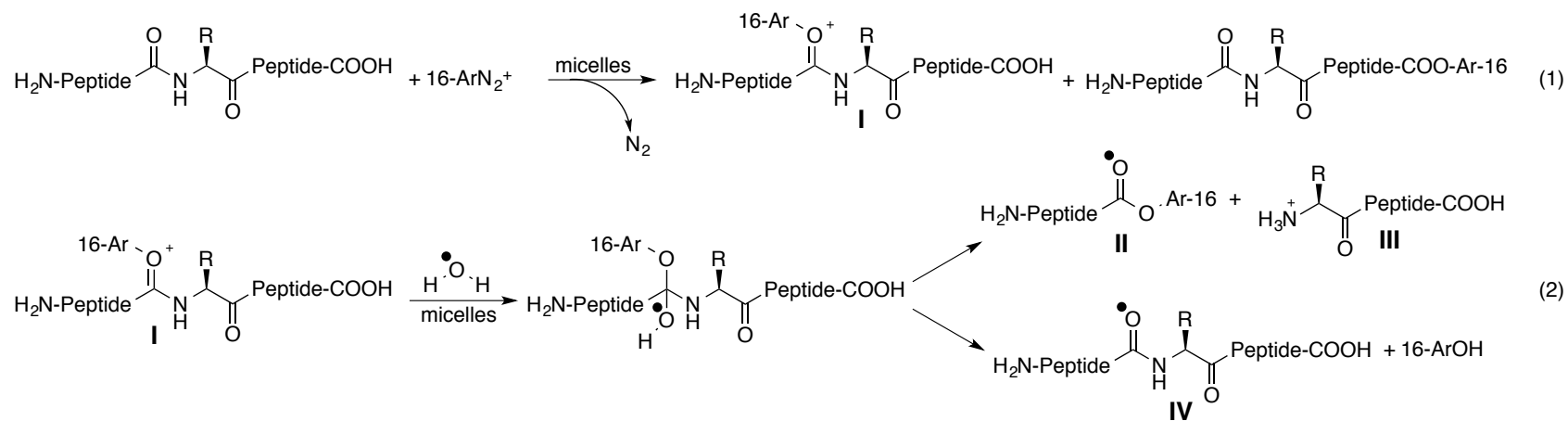
- (1) Kornblum, N. C., G.; Taylor, J. *J Am Chem Soc* **1950**, 72, 3013.
- (2) Kornblum, N. K., A.; Cooper, G. *J Am Chem Soc* **1952**, 74, 3074.
- (3) Laidler, K. J. *Chemical Kinetics*; Harper & Row: New York, 1987.
- (4) Zuman, P.; Patel, R. *Techniques in Organic Reaction Kinetics*; John Wiley & Sons: New York, 1984.
- (5) Romsted, L. S.; Zhang, J.; Zhuang, L. *J Am Chem Soc* **1998**, 120, 10046.
- (6) Zhuang, L., Rutgers, the State University of New Jersey, 1998.

Chapter III. Chemical “Fingerprints” of Peptide-based Amphiphiles - Laying the Foundation for Membrane Peptide Structure Determination

This chapter summarizes a systematic study that uses amphiphiles having multiple amino acids (short peptide) headgroups as model compounds to further demonstrate the potential of chemical imaging approach for determining topologies of peptide at bioimetic interfaces. To achieve this goal, we prepared four dipeptide amphiphiles and one pentapeptide amphiphile that is able to self-assemble in water. Self-assembled aggregates of each one of these short peptide amphiphiles provide primitive models to mimic peptides at biomembrane mimetic interfaces, and the chemical imaging approach was carefully used to probe the interfacial regions of these self-assembled aggregate structures. The degree of cleavage and tagging of the peptide models depends on the distribution of each amide bond and sidechain in the interfacial region. Therefore, analyses of various tagging and fragmentation patterns, *i.e.*, the chemical “fingerprints”, provide information on the estimated locations and local concentrations of peptide backbone and reactive sidechains within the interfacial region of these self-assembled aggregates, and may ultimately provide a unique perspective on the orientations and conformations of peptides in membrane mimetic systems.

3.1. Introduction

Amphiphilic peptide, or peptide amphiphiles (PAs) have found broad applications as bio-functional materials over the past several decades.^{1,2} Inspired by nature and by virtue of highly diverse peptide sequences consisting of, but not limited to, twenty *L*-amino acids, many PAs have been prepared with virtually limitless spatial arrangement of these molecular building blocks stabilized by covalent and non-covalent interactions. New structures are being created and used in material, pharmaceutical and biological technologies.^{1,2} Containing multiple amide bonds and



Scheme 3.1. Step (1): trapping of 16-ArN₂⁺ by the amide O to give an imidate ester intermediate **I** and by the carboxyl group to give an ester. Step (2): hydrolysis of **I** gives rise to two product pairs formed from C-N cleavage (**II** & **III**) and the C-O bond (**IV**) & 16-ArOH. Note that the black spheres indicate the oxygen source as is tracked and observed in some of the products.

various sidechains, these PAs provide an excellent model to mimic peptides at interfaces. Headgroups of peptide amphiphiles dissolved in the colloidal interface are accessible to and, therefore, react with the amphiphile molecular probe within the interface. The products are determined by an analytical protocol that has been described in *Chapter II*.

Scheme 3.1 shows the reaction pathway for the general case of peptide conjugation and amide bond cleavage of a peptide backbone. We assume that selectivities toward different amide bonds in a PA backbone are essentially the same. Note that product **II** is formed by hydrolysis of the imidate ester intermediate, **I**, via a tetrahedral intermediate into pairs of ester/amine and phenol/amide products. Product **II** is observable under the HPLC conditions used. Also, note that product **III** has no chromophore and product **IV** is the starting material. Therefore, quantitative analyses of product yields can be estimated from HPLC chromatographs of chemical imaging reactions with each peptide amphiphile. It should provide meaningful interpretation of their distributions and local concentrations of the peptide headgroups at the interface of their self-assembled aggregates.

To ensure the identification of the products and to explore the product identification protocol at diluted product concentrations, a MALD-TOF mass spectrometry analytical protocol was developed. This MALDI-TOF protocol reduces the quantity of sample solutions used for product identification.

PA-based hydrogels has been studied for a couple of decades.³ In our study, for the first time we discovered a series of PAs with alkyl chains and dipeptide headgroups form gels in aqueous solution, **Figure 3S8-A**. Reducing solution pH triggers gelation, which also depends on PA concentration and temperature.

3.2. Experimental Section

3.2.1. Materials.

All the reagents used in peptide syntheses were purchased from ChemImpex Inc. For information on other materials such as HPLC solvents and dediazonation products, please refer to *Chapter II*. The details on the syntheses of new compounds are listed in *Appendix S1* and *S2*.

3.2.2. Methods.

¹H-NMR spectra were recorded on Varian VNMRs 300 MHz or 400 MHz spectrometers, using either CD₃OD, CDCl₃ or D₂O as solvents. High resolution mass spectra were obtained on a ThermoQuest Finnigan LCQ-DUO mass spectrometer. Surface tensions of C₁₂Ala-Glu were determined by using a du Noüy ring tensiometer, a Fisher Surface Tensiomat (Model 21). Kinetic (UV) measurements were performed on a Perkin-Elmer Lambda 45 UV/VIS spectrophotometer equipped with a Peltier Temperature Programmer 6, operated with UV WinLab 6.0.3 software. HPLC measurements were performed on a Perkin-Elmer Series 200 controlled by TotalChrom Navigator 6.2.1 software. Separations were carried out using Varian Microsorb MV 250 mm length, 5 μm particle size, 100 Å pore size, C₁₈ columns, with *i*-PrOH/MeOH mixtures as eluents. Mass to charge ratios of some of the chemical imaging products were determined using an ABSciex 4800 MALDI-TOF/TOF mass spectrum analyzer. All pH values were measured on a two-buffer standardized Fisher Accumant pH meter.

The protocols of the preparations of calibration curves for dediazonation reaction products, obtaining chemical imaging data in short dipeptide and dipeptide amphiphiles, determining critical micelle concentration, and dediazonation kinetics of C₁₂Ala-Glu, are described in *Chapter II (Experimental Section)*. Product identification by ESI-MS and MALDI-TOF MS are summarized in the *Appendix, S7* of this chapter.

3.3. Results

Several pieces of information are needed to demonstrate the effectiveness of the chemical imaging approach on probing the molecular organizations at the interfacial region of PA self-assembled aggregates. (a) Use water-soluble short peptide model to demonstrate the feasibility of chemical imaging approach in tagging reactive sidechains and fragmenting multiple amide bonds simultaneously. (b) Ensure the PAs' purity and measure their cmcs. (c) Demonstrate that the mechanism of dediazonation is the same in PA colloids as in other amphiphile micelles in *Chapter II*.⁴ (d) Identify the products from chemical imaging reactions, which requires independent synthesis of each product, and authenticate each reaction product peak in the HPLC chromatograms by a spiking experiment. (e) Prepare a calibration curve for each product to convert HPLC peak areas into product yields. (f) Apply the selectivity of each nucleophilic functional group toward 16-ArN_2^+ relative to water in aqueous solution to convert product yields from dediazonation of 16-ArN_2^+ into the local concentrations of each nucleophile, i.e., each amide bond and reactive sidechain on the peptide backbone.

3.3.1. Dediazonation Reactions in the Aqueous Solution of Water-Soluble Short Peptide Model.

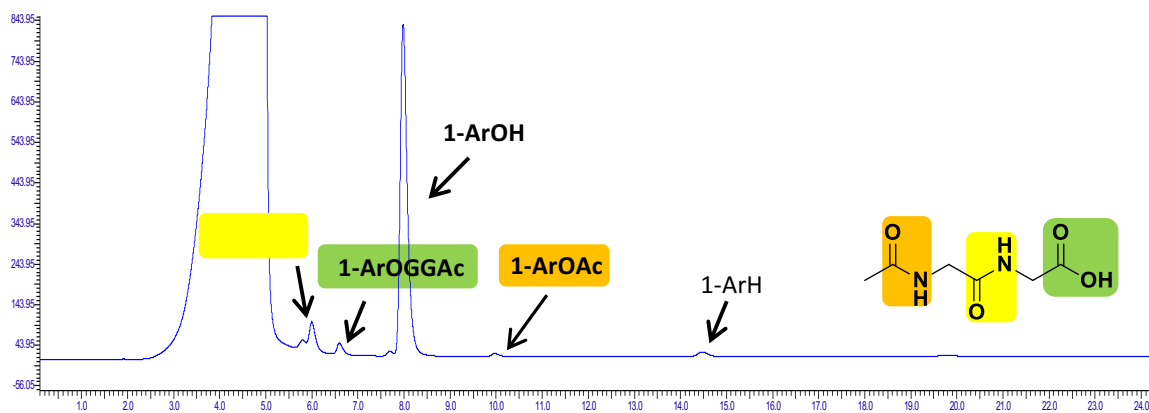


Figure 3.1. Representative HPLC Chromatogram for Reaction of Ac-Gly-Gly at a Concentration of 5.0 M at pH 4.0.

Chemical imaging reactions were initially carried out in concentrated aqueous solutions of

the *N*-capped dipeptide, Ac-Gly-Gly, a short peptide model. A short chain analogue of the long chain chemical probe, 2, 4, 6 - trimethylbenzenediazonium ion, was synthesized and used.⁵ The advantage of using Ac-Gly-Gly as the candidate peptide is its high water solubility that makes it a potentially good model to mimic the local concentration of peptide at self-assembled aggregates. Note that three expected ester products from trapping of the two amide carbonyls and the terminal carboxylate group were identified by HPLC spiking experiments, **Figure 3.1**. Observed and normalized product yields were summarized in **Table 3.1**.

Table 3.1. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 5.0 M Ac-Gly-Gly at pH 4.0 and 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%)*	Normalized Yield (%)*
1-ArOH	8	58.9	84.60
1-ArOGGAc	6	6.4	9.19
1-ArOGAc	6.3	3.4	4.88
1-ArOAc	10	0.92	1.32

[1-ArN₂⁺] = 5e-3 M, %Total = 72%.

*80% MeOH, 0.6 mL min⁻¹.

3.3.2. Dediazonation Reactions at Self-Assembled Aggregate Interfaces.

A series of short peptide amphiphiles were carefully designed, synthesized, purified, with their purity ensured and their structure characterized by multiple analytical methods including HPLC, ESI-MS and ¹H-NMR (See the *Appendix*). For solid-phase peptide syntheses, please refer to the *Appendix S1*. Note that each one of the four dipeptide amphiphiles with various peptide sequences as in the headgroups, C₁₂Ala-Glu (C₁₂AE), C₁₂Gly-Glu (C₁₂GE), C₁₂Ala-Phe (C₁₂AF), C₁₂Gly-Phe (C₁₂GF) and the peptapeptide amphiphile, C₁₂Ala-Phe-Glu-Glu-Glu (C₁₂AFEEE), may have its unique organization state at its self-assembled aggregate interface. Therefore,

chemical imaging approach is potentially able to reveal topological information of the self-assembled aggregates of each PA.

3.3.2.1. Surface Tension Determination.

Surface tension measurement was carried out to ensure the purity and determine the cmc of one candidate peptide amphiphile, C₁₂Ala-Glu (For detail of the cmc determination, see *Chapter II*). Generally, stock solution of C₁₂AE was prepared and diluted serially at various concentrations. A breaking point was observed and the cmc of C₁₂AE at room temperature and ambient pH was determined to be at ca. 14 mM (**Figure 3.2**). The solution pH changed from 10.6 to 7.6 along with the dilution of micellar solution. The surface tension profile matches those obtained with surfactants at high purity. The determination of cmc is used to help prepare PA micellar solutions whose concentration is typically 4 - 7 times higher than the cmc value. See *Section 3.3.2.3*.

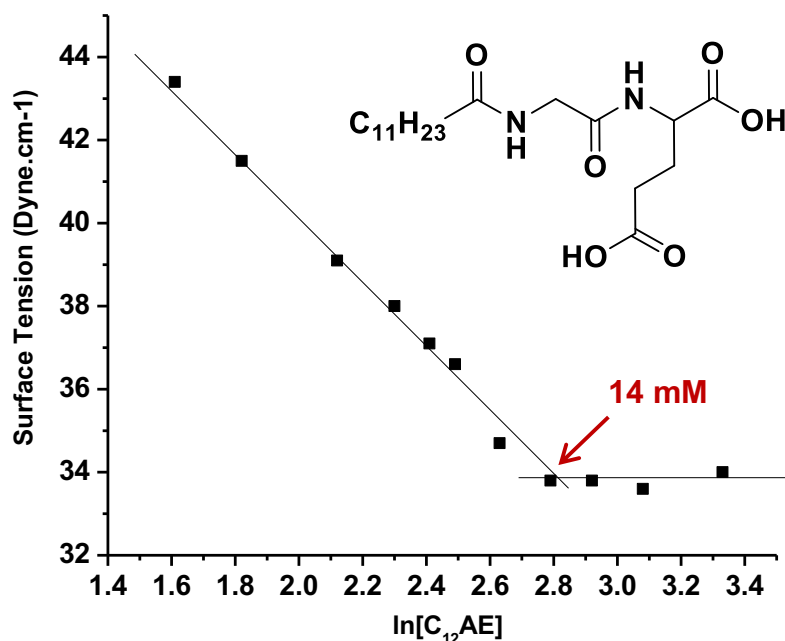


Figure 3.2. Surface Tension Measurement of C₁₂AENa at 27°C (initial pH 10.6, final pH 7.6, Cmc = 14 mM).

3.3.2.2. Reaction Kinetics.

Values for k_{obs} and half-lives for dediazonation were determined in C₁₂Ala-Glu micelles at two different pHs to examine the reaction kinetics in two different types of aggregates, and to ensure that heterolytic dediazonation mechanism behaves the same as amino acid amphiphiles in *Chapter II*, and was not changed with the solution pHs. Values of k_{obs} were obtained from the measured decrease in the absorbance of the arenediazonium ion at $40 \pm 0.1^\circ\text{C}$, $\lambda = 285.5 \text{ nm}$ (λ_{max} for 16-ArN₂⁺) with time and from plots of the data by using the integrated first order rate law. The procedure for obtaining k_{obs} and $t_{1/2}$ are summarized in *Chapter II (the experimental section)*. **Table 3.2** summarizes the results. Note that the estimated half-life of the reaction in a viscous hydrogel at pH 4.8, is more than doubled than that of pH 6.2 in solution. This indicates that the existence of noncovalent interactions, such as hydrogen bonding may have a moderate impact on the kinetics of the dediazonation reactions at the interfacial region of the self-assembled aggregates (see *Section 3.3.2.4 and discussion*).

Table 3.2. Observed Rate Constant, k_{obs} , and Half Life, $t_{1/2}$, for Dediazonation of 16-ArN₂⁺ ($6 \times 10^{-4} \text{ M}$) in 0.1 M C₁₂Ala-Glu at pH 6.2 (solution) and pH 4.8 (gel) at 40°C .

Solution pH	$10^4 k_{\text{obs}} (\text{s}^{-1})^{\text{a}}$	R^2	$t_{1/2} (\text{min})^{\text{a,b}}$
6.2	4.51	0.9986	25.6
4.8	1.99	0.9992	58.2

a. Average value of $k_{\text{obs}} = (2.64 \pm 0.1) \times 10^{-4} \text{ s}^{-1}$, Average deviation: $\pm 3.8\%$. Average half-life: 44.7 min.

b. Number of $t_{1/2}$ followed: 3 - 5.

3.3.2.3. Chemical Imaging Experiments by HPLC.

Chemical imaging reactions were carefully carried out in aqueous solutions for each one of the four dipeptide amphiphiles at various pHs. For example, at concentrations 4 - 7 times higher than the cmc of the PA, chemical imaging reactions of dipeptide amphiphile, C₁₂Gly-Glu, were

run in its colloidal solutions at 40°C and different pHs. Reaction samples were injected into HPLC producing a series of peaks with baseline separations of impurity peaks under at the eluent condition reported previously (See *Chapter II, Section 2.2*). Some of the peaks correspond to the key products that reflect their respective percent yields. Phenol product is separated and detected

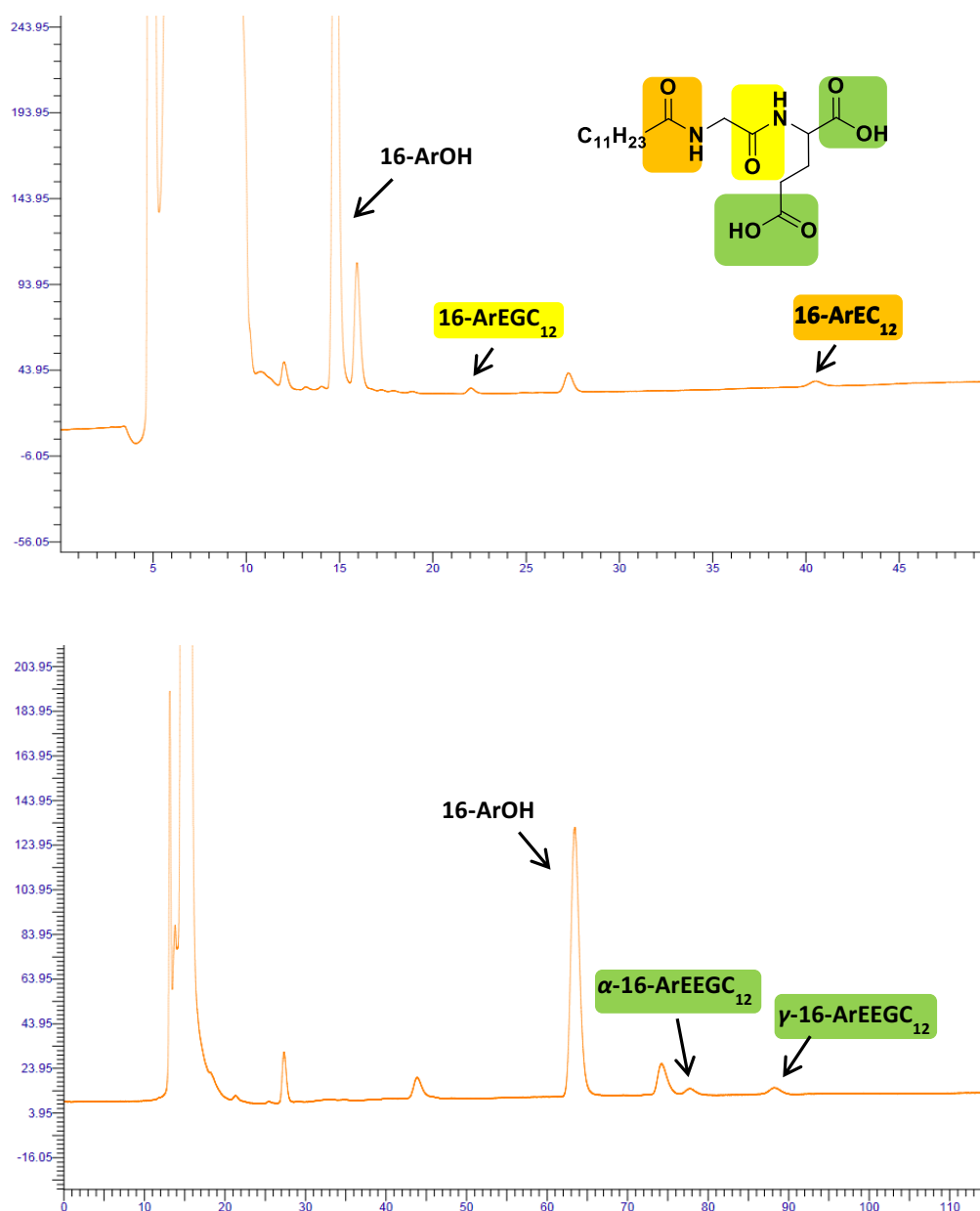


Figure 3.3. HPLC Chromatogram for Dediazotiation of 16-ArN₂⁺ in 50 mM C₁₂Gly-Glu Aqueous Solution at pH 6.0 at Two Different HPLC Eluent Conditions [Top: 35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA; Bottom: 0 or 10% IPA, 0.2 mL min⁻¹, 0.05% (v/v) TFA].

at a retention time of 13 - 14 minute at a typical eluent condition (35% IPA and 65% MeOH, isocratic) by HPLC with UV detector. A small peak was found at ca. 20 minute and another one appears at 40 - 43 minute, **Figure 3.3**. Both peaks correspond to the products formed by the cleavage of two different amide bonds. The products obtained by tagging of two terminal carboxylate groups were separated and detected at a different eluent condition (0 or 10% IPA, isocratic). This is because each of the corresponding ester products has one extra free carboxylic group that is potentially ionizable and thus gives products with different retention times detected by HPLC. The overall HPLC chromatograms provide one unique set of peaks indicating the local concentrations of H₂O, amide bonds and carboxylate side chains at the C₁₂Gly-Glu aggregate interface, which is termed as “chemical fingerprints”. The detailed HPLC eluent conditions and product yields are summarized in **Tables 3.3 - 3.4**.

Table 3.3. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.1 M C₁₂Gly-Glu Aqueous Solution at pH 6 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	63.0	92.85
16-ArGE	19	0.13	0.59
16-ArEEGC₁₂ (alpha)	55**	1.97	2.91
16-ArEEGC₁₂ (gamma)	62**	1.91	2.82
16-ArEC₁₂	37-38	0.84	3.85

[16-ArN₂⁺] = 8e-4 M, %Total = 88.9% (78.2% when 50 µL injection, no TFA buffer).

*35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA.

**0 or 10% IPA, 0.2 mL min⁻¹, 0.05% (v/v) TFA.

At rt and 40°C, results are the same.

Table 3.4. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.1 M C₁₂Gly-Glu Aqueous Solution at pH 4.5 (Gel) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	72.8	78.85
16-ArGE	19	0.19 (0.13)	0.71
16-ArEEGC₁₂ (alpha)	55**	3.10 (3.10)	3.76
16-ArEEGC₁₂ (gamma)	62**	2.87 (3.34)	3.48
16-ArEC₁₂	37-38	3.51 (3.08)	13.20

[16-ArN₂⁺] = 8e-4 M, %Total = 102.2%. (94.4% when 50 µL injection, no TFA buffer)

*35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA.

**10% IPA, 0.2 mL min⁻¹, 0.05% (v/v) TFA.

()40°C viscous solution.

A similar distribution of products from reactions with another dipeptide amphiphile, C₁₂Ala-Glu, was observed by HPLC analyses (**Figure 3.4**), with the product yields summarized in

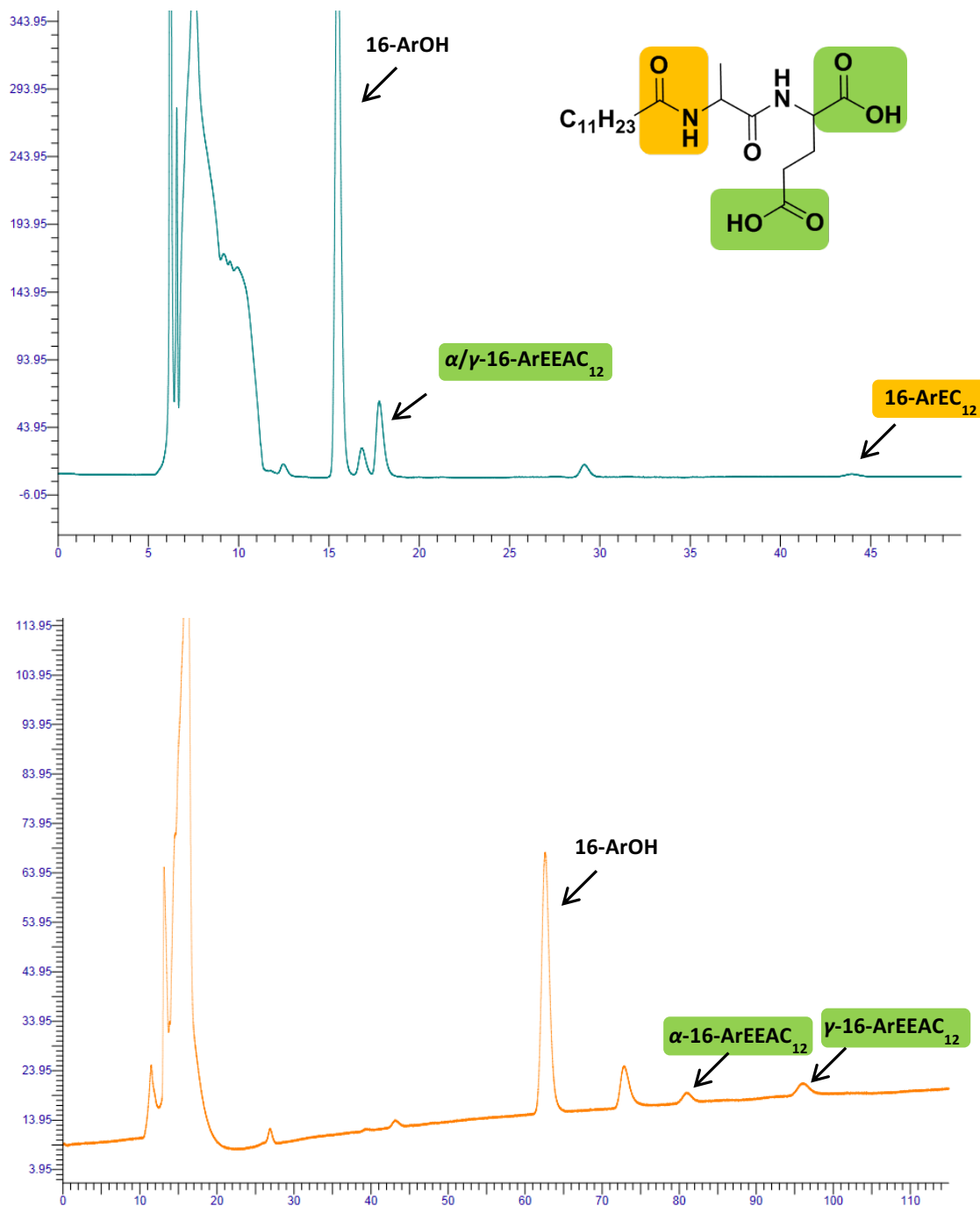


Figure 3.4. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.1 M C₁₂Ala-Glu Aqueous Solution at pH 6.0 at Two Different HPLC Eluent Conditions [Top: 35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA; Bottom: 0 or 10% IPA, 0.2 mL min⁻¹, 0.05% (v/v) TFA].

Tables 3.5 - 3.6. It should be noted that products from cleavage of the amide bond between alanine (Ala) and glutamic acid (Glu) in C₁₂Ala-Glu was not identified. In contrast, its analogue was identified previously in reactions of C₁₂Gly-Glu (see **Figure 3.3**, **Tables 3.3 - 3.4**), even though the product yield is significantly smaller than those from cleavage at the same position in C₁₂Ala-Phe and C₁₂Gly-Phe. This may indicate the variation of the orientation of amino acid side chains with the change of amino acid in peptide sequences. See *discussion*.

Table 3.5. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.1 M C₁₂Ala-Glu Aqueous Solution at pH 6 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	61.2	90.59
16-ArEEAC₁₂ (alpha)	50**	1.40	2.12
16-ArEEAC₁₂ (gamma)	58**	2.83	4.28
16-ArEAC₁₂	19-20	***	***
16-ArEC₁₂	39-40	0.64	3.01

[16-ArN₂⁺] = 8e-4M, %Total = 83.9% (79.5% when 50 μL injection, no TFA buffer).

*35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA.

**0 or 10% IPA, 0.2 mL min⁻¹, 0.05% (v/v) TFA.

***Small enough to ignore (NA).

Table 3.6. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.1 M C₁₂Ala-Glu Aqueous Solution at pH 4.5 (Gel) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	40.9	82.14
16-ArEEAC₁₂ (alpha)	50**	1.93	4.09
16-ArEEAC₁₂ (gamma)	58**	3.33	7.06
16-ArEAC₁₂	19-20	***	***
16-ArEC₁₂	39-40	1.02	6.71

[16-ArN₂⁺] = 8e-4M, %Total = 54% (lower than normal due to the surfactant effect on HPLC).

*35%IPA, 0.4 mL min⁻¹, 0.05% TFA (v/v).

**0 or 10% IPA, 0.2 mL min⁻¹, 0.05% TFA (v/v).

***Small enough to ignore (NA).

Chemical imaging experiments were arranged in the other two dipeptide amphiphiles, C₁₂Ala-Phe and C₁₂Gly-Phe. Tagging and fragmentation patterns of carboxylate groups and amide bonds in reactions of these PAs generally match those obtained from reactions with C₁₂Gly-Glu. The results are summarized in **Figures 3.5 - 3.6** and **Tables 3.7 - 3.8**.

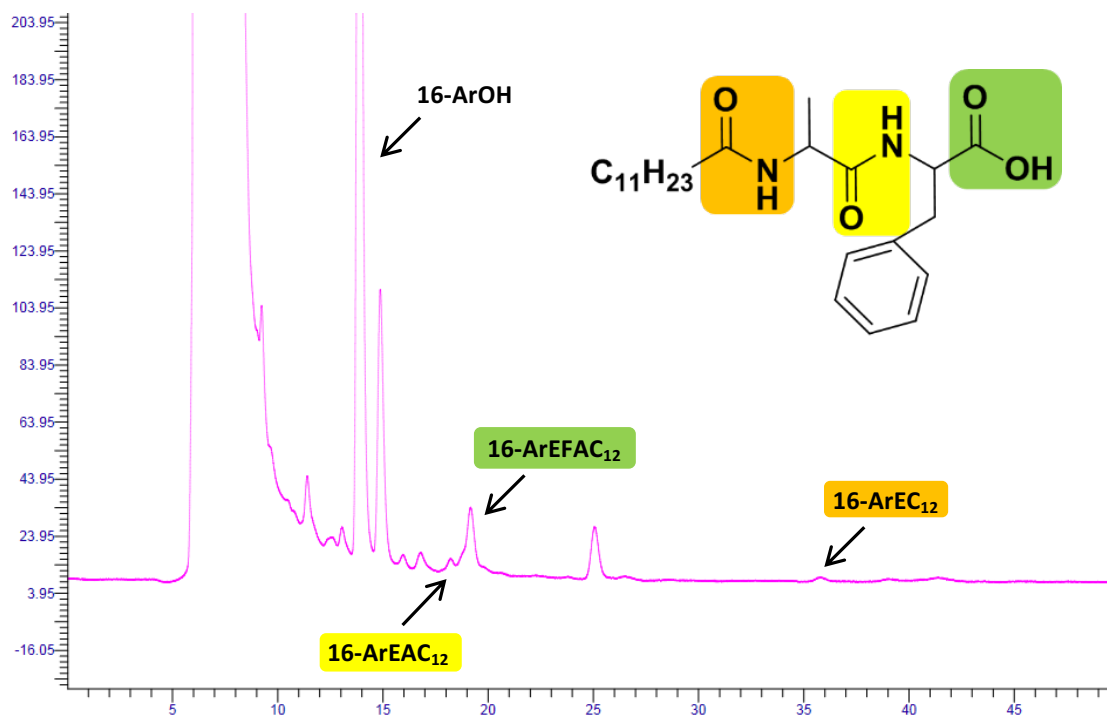


Figure 3.5. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.08 M C₁₂Ala-Phe Aqueous Solution at pH 6.8.

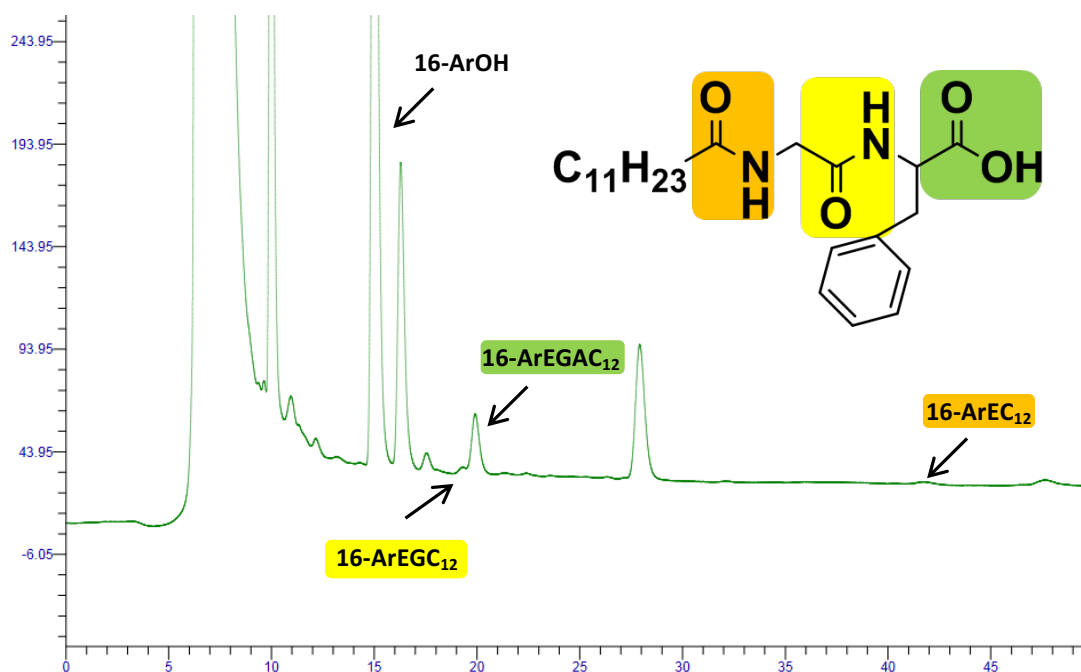


Figure 3.6. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.05 M C₁₂Gly-Phe Aqueous Solution at pH 6.5.

Table 3.7. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.08 M C₁₂Ala-Phe Aqueous Solution at pH 6.8 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%)*	Normalized Yield (%)*
16-ArOH	14-15	69.3	93.08
16-ArEAC₁₂	19	0.56	2.42
16-ArEFAC₁₂	20	2.01	2.79
16-ArEC₁₂	37-38	0.40	1.71

[16-ArN₂⁺] = 4e-4 M, %Total = 101.6%.

*35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA.

Table 3.8. HPLC Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.05 M C₁₂Gly-Phe Aqueous Solution at pH 6.5 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Percent Yield (%)*	Normalized Yield (%)*
16-ArOH	14-15	71.0	90.94
16-ArGE	19	0.73	3.01
16-ArEFGC₁₂	20	2.92	3.88
16-ArEC₁₂	37-38	0.53	2.17

[16-ArN₂⁺] = 6e-4 M, %Total = 106.4%.

*35% IPA, 0.4 mL min⁻¹.

A penta-peptide amphiphile, C₁₂AFEEEE, was synthesized and used as the peptide model at the interface. The three consecutive glutamic acid (E) is used to increase its solubility in water. Chemical imaging experiments were arranged in the micellar solutions of C₁₂AFEEEE. Three key ester products were identified (**Figure 3.7**) and their product yields were determined and summarized in **Table 3.9**. Because some of the potential ester products are ionic compounds whose separation need a better understanding of their retention behavior in the HPLC column and the optimized eluent condition to resolve the peak for each unidentified product, future studies are needed to obtain the tagging and fragmentation patterns of chemical imaging reactions of this penta-peptide amphiphile in a more accurate manner.

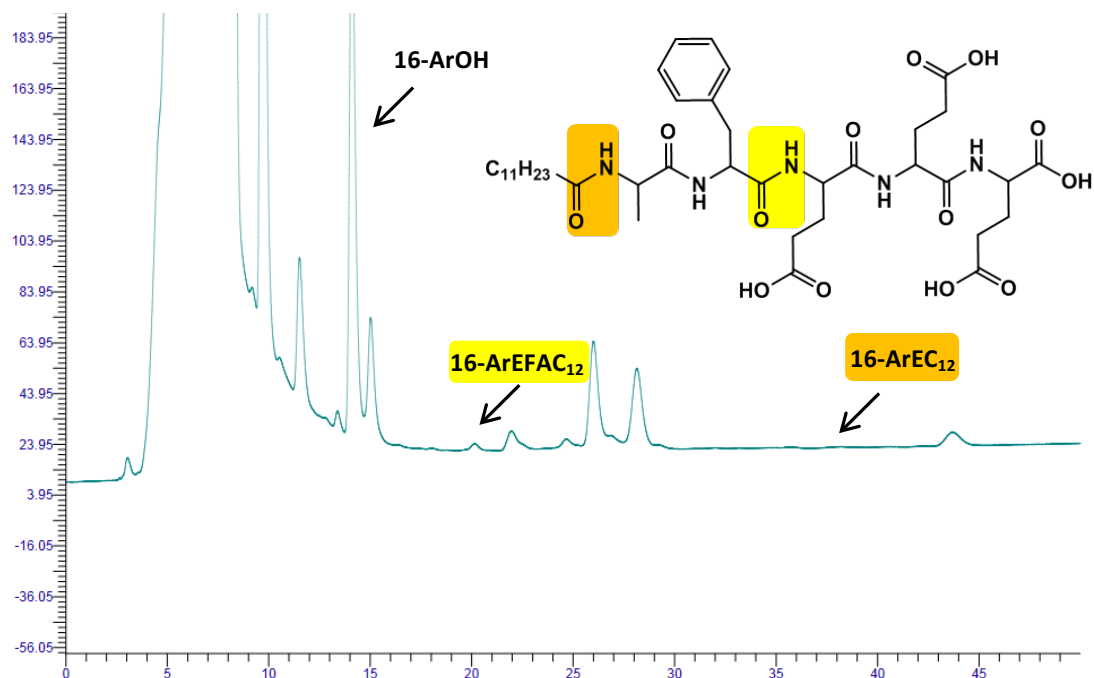


Figure 3.7. HPLC Chromatogram for Dediazonation of 16-ArN₂⁺ in 0.08 M C₁₂Ala-Phe-Glu-Glu Aqueous Solution at pH 6.0.

Table 3.9. HPLC Peak Area, Observed and Normalized Yields for the Reaction of 16-ArN₂⁺ with 0.08 M C₁₂Ala-Phe-Glu-Glu-Glu Aqueous Solution at pH 6.0 (Solution) at 40 ± 0.1°C with a Reaction Time of 12 hours.

Abbreviation	Retention Time (min)	Peak Area (μv.s)	Percent Yield (%)*	Normalized Yield (%)*
16-ArOH	14-15	4294809	63.5	98.07
16-ArEAC ₁₂	19-20	7977	0.09	0.14
16-ArEFAC ₁₂	20-21	71656	0.81	1.25
16-ArEC ₁₂	38-39	31411	0.35	0.54

[16-ArN₂⁺] = 4e-4 M, %Total = 91%.

*35% IPA, 0.4 mL min⁻¹, 0.05% (v/v) TFA.

3.3.2.4. Peptide Hydrogelation.

Solutions were also prepared at various pHs to investigate pH effects on the molecular organization in the interfacial region of the self-assembled aggregates. It should be noted that tremendous amount of studies have demonstrated that self-assembly of the PAs can promote formation of nano-structures such as cylindrical micelles or nanofibers, and, macroscopically, the hydrogel networks.³ The self-assembly process can be triggered by the change of solution pH or the addition of electrolytes, which normally appears in PAs having oligopeptide headgroups

consisting of ca. ten to a hundred amino acids.³ Interesting and to our surprise, however, in our studies except C₁₂Gly-Phe, aqueous solutions of the three ultra-short dipeptide PAs, i.e., C₁₂Gly-Glu, C₁₂Ala-Glu and C₁₂Ala-Phe, showed a pH-triggered reversible solution-gel transition. For example, 100 μ L of C₁₂GE aqueous solution at 100 mM and pH 6.0 with no added salt was prepared at room temperature. Aliquots of 1 M HCl were added to determine critical gelation pH at room temperature. Solution became viscous below pH 5 and formed a gel that was stuck at the bottom of a glass vial at pH 4.6. The gel is homogenous but a bit opaque. The critical gelation pH dropped below 4.0 at 40°C and became viscous at ca. pH 4. Gelation also depends on temperature and PA concentration. Typically, for each PA sample that was prepared at 20-50 mM and at room temperature, hydrogelation appears at a critical pH of 4.8 for PAs having glutamic acid, i.e., C₁₂AE and C₁₂GE.

Tables 3.3 - 3.4 summarize the corresponding product yield variations of the chemical imaging reactions in solutions of C₁₂Gly-Glu (see also C₁₂Ala-Glu in **Tables 3.5 - 3.6**) both in micellar solutions and in gels. Generally, as the pH decreases, solution-gel transition was triggered, concurrent with an apparent decrease of the yield of phenol product (16-ArOH) and a significant increase of the yield of ester products. This may indicate the solution to gel transition is a process in which the degree of hydration at the colloidal interface was reduced, accompanied with an increase of the local concentrations of the headgroups of these peptide amphiphiles. See *discussion* for a comprehensive explanation. For detailed information of hydrogelation, see *Appendix S8*.

3.3.2.5. Product Identifications by MALDI-TOF Mass Spectrometry.

To characterize a series of key products with their precise molecular weights, and to enhance the sensitive of the chemical imaging approach, such that reaction products whose yields are beyond the detection limit of UV lamp for HPLC, a protocol was being developed using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF).

Sample products from chemical imaging reactions were concentrated at least 10 fold higher, mixed with selected matrix compound and additives at appropriate mole ratios, sampled on a steel well-plate, and detected by instrument with a series of carefully optimized parameters (For a detailed description of MALDI-TOF protocol along with the detection method, see the *Appendix S5*.) Molecular weights of all the key ester products, except 16-ArEC₁₂, was carefully examined by MALDI-TOF, and confirmed to match their calculated values (**Figures 3.8 – 3.14**).

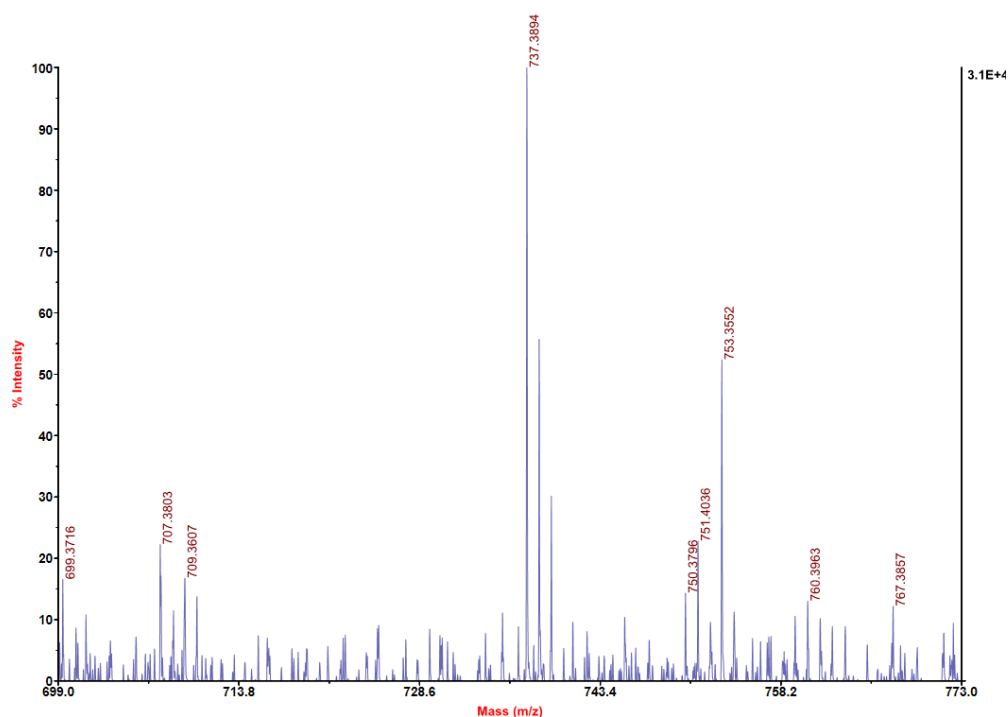


Figure 3.8. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEEGC₁₂ Gamma Ester (Note: $[M+Na]^+ = 737.4$ Da). Sample was Collected and Concentrated from Reactions of C₁₂Gly-Glu at a Concentration of 50 mM at pH 6.5.

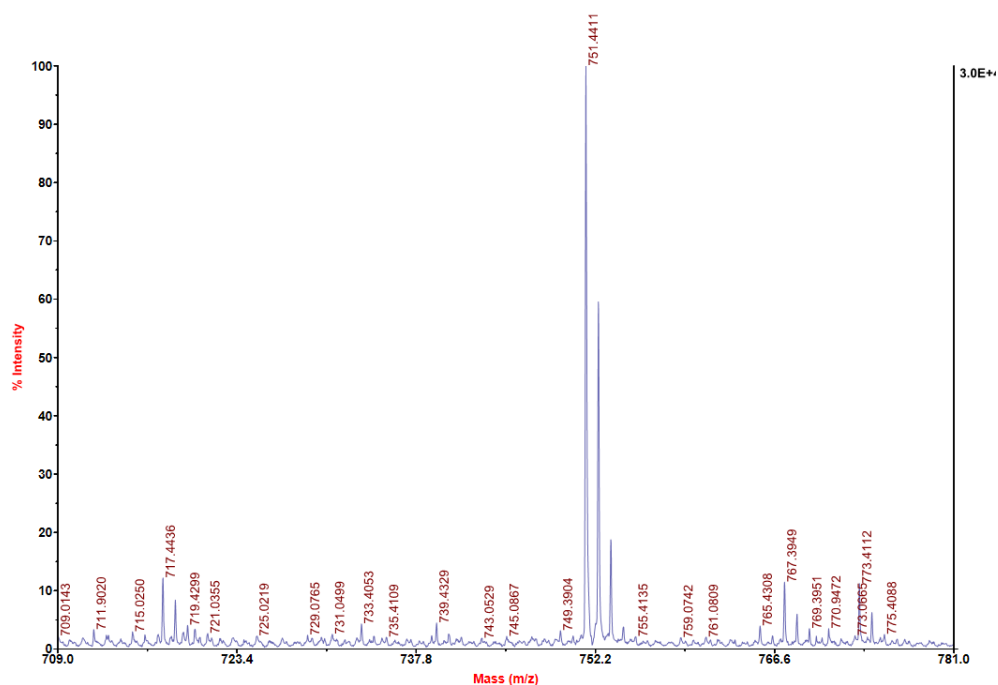


Figure 3.9. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEEAC₁₂ Alpha Ester. Note: $[M+Na]^+ = 751.4$ Da). Sample was Collected and Concentrated from Reactions of C₁₂Ala-Glu at a Concentration of 50 mM at pH 6.5.

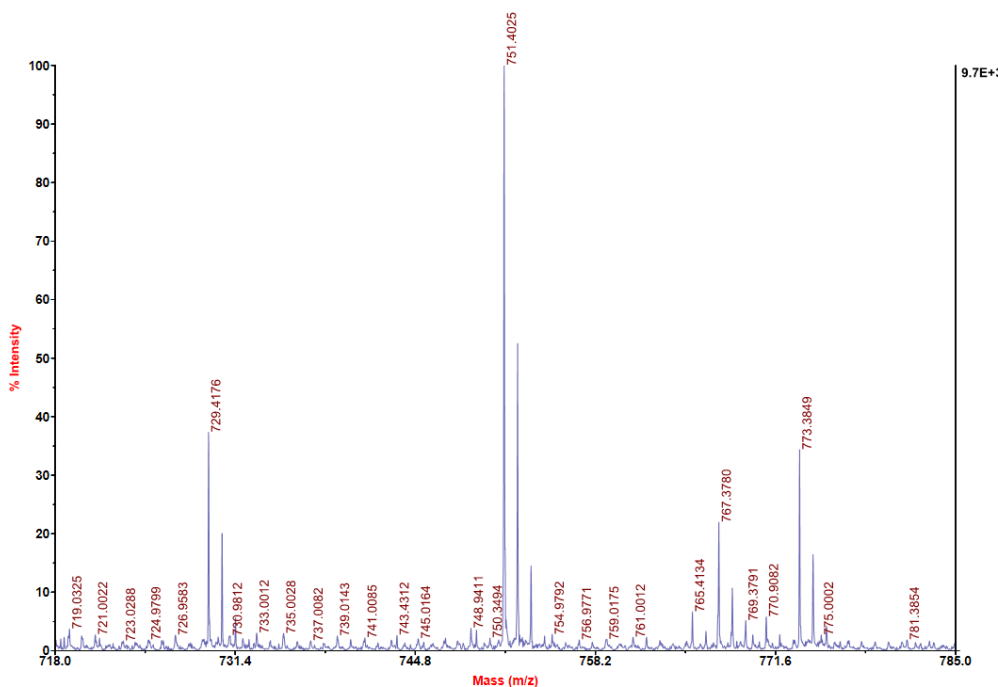


Figure 3.10. MALDI-TOF Spectrum of the Carboxylate ester, 16-ArEEAC₁₂ Gamma Ester (Note: $[M+Na]^+ = 751.4$ Da). Sample was Collected and Concentrated from Reactions of C₁₂Ala-Glu at a Concentration of 50 mM at pH 6.5.

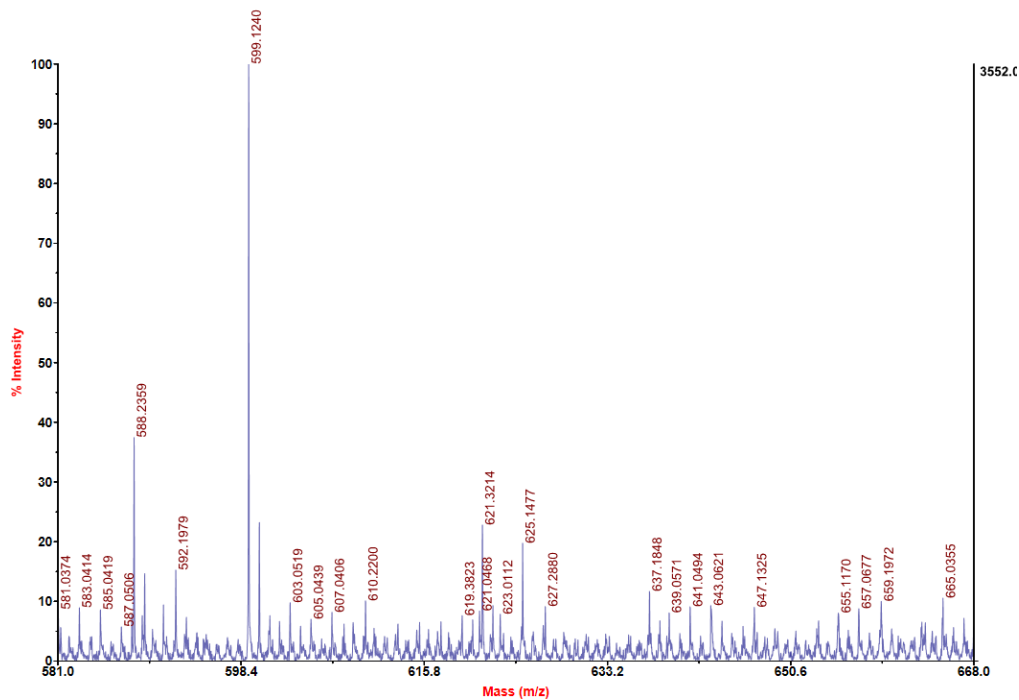


Figure 3.11. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEAC₁₂ (Equivalent to 16-ArAE. See *Chapter II*. Note: $[M+H]^+ = 599.12$ Da). Sample was Collected and Concentrated from Reactions of C₁₂Ala-Phe at a Concentration of 50 mM at pH 6.5.

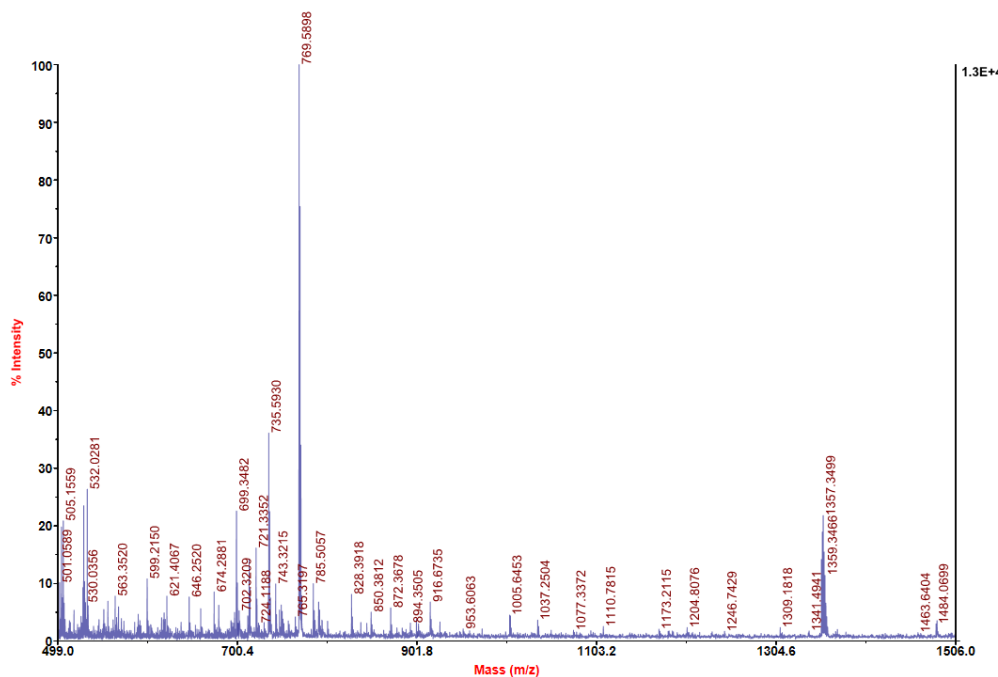


Figure 3.12. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEFAC₁₂ (Note: $[M+H]^+ = 769.6$ Da). Sample was Collected and Concentrated from Reactions of 16-ArN₂⁺ in 50 mM C₁₂Ala-Phe Aqueous Solution at pH 6.5.

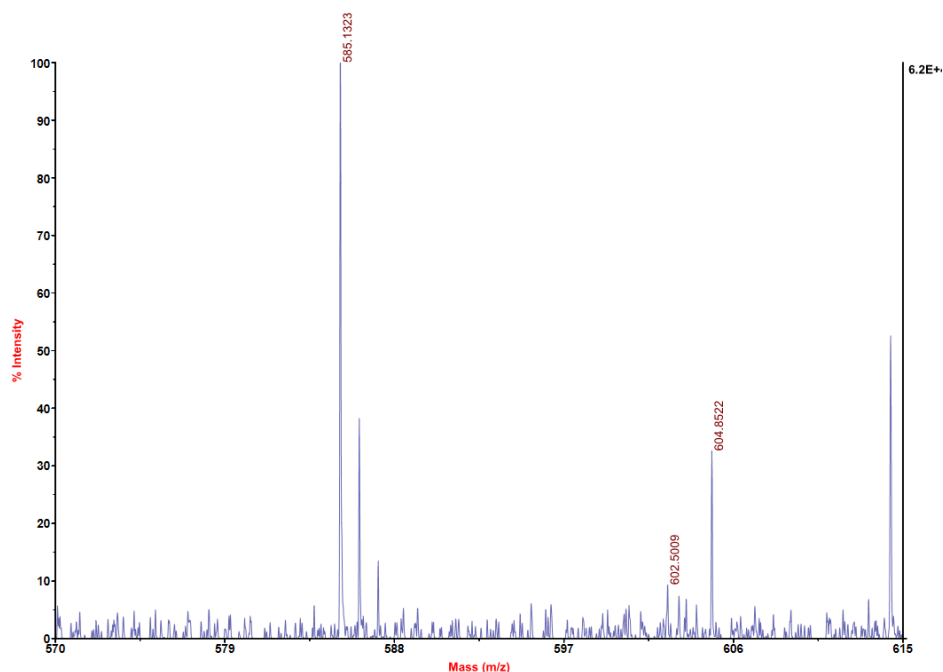


Figure 3.13. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEGC₁₂ (Note: $[M+H]^+ = 586.1$ Da). Sample was Collected and Concentrated from Reactions of 16-ArN₂⁺ in 50 mM C₁₂Gly-Phe Aqueous Solution at pH 6.5.

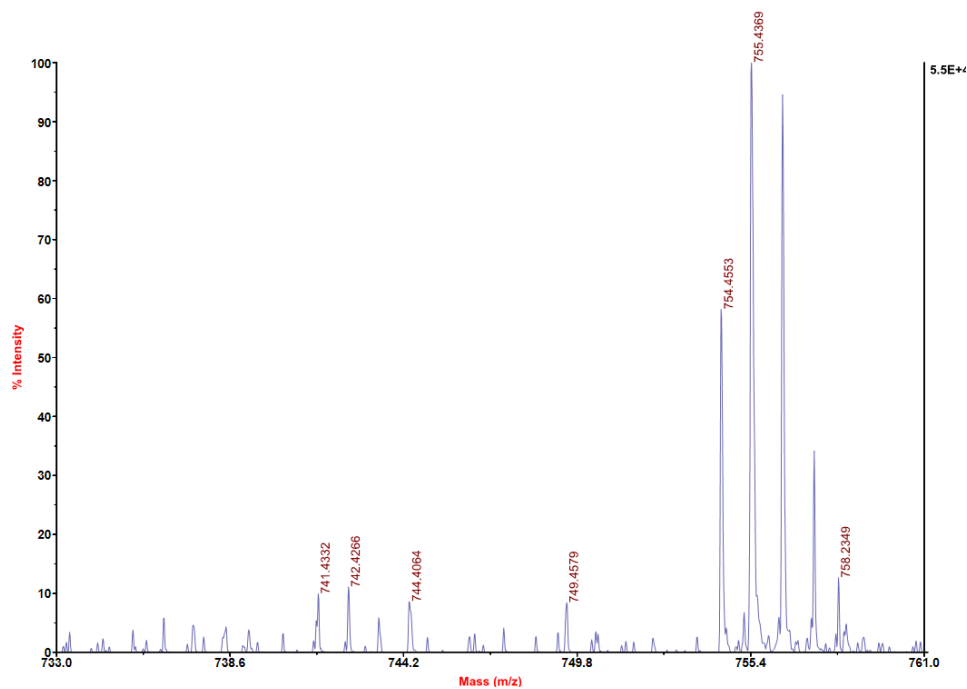


Figure 3.14. MALDI-TOF Spectrum of the Carboxylate Ester, 16-ArEFGC₁₂ (Note: $[M+Na]^+ = 755.4$ Da). Sample was Collected and Concentrated from Reactions of 16-ArN₂⁺ in 50 mM C₁₂Gly-Phe Aqueous Solution at pH 6.5.

3.3.2.6. Estimated Local Concentrations of Amide Bonds and Carboxylate Headgroups of PAs.

Previous studies¹⁰ demonstrated that the selectivity of arenediazonium probe's reaction with carboxylate group is the same as H₂O (1:1), while selectivity of amide carbonyl oxygen versus H₂O is 0.63.⁶ Having these coefficients combined with a calibration equation that was published previously,⁷ the local concentrations of interfacial H₂O, amide bonds and terminal carboxylate sidechains were estimated for each PA, and are summarized in **Tables 3.10 - 3.13**. Local concentrations of carboxylate sidechains among all four PAs are estimated and averaged at a normal range of 1 - 2.5 M, with a few exception of up to 3.9 M at acidic pH. Local concentration of amide carbonyl oxygen, however, is highly variable (ranging across 1 - 12 M) and dependent on peptide sequence and solution pH (see *discussion*). It should be noted that the estimated local concentration of H₂O is closely relevant to those of functional groups of amphiphiles at interfaces, i.e., the higher the local concentration of the surfactant headgroups, the lower the degree of interfacial hydration. This observation provides important evidence that reorganization of surfactants in self-assembled aggregate solution is concurrent with the change of the degree of hydration at the interface of self-assembled aggregates (see *discussion*).

Table 3.10. Estimated Interfacial Molarities¹ of the Nucleophiles from Chemical Imaging in Aqueous 0.100 M of C₁₂Gly-Glu at pH 4.5 (Gel) and pH 6.0 (Solution), respectively.

Amphiphile (pH)	[COO ⁻] _α	[COO ⁻] _γ	[H ₂ O]	[Peptide I] ²	[Peptide II] ²
C ₁₂ Gly-Glu (4.5)	2.02	1.87	42.36	11.26	0.61
C ₁₂ Gly-Glu (6.0)	1.56	1.52	49.88	3.28	0.51

¹%I-ArGE = 1.796 [CO₂⁻]_w + 0.130

²[Peptide I] and [Peptide II] correspond to the local concentration of amide bond that is closest to the hydrocarbon (I) and the one that is closest to the C-terminal (II).

Table 3.11. Estimated Interfacial Molarities¹ of the Nucleophiles from Chemical imaging in Aqueous 0.100 M of C₁₂Ala-Glu at pH 4.0 (Gel) and pH 6.0 (Solution), respectively.

Amphiphile (pH)	[COO ⁻] _α	[COO ⁻] _γ	[H ₂ O]	[Peptide I] ²	[Peptide II] ²
C ₁₂ Ala-Glu (4.0)	2.21	3.82	44.38	5.76	na
C ₁₂ Ala-Glu (6.0)	1.15	2.32	49.14	2.59	na

$$^1\%1\text{-ArGE} = 1.796 [\text{CO}_2^-]_w + 0.130$$

²[Peptide I] and [Peptide II] correspond to the local concentration of amide bond that is closest to the hydrocarbon (I) and the one that is closest to The C-terminal (II).

Table 3.12. Estimated Interfacial Molarities¹ of the Nucleophiles from Chemical Imaging in Aqueous 0.080 M of C₁₂Ala-Phe at pH 5-6 (Gel) and pH 6.8 (Solution), respectively.

Amphiphile (pH)	[COO ⁻]	[H ₂ O]	[Peptide I] ²	[Peptide II] ²
C ₁₂ Ala-Phe (5-6)	3.96	49.29	1.29	0.97
C ₁₂ Ala-Phe (6.8)	1.52	50.71	1.48	2.10

$$^1\%1\text{-ArGE} = 1.796 [\text{CO}_2^-]_w + 0.130$$

²[Peptide I] and [Peptide II] correspond to the local concentration of amide bond that is closest to the hydrocarbon (I) and the one that is closest to The C-terminal (II).

Table 3.13. Estimated Interfacial Molarities¹ of the Nucleophiles from Chemical Imaging in Aqueous 0.050 M of C₁₂Gly-Phe at pH 6.5.

Amphiphile (pH)	[COO ⁻]	[H ₂ O]	[Peptide I] ²	[Peptide II] ²
C ₁₂ Gly-Phe (6.5)	2.09	48.94	1.85	2.57

$$^1\%1\text{-ArGE} = 1.796 [\text{CO}_2^-]_w + 0.130$$

²[Peptide I] and [Peptide II] correspond to the local concentration of amide bond that is closest to the hydrocarbon (I) and the one that is closest to The C-terminal (II).

3.4. Discussion

It should be noted that a correlation exists between the above-mentioned pH-triggered gelation and the local concentrations of interfacial substrates. In hydrogels, the local concentration of H₂O is lower and is accompanied with a significant increase in concentrations of carboxylate groups and amide bonds (**Tables 3.10 - 3.12**). Consistent with our general understanding of molecular organization of self-assemblies, headgroups in aggregates, such as anionic carboxylate micelles, are more loosely packed at high solution pH ($> pK_a$ of the PAs), due to the electrostatic repulsion between negatively charged headgroups or stronger hydration. However, interfacial headgroups in hierarchical self-assemblies such as long micelles, vesicles or nano-fibers/sheets, are closely-packed, probably due to a reduced degree of hydration of interfacial headgroups at a lower pH ($< pK_a$ of the PAs). For example, for PAs including glutamic acid (Glu), the local concentrations of the amide bonds and carboxylate groups were observed to increase as the solution pH decreases during the sol-gel transition (**Tables 3.10 - 3.11**). This mixture of interfacial functional groups in PAs should also contribute to the formation of the hydrogelation networks, stabilized by weak hydrogen bonding interactions. On the other hand, gelation of C₁₂Ala-Phe is concurrent only with an increase of the local concentration of terminal carboxylate groups, but not amide bonds. This implies that the hydrogelation network of C₁₂Ala-Phe is primarily stabilized by aromatic π - π stacking (owing to the aromatic sidechain of phenylalanine). Alanine is incorporated in its peptide sequence, brings chirality and rigidity to the peptide backbone, and therefore, may contribute to stabilization of the gelation network (note that its structural analogue, C₁₂Gly-Phe, with an achiral amino acid in its peptide sequence, was not an ideal hydro-gelator, but formed precipitates at an acidic solution pH).

MALDI-TOF provides an alternate but more convenient detection method that could potentially contribute to a more sensitive analytical protocol for the identification and quantification of the ester products. Future research is needed to optimize the analytical condition

and instrumental parameters, such that products could be detected at one or multiple orders of magnitude lower concentrations than those in the current approach.

In general, chemical “fingerprints” of a series of short PAs were successfully obtained that provide information on molecular organization and distribution of amide bonds and carboxylate groups at the interface of peptide-based self-assemblies. Based on the data summarized in **Tables 3.10 - 3.13**, we summarized the information below.

a). The alanine (Ala) sidechain, $-\text{CH}_3$, increases the hydrophobicity of a PA, leading to a larger fraction of amide bond and carboxylate group buried in the hydrocarbon core of self-assemblies such as micelles. This observation also matches our previous determination on the local concentration of amide bonds of amino acid amphiphiles (see *Chapter II*). Replacing alanine with glycine in peptide sequence minimizes this effect because peptide backbone distribution is shifted further into the interfacial region that is more reactive arenediazonium ion chemical probe.

b). Glutamic acid (Glu) enhances the hydrophilicity of a PA, due to its di-carboxylate groups. This leads to a larger fraction of carboxylate group extended toward the bulk region instead of remaining located in the interfacial region.

c). The product from cleavage of the amide bond between alanine (Ala) and glutamic acid (Glu) in $\text{C}_{12}\text{Ala-Glu}$ were not identified because chemical imaging experiments showed no signs of this product (**Table 3.11 & Figure 3.16**, bond marked in blue). In contrast, its structurally similar counterpart was identified in reactions of $\text{C}_{12}\text{Gly-Glu}$ (**Table 3.10 & Figure 3.17**, bond marked in blue), even though the product yield is significantly smaller than those from cleavage at the same position in $\text{C}_{12}\text{Ala-Phe}$ and $\text{C}_{12}\text{Gly-Phe}$. This may be because of the orientation of the γ -carboxylate sidechain in glutamic acid. The presence of γ -carboxylate sidechain may shield the amide bond in its vicinity and reduce amide bond reactivity with the chemical probe, which is probably due to the steric hindrance caused by the hydrogen bonding interaction between γ -carboxylate sidechain and the amide bond (**Figure 3.16**, hydrogen bond not specified). Note that

the product formed by reaction of γ -carboxylate sidechain in C₁₂Ala-Glu, has a significantly estimated local concentration than that of α -carboxylate (**Table 3.11**). However, these two headgroups were more evenly distributed in the reactions of C₁₂Gly-Glu (**Table 3.10**), implying that the orientation of the two carboxylates at interface are moderately different in C₁₂Ala-Glu (with γ -carboxylate sidechain orienting back to the amide carbonyl, **Figure 3.17**), but essentially the same in C₁₂Gly-Glu. Replacing alanine (as in C₁₂Ala-Glu) with an achiral amino acid, glycine (as in C₁₂Gly-Glu), however, not only endows PA with more hydrophilicity, but provides the peptide segment with less rigidity, such that the amide bond becomes more accessible to 16-ArN₂⁺ (**Figure 3.17**), leading to a substantially higher percent yield of the corresponding ester product.

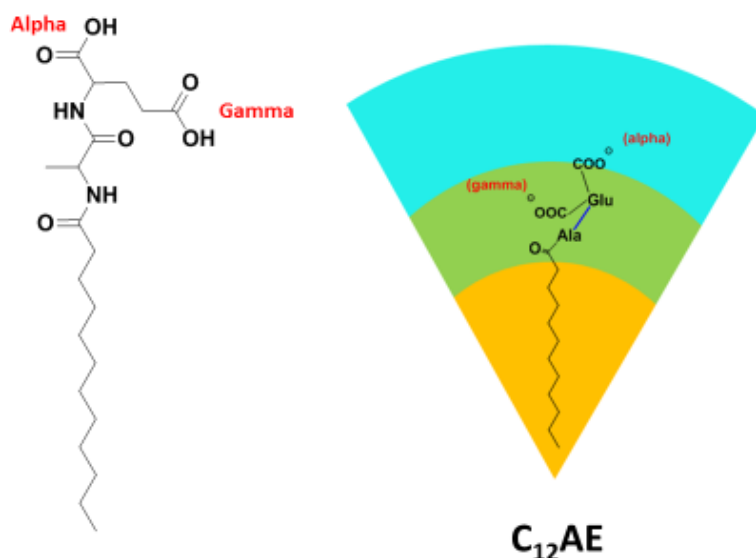


Figure 3.15. Chemical Structure of Short Peptide Amphiphile, C₁₂Ala-Glu and Predicted Orientation of Cleavable Amide Bonds and Carboxylate Sidechains at Interface of Self-Assemblies. Note the Proposed Pseudophase Model: Micelle Core (Orange), Interfacial Region (Green) and the Bulk Aqueous Region (Blue).

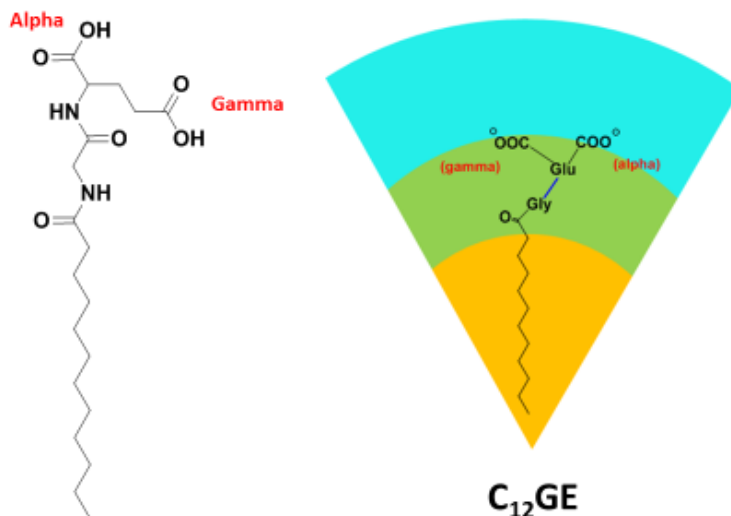


Figure 3.16. Chemical Structure of Short Peptide Amphiphile, $C_{12}Gly-Glu$, and Predicted Orientation of Cleavable Amide Bonds and Carboxylate Sidechains at Interface of Self-Assemblies. Note the Proposed Pseudophase Model: Micelle Core (Orange), Interfacial Region (Green) and the Bulk Aqueous Region (Blue).

3.5. Summary

Chemical “fingerprints” of a short peptide, and its long chain amphiphilic analogues (PAs) in their self-assembled aggregate aqueous solutions were successfully obtained. The results provide information on the interfacial hydration and the local concentrations of multiple amide bonds and terminal carboxylate groups. Product yields from reaction with $16-ArN_2^+$ show that the local concentrations of carboxylate sidechains are averaged at 1.0 - 2.5 M at a solution pH of 6 - 7, but approach ~ 4.0 M at pH 4.5 when the dipeptide amphiphiles form hydrogels. Local concentrations of amide bonds are generally dependent on the peptide sequences, ranging across 1.0 - 3.0 M at a solution pH of 6 - 7, but are 2 - 3 times higher at gelation pHs. Glutamic acid, with both carboxylate headgroups oriented away from the hydrocarbon core, makes its neighboring amide bonds more soluble in the interfacial region, and therefore, more accessible to the reactions with $16-ArN_2^+$. Product yields also show that the presence of glutamic acid minimizes the probe’s accessibility to the amide bond between itself and the nearby amino acid,

perhaps due to steric hindrance caused by hydrogen bonding between the γ -carboxylate side-chain and its nearby amide bond.

Chemical imaging approach is beginning to provide unique insights into molecular organization of peptides at interfacial region of membrane mimics, and may ultimately provide a unique perspective on the orientations and conformations of membrane proteins or peptides in membrane mimetic systems.

Bibliography

- (1) Peptide-Based Materials. Topics in Current Chemistry (Book 310). Deming, T. volume editor. Springer; 2012.
- (2) Ulijn, R. Peptide-Based Materials via Molecular Self-Assembly. In: Castillo, J, Sasso L, Svendsen, WE, editors. Self-Assembled Peptide Nanostructures: Advances and Applications in Nanobiotechnology. Pan Stanford Publishing; 2012. p. 67-91.
- (3) Altunbas, A.; Pochan, D. Peptide-Based and Polypeptide-Based Hydrogels for Drug Delivery and Tissue Engineering. In Peptide-Based Materials: Topics in Current Chemistry (Book 310). Deming, T. volume editor. Springer; 2012.
- (4) Zhang, Y.; Romsted, L.; Zhuang, L.; de Jong, S. Simultaneous Determination of Interfacial Molarities of Amide Bonds, Carboxylate Groups, and Water by Chemical imaging in Micelles of Amphiphiles Containing Peptide Bond Models. *Langmuir*, **2013**, 29, 534-544.
- (5) Romsted, L. Do Amphiphile Aggregate Morphologies and Interfacial Compositions Depend Primarily on Interfacial Hydration and Ion-Specific Interactions? The Evidence from Chemical imaging. *Langmuir*, **2007**, 23, 414-424.
- (6) Romsted, L.; Zhang, J.; Zhuang, L. Mechanism of Reaction of an Arenediazonium Ion in Aqueous Solutions of Acetamide, *N*-Methylacetamide, and *N, N*-Dimethylacetamide. A Potential Method for Chemically Tagging Peptide Bonds at Aggregate Interfaces. *J Am Chem Soc* **1998**, 120, 10046-10054.
- (7) Zhuang, L. Mechanism of Reaction of Arenediazonium Ions with the Amide Bond. Ph.D. Thesis. Rutgers, The State University of Jersey, 1998.

Appendix

Section S1. Syntheses of peptide amphiphiles.

Syntheses of C₁₂AE, **Scheme S1**, and its two intermediates, α -OtBu-C₁₂AE and γ -OtBu-C₁₂AE, C₁₂AF, C₁₂GE, C₁₂GF were based on a solid-phase peptide synthesis protocol that is summarized below.

Measure and transfer 0.5 mmol 2-chlorotrityl chloride resin into peptide synthesis glass vessel, then add 2.5 mL anhydrous dichloromethane (DCM) into the vessel. Apply N₂ pressure for 5 min to make resin well swelled. Remove DCM from the vessel by applying N₂ pressure.

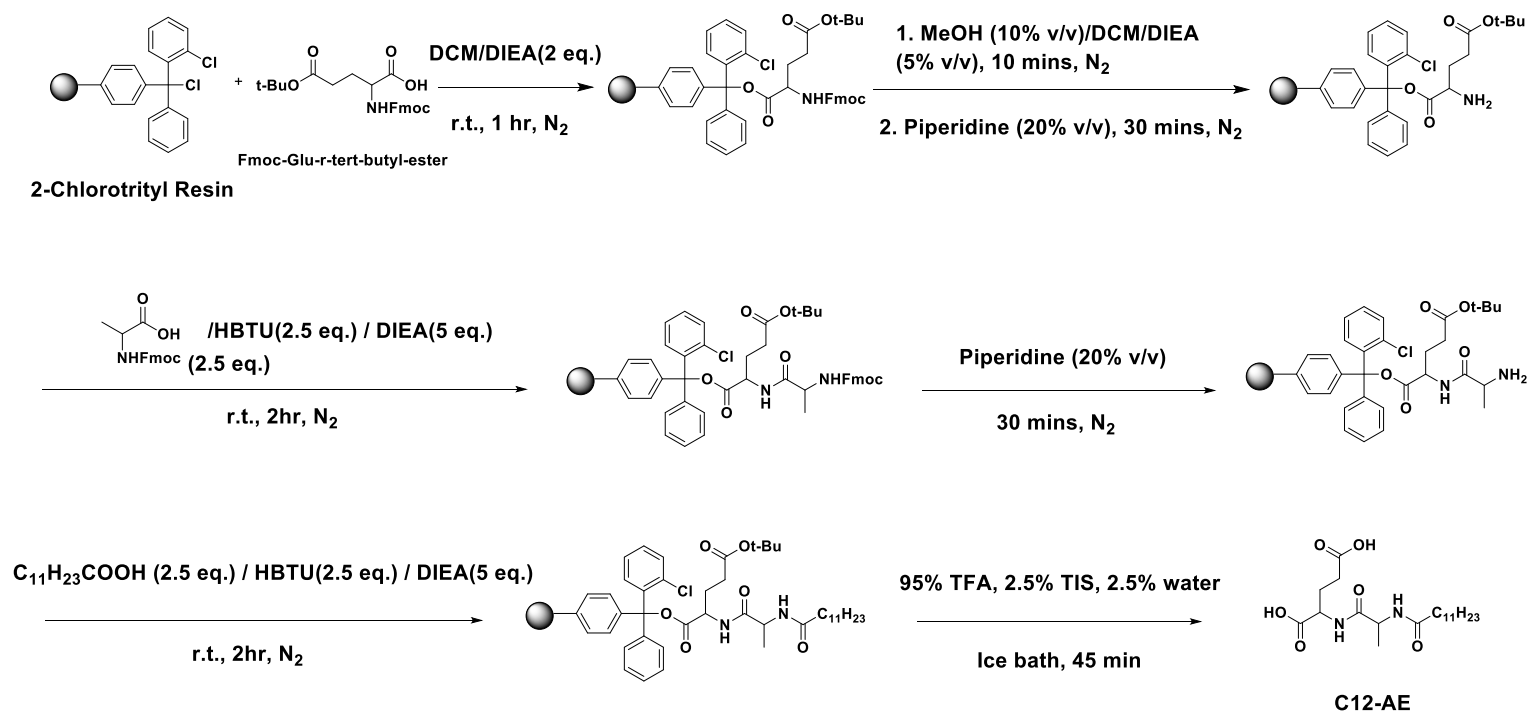
Step 1: DIEA (0.6 mmol) was added to a 2 mL solution of 0.6 mmol Fmoc-protected amino acid dissolved in anhydrous DCM. This mixture was transferred to the vessel and an extra aliquot of 0.6 mmol DIEA added, and the reaction proceeded for 1 hr under N₂ protection.

Step 2: the liquid phase was removed from the vessel and the resin washed for 5 times, with 5 mL anhydrous DCM for 1 min at each time. Then 3 mL of a mixture of anhydrous DCM, DIEA and anhydrous MeOH (volume ratio: 17:1:2) is added and the reaction takes place for 10 min with N₂ protection.

Step 3: the liquid phase was removed from the vessel and the resin washed as in **Step 2**, then with DMF. Then 5 mL 20% (%vol) piperidine in DMF solution is added and the reaction takes place for 30 min with N₂ protection. Then the resin is washed first with 5 mL 20% piperidine in DMF, then with 5 mL DMF, 1 min each time for a total of 8 times. Dump all the liquid.

Step 4: to the resin is added a mixture of the 2nd Fmoc-protected amino acid (1.5 mmol, or the long chain carboxylate acid: C₁₁H₂₃COOH), HBTU (1.5 mmol) and DIEA (3 mmol) dissolved in 3 mL anhydrous DMF. The reaction takes place for 2 hr with N₂ protection.

Repeat *Step 3 & 4* as needed.



Scheme 3S-1. Example synthetic route of peptide amphiphile, C₁₂AE.

Then have the resin washed 8 times with DMF, then 5 times with DCM

To the vessel is added a mixture (3 mL) of 95% TFA, 2.5% TIS and 2.5% dl H₂O and have the reaction take place for at least 45 min. For PAs that have *t*-Bu protecting groups, a slightly different treatment was used. Typically, multiple aliquots of 1% TFA in DCM solution was added and used to wash resin for 1 min, with the filtrate collected. Transfer the filtrate and wash the residue with TFA or DCM multiple times. Collect all the filtrates and evaporate the solvents with a rotary evaporator (with toluene being added at least twice to help get rid of excessive TFA). The crude product is further purified by HPLC or recrystallization (with EtOH). Basically, crude product was dissolved in MeOH. An aliquot of 1 M HCl was added to adjust the pH to be acidic at ca. 1 - 1.5. Solvent was evaporated and re-dissolved in ethyl acetate and washed with distilled water twice. Ethyl acetate was evaporated and then the remaining product was rinsed with aliquots of hexanes three to five times to give white product. Products undergo further purification by HPLC depending on the purity.

***a.* N-lauroyl-L-alanyl-L-glutamic acid. C₁₂AE.**

Mw (calc.): 400.5 g mol⁻¹. ¹H NMR (300 MHz, CD₃OD): δ ppm 0.89 (3H, t), 1.32 (16H, bs), 1.60 (5H, m), 1.94 (2H, m), 2.22 (2H, m), 2.41 (2H, t), 4.38 (2H, m), 8.12 (1H, d), 8.21 (1H, d). HRMS (ESI) *m/z* (found): 399.22 g mol⁻¹ (M⁻), **Figure 3S1-A**.

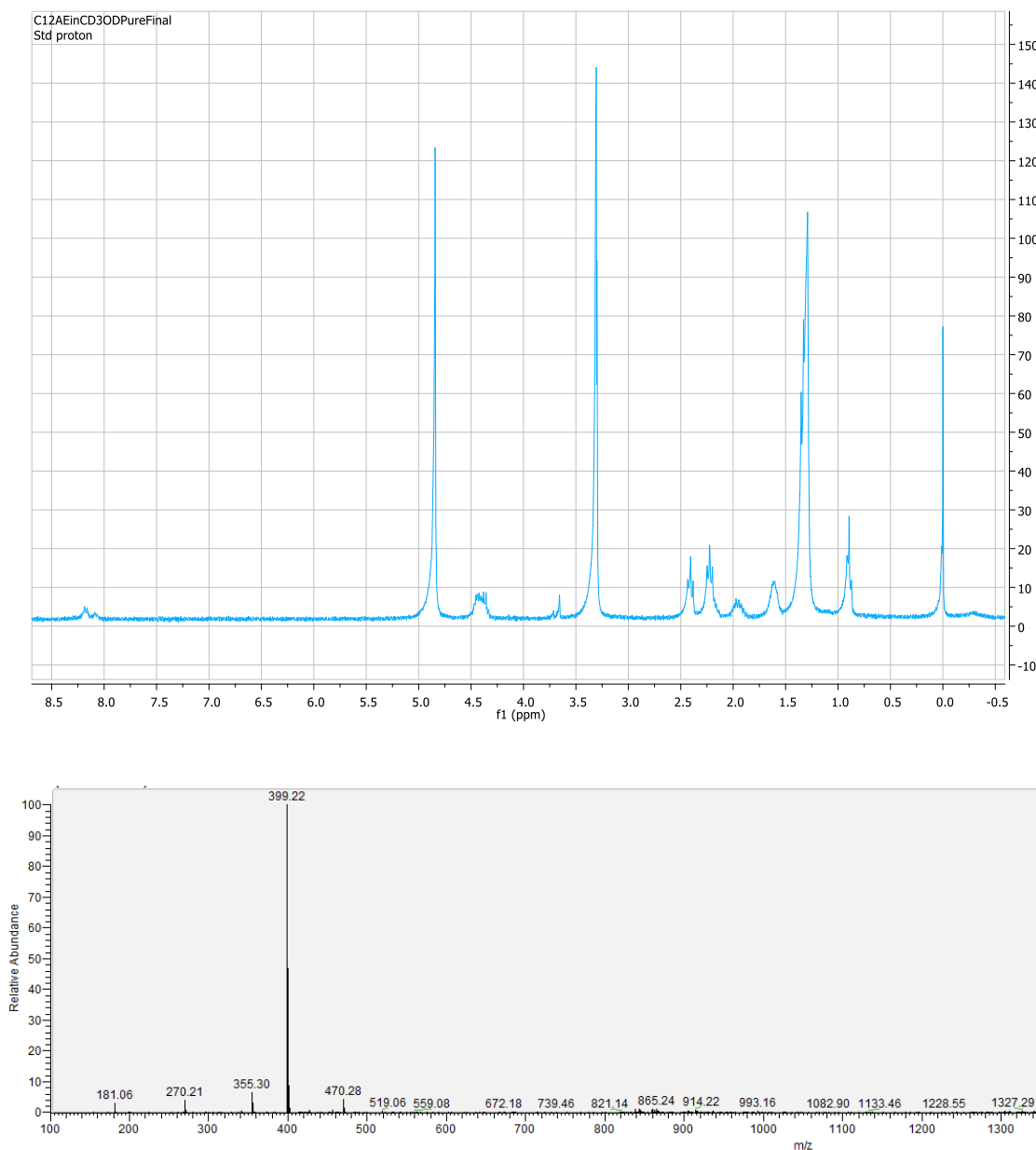


Figure 3S1-A. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for *N*-Lauroylalanyl Glutamic Acid, C_{12}AE .

b. *N*-lauroylglycyl-*L*-glutamic acid. C_{12}GE .

Mw (calc.): 386.5 g mol^{-1} . ^1H NMR (300 MHz, CD_3OD): δ ppm 0.90 (3H, t), 1.29 (16H, bs), 1.62 (2H, m), 1.97 (2H, m), 2.26 (2H, t), 3.88 (2H, d), 4.47 (1H, m), 8.14 (2H, d). HRMS (ESI) m/z (found): $385.85 \text{ g mol}^{-1} (\text{M}^-)$, **Figure 3S1-B.**

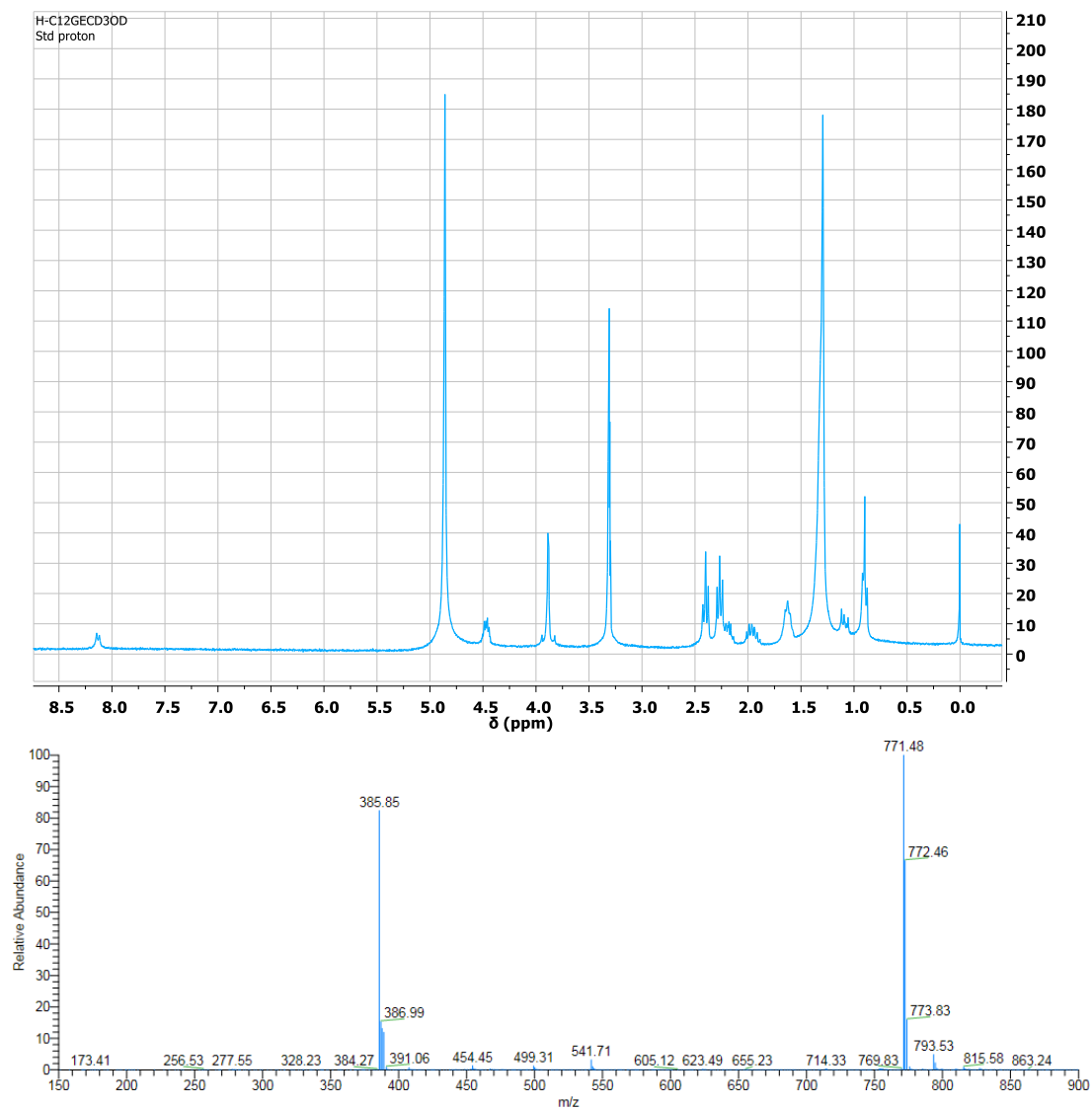


Figure 3S1-B. ¹H-NMR (Top) and ESI-MS (Bottom) Spectrum for *N*-Lauroylglycyl Glutamic Acid, C₁₂GE.

c. *N*-lauroyl-*L*-alanyl-*L*-phenylalanine. C₁₂AF.

Mw (calc.): 418.6 g·mol⁻¹. ¹H NMR (300 MHz, CD₃OD): δ ppm 0.89 (3H, t), 1.29 (16H, bs), 1.58 (5H, m), 2.18 (2H, t), 2.41 (2H, t), 3.00 and 3.19 (2H, m), 4.62 (1H, m), 7.23 (5H, m). HRMS (ESI) *m/z* (found): 417.24 g·mol⁻¹ (M⁺), **Figure 3S1-C**.

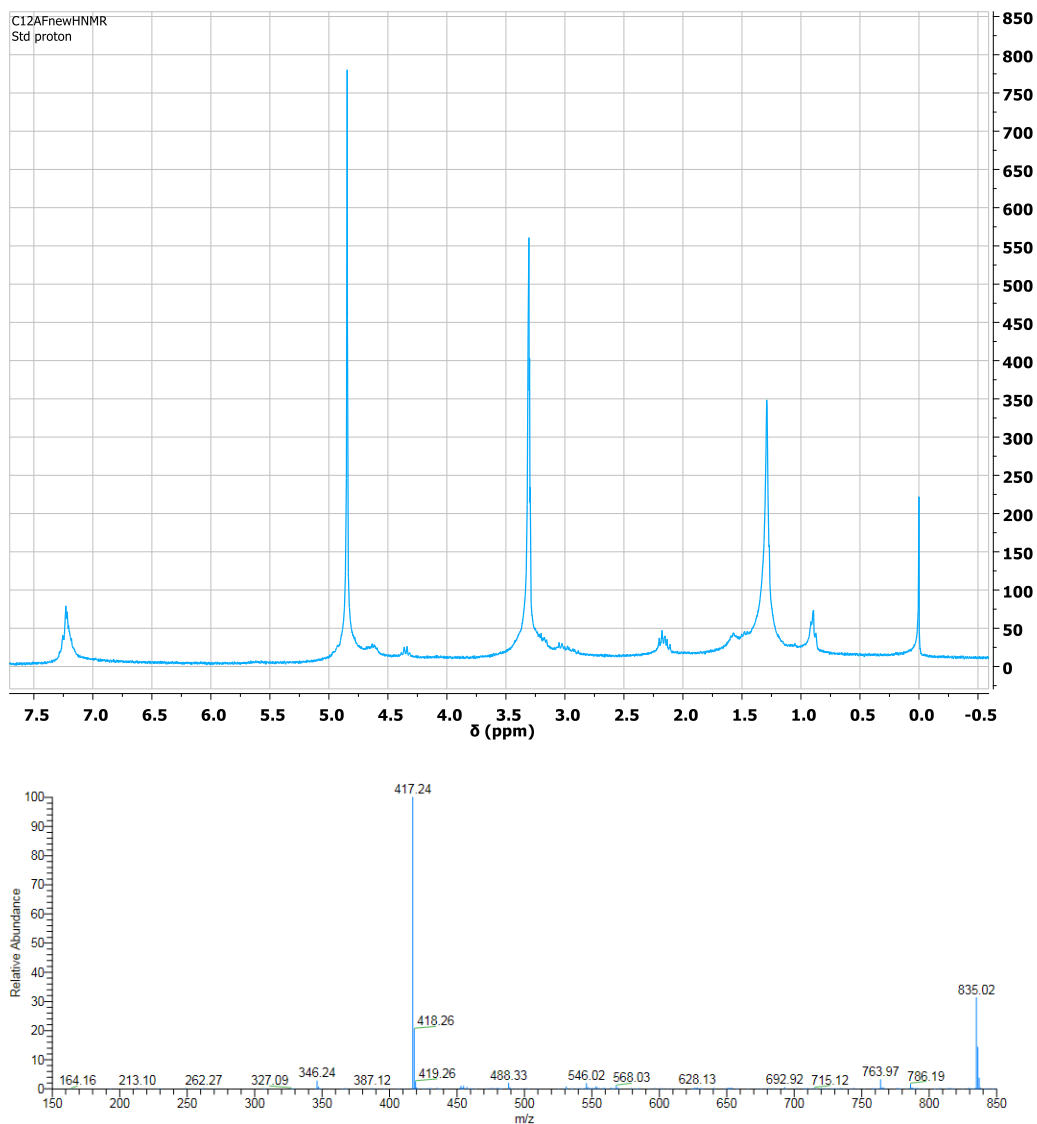


Figure 3S1-C. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for *N*-Lauroylalanyl Phenylalanine, C_{12}AF .

d. *N*-lauroylglycyl-*L*-phenylalanine. C_{12}GF .

Mw (calc.): 404.6 g mol^{-1} . ^1H NMR (300 MHz, CD_3OD): δ ppm 0.89 (3H, t), 1.29 (16H, bs), 1.58 (2H, m), 2.21 (2H, t), 3.03 and 3.16 (2H, m), 3.75 and 3.86 (2H, d), 4.58 (1H, m), 7.23 (5H, m). HRMS (ESI) m/z (found): $403.41 \text{ g mol}^{-1}$ (M^+). **Figure 3S1-D.**

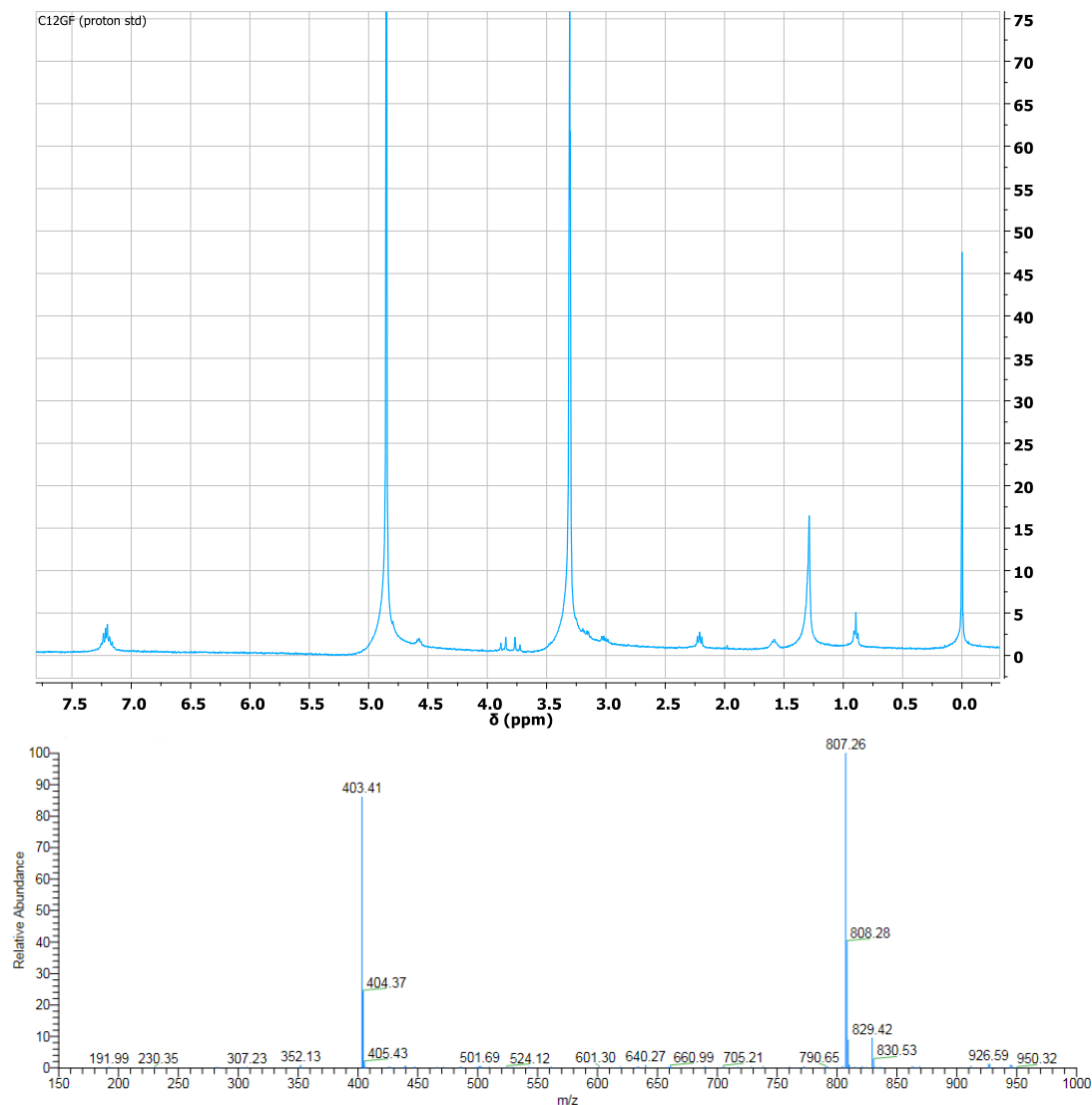


Figure 3S1-D. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for *N*-Lauroylglycyl Phenylalanine, C₁₂GF.

e. N-lauroyl-L-alanyl-L-phenylalanyl-L-glutamyl-L-glutamyl-L-glutamic acid, C₁₂AFEEEE.

Mw (calc.): 805.41 g·mol⁻¹. ^1H NMR (400 MHz, CD₃OD): δ ppm 0.89 (3H, t), 1.27 (16H, bs), 1.54 (5H, m), 1.98 (4H, m), 2.16 (4H, m), 2.33 (2H, t), 2.41 (4H, t), 3.02 and 3.14 (2H, m), 4.24 (2H, m), 4.33 (1H, m), 4.39 (1H, m), 4.55 (1H, m), 7.26 (5H, m). HRMS (ESI) m/z (found): 804.36 g·mol⁻¹ (M⁻). **Figure 3S1-E.**

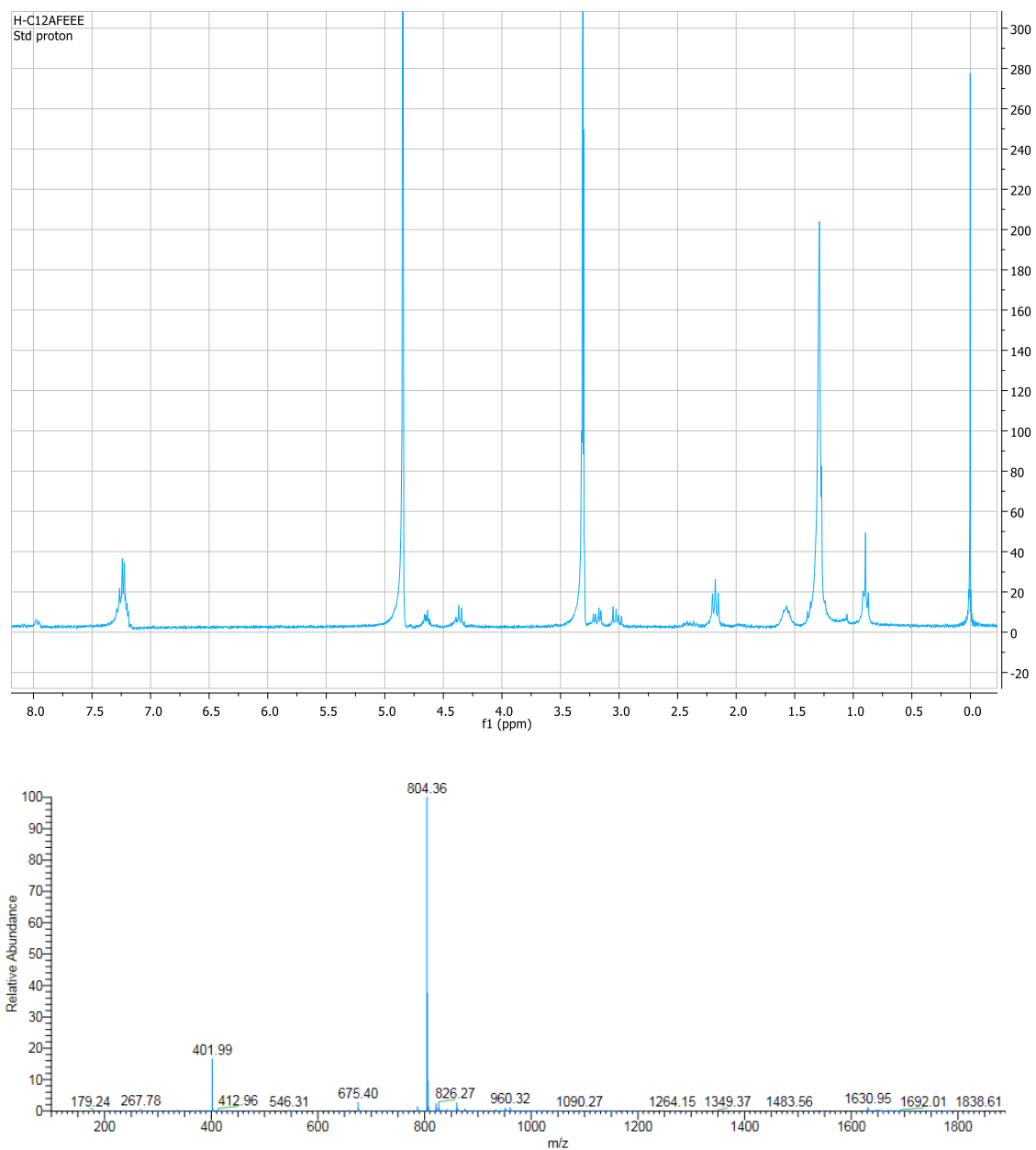


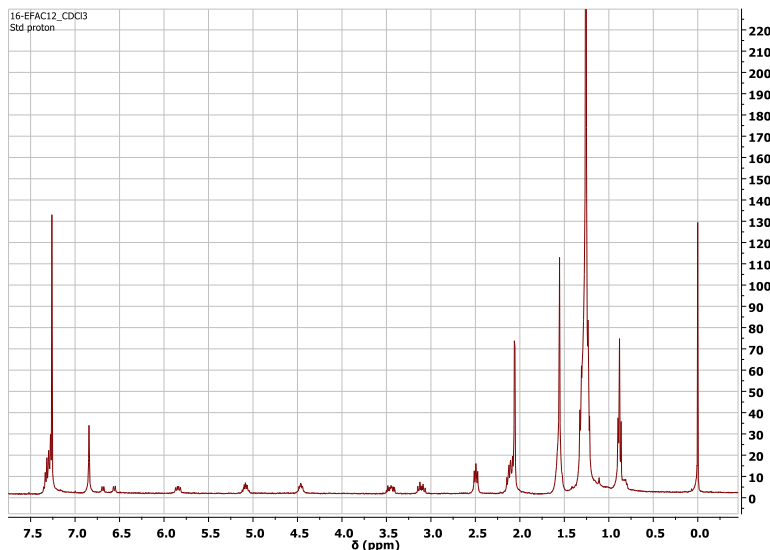
Figure 3S1-E. ¹H-NMR (Top) and ESI-MS (Bottom) Spectrum for *N*-Lauroyl-*L*-Alanyl-*L*-Phenylalanyl-*L*-Glutamyl-*L*-Glutamyl-*L*-Glutamic Acid, C₁₂AFEEE.

Section S2. Syntheses and purification procedures for the preparation of arenediazonium ion, dediazonation products and other related products, and their ^1H NMR and ESI-MS spectra, respectively.

a. Synthetic routes for some of the products, 1- ArN_2^+ , 1- ArOH , 1- ArOAc , 16- ArNH_2 , 16- ArN_2^+ , 16- ArOH , 16- ArNHAc and 16- ArInd , have been published. See Chapter II.

b. 4-*n*-Hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alanyl-*L*-phenylalaninate, 16- ArEFAC_{12} .

DIC (7.0 μL , 1.5 eq.) and DMAP (1.75 mg, 0.5 eq.) were added to a solution of 16- ArOH (10 mg, 0.289 mmol) and *N*-lauroylalanine (25 mg, 2 eq.) in dichloromethane (DCM, 5.0 mL). The mixture was stirred at r.t. overnight, solid removed by filtration and the filtrate evaporated to give a white solid. The solid was dissolved in EtOAc, which was washed successively with saturated Na_2CO_3 , NH_4Cl and NaCl solution, dried over Na_2SO_4 and then with the EtOAc evaporated. 20 mg (80.2%) of white solid was obtained after column chromatography using 10% ethyl acetate/hexanes. Mw (calc.): 747.9 g mol^{-1} . ^1H NMR (300 MHz, CDCl_3): δ ppm 0.88 (6H, t), 1.27 (42H, bs), 1.55 (2H, m), 2.10 (6H, s), 2.15 (2H, t), 2.49 (2H, t), 3.11 and 3.45 (2H, m), 4.47 (1H, m), 5.08 (1H, m), 6.56 (1H, d), 6.69 (1H, d), 6.84 (2H, s), 7.25 (5H, m). HRMS (ESI) m/z (found): 769.35 g mol^{-1} ($[\text{M}+\text{Na}]^+$). **Figure 3S2-B.**



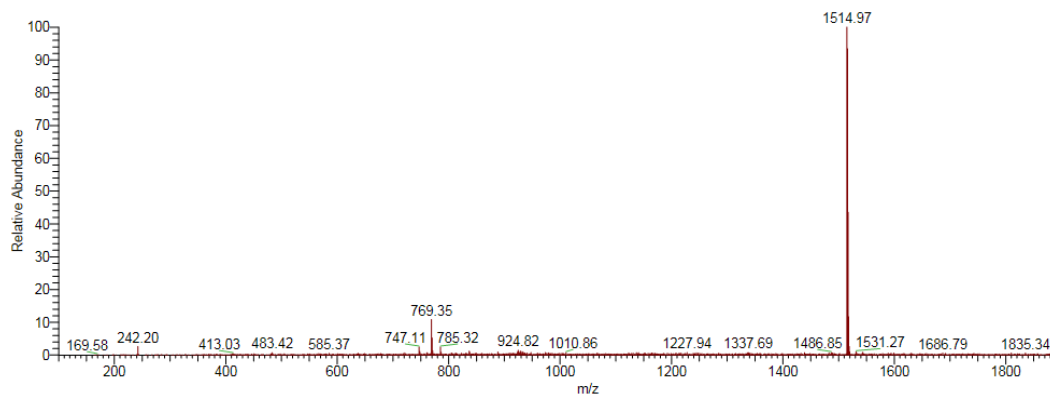


Figure 3S2-B. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for 4-*n*-hexadecyl-2, 6-dimethylphenyl-*N*-lauroylalanyl phenylalaninate, 16-ArEFAC₁₂.

c. 4-*n*-Hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alanyl-*L*-glutamate, 16-ArEEAC₁₂

(alpha ester).

DIC (104 μL , 3 eq.) and DMAP (26.8 mg, 1 eq.) were added to a solution of 16-ArOH (100 mg, 0.289 mmol) and γ -OtBu-*N*-lauroylalanyl glutamic acid (152 mg, 2 eq.) in dichloromethane (DCM, 10 mL). The mixture was stirred at room temperature overnight, solid removed by filtration and the filtrate evaporated. The solid was dissolved in EtOAc, which was washed with NH_4Cl solution twice, dried over Na_2SO_4 and then with the EtOAc evaporated to give 200 mg (88.1%) of solid, Mw (calc.): 729.1 g mol^{-1} . Molecular weight and purity of the resulting intermediate product was checked by ESI-MS. Then 50 mg of the solid was dissolved in a mixture of 35% TFA, 60% DCM, 2.5% TIS and 2.5% H_2O at 0°C , then room temperature overnight. Resulting product was washed with aliquots of hexanes to give white solid. Mw (calc.):

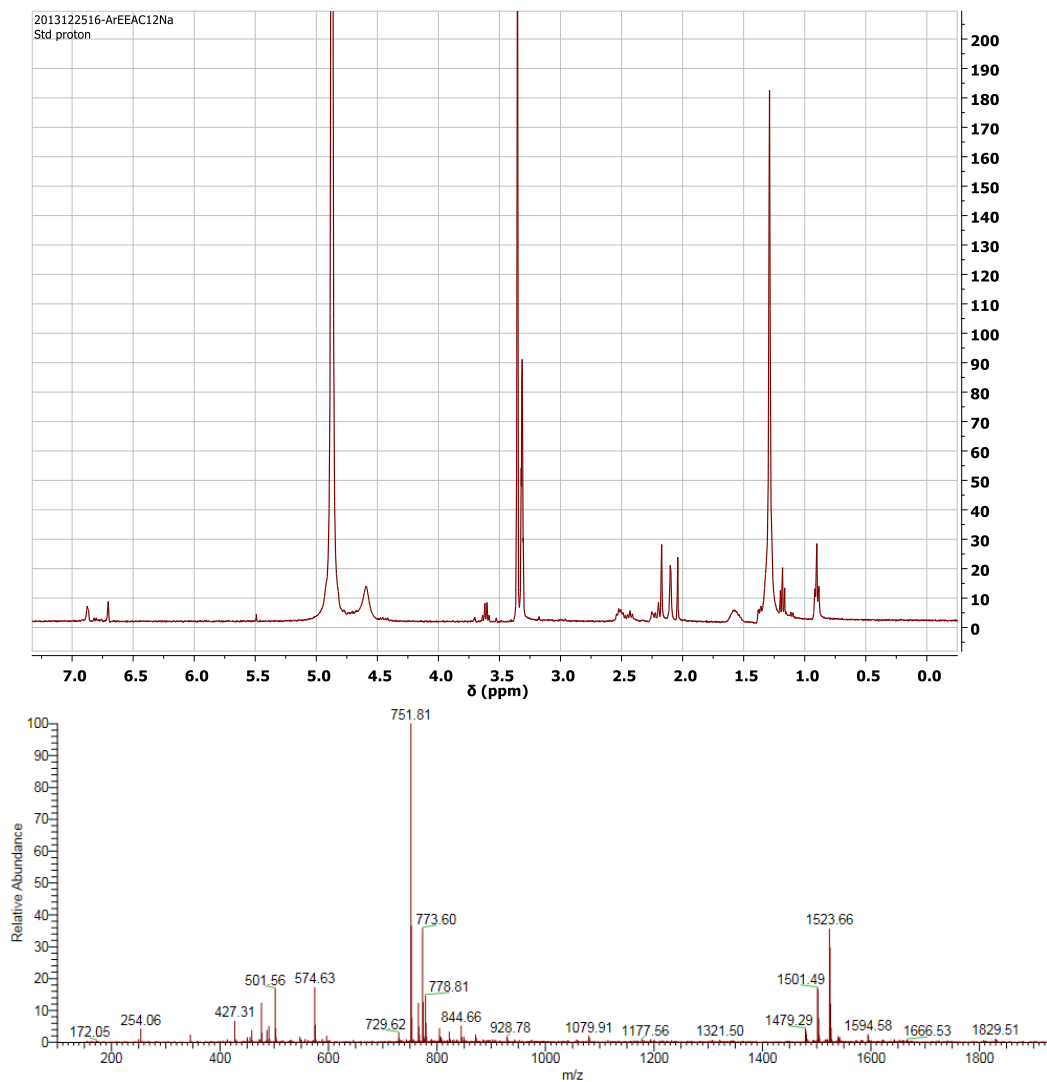


Figure 3S2-C. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for 4-*n*-Hexadecyl-2, 6-dimethylphenyl-*N*-lauroylalanyl glutamate, 16-ArEEAC₁₂ (Alpha Ester).

729.1 $\text{g}\cdot\text{mol}^{-1}$. ^1H NMR (400 MHz, CD_3OD): δ ppm 0.82 (6H, t), 1.24 (42H, bs), 1.51 (7H, m), 2.05 (2H, t), 2.09 (8H, m), 2.14 (2H, t), 2.40 (2H, t), 4.51 (2H, m), 6.62 (2H, s), 6.80 (2H, b). HRMS (ESI) m/z (found): 751.81 $\text{g}\cdot\text{mol}^{-1}$ ($[\text{M}+\text{Na}]^+$). **Figure 2S2-C.**

d. 4-*n*-Hexadecyl-2, 6-dimethylphenyl-*N*-lauroyl-*L*-alanyl-*L*-glutamate, 16-ArEEAC₁₂ (gamma ester).

DIC (104 μL , 3 eq.) and DMAP (26.8 mg, 1 eq.) were added to a solution of 16-ArOH (100 mg, 0.289 mmol) and γ -OtBu-*N*-lauroylalanyl glutamic acid (152 mg, 2 eq.) in dichloromethane

(DCM, 10 mL). The mixture was stirred at room temperature overnight, solid removed by filtration and the filtrate evaporated. The solid was dissolved in EtOAc, which was washed with NH_4Cl solution twice, dried over Na_2SO_4 and then with the EtOAc evaporated to give 200 mg (88.1%) of solid, Mw (calc.): 785.2 g mol^{-1} . Molecular weight and purity of the resulting intermediate product was checked by ESI-MS. Then 50 mg of the solid was dissolved in a mixture of 35% TFA, 60% DCM, 2.5% TIS and 2.5% H_2O at 0°C , then room temperature overnight. Resulting product was washed with aliquots of hexanes to give white solid. Mw (calc.): 729.1 g mol^{-1} . HRMS (ESI) m/z (found): $751.56 \text{ g mol}^{-1}$ ($[\text{M}+\text{Na}]^+$). **Figure 3S2-D.**

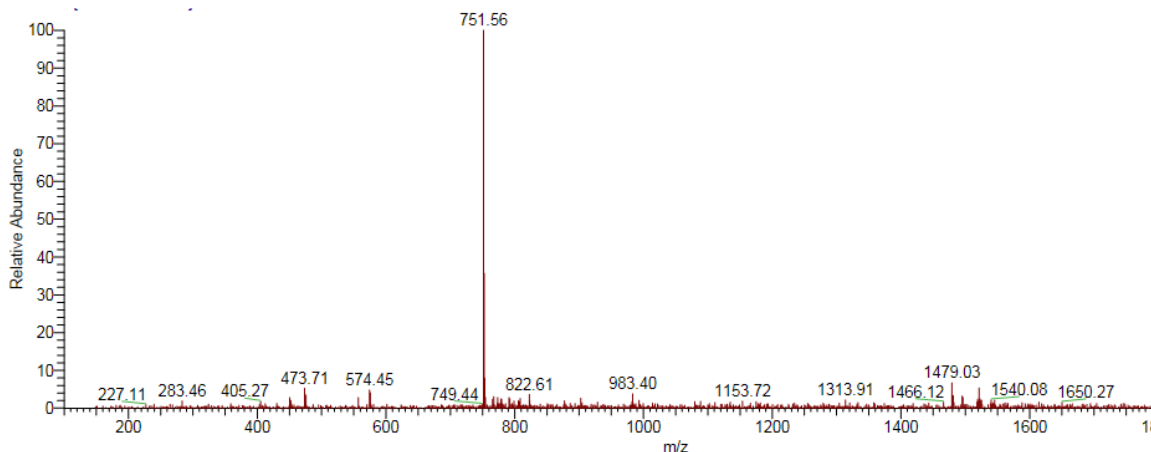


Figure 3S2-D. ESI-MS Spectrum for 4-*n*-Hexadecyl-2,6-dimethylphenyl-*N*-lauroylalanyl glutamate, 16-ArEEAC₁₂ (gamma ester).

e. 2, 4, 6-Trimethylphenyl-*N*-acetyl glycine, 1-ArOGAc.

DIC (1.71 mL, 1.5 eq.) and DMAP (450 mg, 0.5 eq.) were added to a solution of 1-ArOH (1.0 g, 1 eq.) and *N*-acetyl glycine (1.72 g, 2 eq.) in dichloromethane (DCM, 10 mL). The mixture was stirred at r.t. overnight, solid removed by filtration and the filtrate evaporated to give an off-white solid. The solid was dissolved in EtOAc, which was washed successively with saturated Na_2CO_3 and NaCl solution, dried over Na_2SO_4 and then with the EtOAc evaporated. Product was purified by column chromatography using 50% ethyl acetate/hexanes, and recrystallized by dissolving in hot methanol and cool down to room temperature then 0°C for 3 hours. 1.7 g (98.3%) of white

crystal was obtained. Mw (calc.): 235.3 g·mol⁻¹. ¹H NMR (300 MHz, CD₃OD): δ ppm 2.02 (3H, s), 2.08 (6H, s), 2.25 (3H, s), 4.21 (2H, d), 6.87 (2H, s), NH not observed. See ¹H NMR spectrum,

Figure 3S2-E.

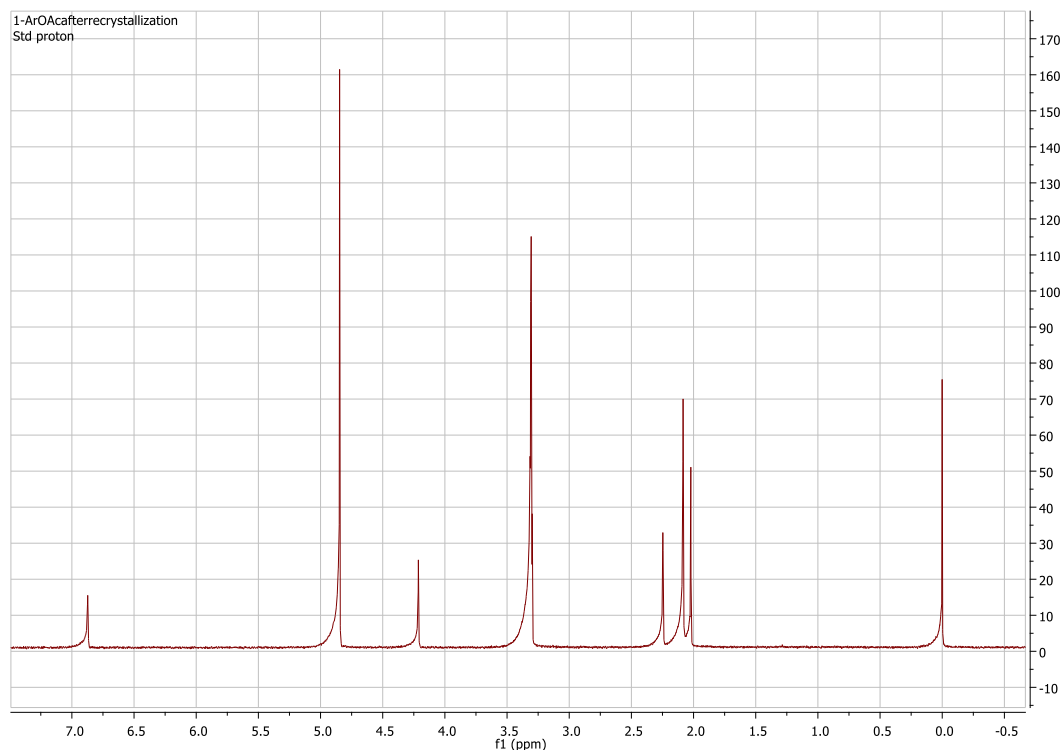


Figure 3S2-E. ¹H-NMR (Top) and ESI-MS (Bottom) Spectrum for 2,4,6-trimethylphenyl-*N*-acetyl glycine, 1-ArOGAc.

f. 2, 4, 6-Trimethylphenyl-*N*-acetyl glycylglycine, 1-ArOGGAc.

DIC (137 μL, 1.5 eq.) and DMAP (36 mg, 0.5 eq.) were added to a solution of 16-ArOH (80 mg, 0.289 mmol) and *N*-acetylglycylglycine (102.3 mg, 2 eq.) in dichloromethane (DCM, 10 mL). The mixture was stirred at r.t. overnight, solid removed by filtration and the filtrate evaporated to give a white solid. The solid was dissolved in EtOAc, which was washed successively with saturated Na₂CO₃, NH₄Cl and NaCl solution, dried over Na₂SO₄ and then with the EtOAc evaporated. 160 mg (93.2%) of white solid was obtained after column chromatography using 10% ethyl acetate/hexanes. Mw (calc.): 292.3 g·mol⁻¹. ¹H NMR (300 MHz, CDCl₃): δ ppm 2.05 (3H, s),

2.10 (6H, s), 2.24 (3H, s), 4.01 (2H, d), 4.34 (2H, d), 6.23 (1H, t), 6.61 (1H, t), 6.87 (s, 2H). See ^1H NMR spectrum, **Figure 3S2-F**.

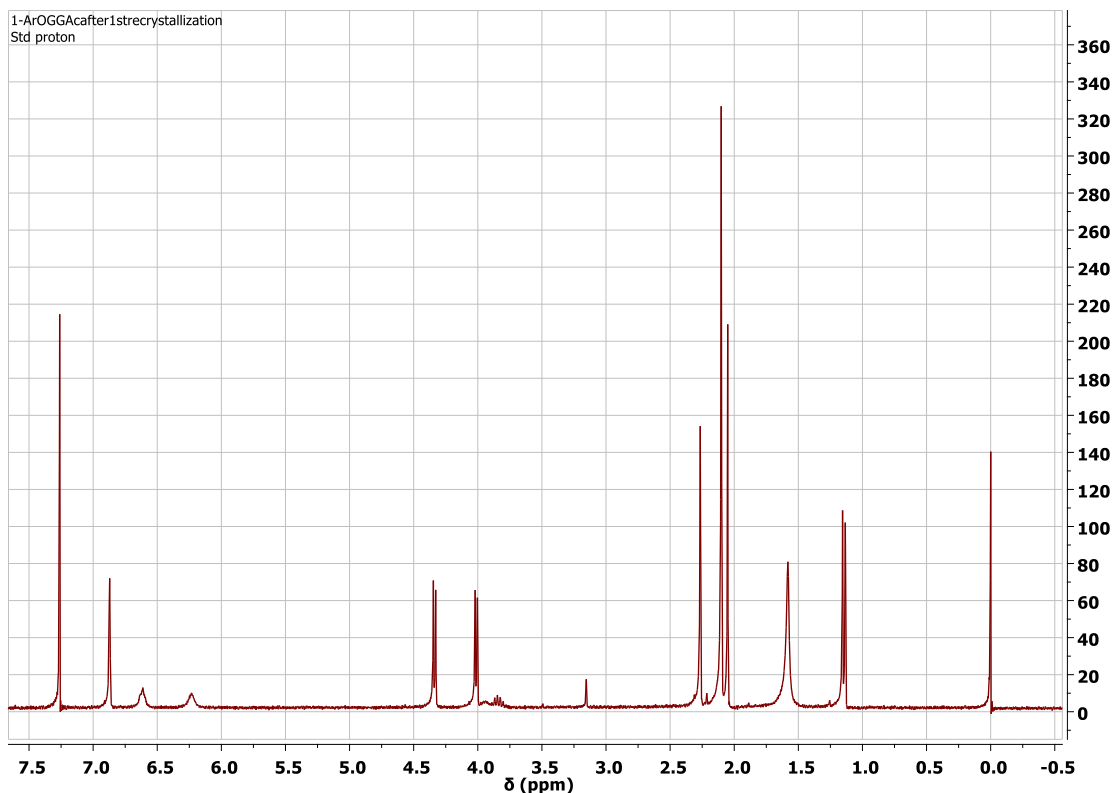


Figure 3S2-F. ^1H -NMR (Top) and ESI-MS (Bottom) Spectrum for 2, 4, 6-trimethylphenyl-*N*-acetyl glycylglycine, 1-ArOGGAc.

Section S3. Protocol and conditions of the synthesized chemical imaging product characterization via ESI-MS and MALDI-TOF MS.

ESI-MS. The purity of synthesized peptide amphiphiles and some of the key products and byproduct was checked with a quadrupole ion trap (LCQ) mass spectrometer. The electrospray needle voltage was set at -4.0 kV. The capillary temperature was set at 250°C. Typically, an 1×10^{-5} to 5×10^{-5} M solution of each peptide amphiphile or product was prepared in a 50% IPA – 50% MeOH solution with either 0.1% v/v formic acid or ammonium hydroxide as the additive. Positive or negative mode was used for mass detection. Note that some side products such as 16-

ArH, 16-ArF and 16-ArCl were not detectable under the optimized condition. Mass spectra were summarized in Section *S1* and *S2* of this chapter and also *Chapter II*.

MALDI-TOF mass spectrometry. Molecular mass to charge ratio of some of the key chemical imaging products were obtained with a MALDI-TOF/TOF mass spectrometer. 2,5-Dihydroxyacetophenone (DHAP) was selected and used as the matrix compound. Ions were produced by irradiation of the sample with nitrogen laser (337 nm). Profiling of product ions was achieved in the reflector of positive mode. The accelerating voltage was set at 20 kV. A sum of 50 shots was collected for each spectrum.

With a careful and systematic optimization, 2, 6-dihydroxyacetone phosphate was selected as the most ideal matrix compound among a series of candidate reagents. MALDI analysis samples were prepared as follows: 7.6 mg of 2, 5-DHAP was dissolved in 375 μ L ethanol. Then 125 μ L of an 18 mg/mL aqueous solution of diamonium hydrogen citrate was added to obtain a mixed matrix solution. Samples from chemical imaging reactions were initially separated by HPLC. Fractions were serially collected and concentrated at a concentration range of ca. 10^{-4} molar. 2 μ L of this sample solution with 2 μ L 2% TFA in either aqueous or methanol solution, and 2 μ L of the above freshly prepared matrix solution were added. 0.85 μ L of the crystal suspension was deposited on the MALDI target and air dry.

MALDI mass spectra were obtained with Ab Sciex 4800. Ions were produced by irradiation of the sample with nitrogen laser (337 nm). Profiling of product ions was achieved in the reflector positive mode. The accelerating voltage was 20 kV. The sum of 50 shots was collected for each spectrum.

Section S4. Calibration curves for reaction products.

Table 3S4-1. Equations Used to Fit HPLC Calibration Curves of the Relevant Key Products.^a

Reaction Product	Calibration Equation ^b	R ²
16-ArOH	$y=9.998 \times 10^{10} x$	0.9998
16-ArGE	$y=1.090 \times 10^{11} x$	1.0000
16-ArEC₁₂	$y=9.997 \times 10^{10} x$	0.9945
16-ArInd	$y=1.404 \times 10^{11} x$	0.9980
16-ArAE	$y=1.199 \times 10^{11} x$	1.0000
16-ArEFAC₁₂	$y=1.012 \times 10^{11} x$	0.9960
16-ArEEAC₁₂ (alpha)	$y=1.102 \times 10^{11} x$	0.9990
16-ArEEAC₁₂ (gamma)	$y=1.088 \times 10^{11} x$	0.9940
1-ArOH	$y=6.155 \times 10^{10} x$	1.0000
1-ArOAc	$y=5.745 \times 10^{10} x$	1.0000
1-ArOGAc	$y=4.353 \times 10^{10} x$	0.9994
1-ArOGGAc	$y=4.656 \times 10^{10} x$	0.9989

a. HPLC eluting solvent for 16-ArX: *i*-PrOH/MeOH, 35%/65% (v/v), or 45%/55% (v/v). Flow rate: 0.40 mL/min. For 16-ArEEAC₁₂: MeOH 100%, 0.05% v/v TFA, 0.2 mL/min. For 1-ArX: MeOH/H₂O, 80%/20% (v/v), 0.60 mL/min. All calibration curves were made at an injection volume of 100 μ L, reflecting peak area – conc. relationship at 0.4 mL/min, and at a wavelength of 220 nm UV detection.

b. Units: y-peak area (μ V's), x-concentration (molarity), and R² (correlation coefficient). The y intercept values are very small and not used in the calculations.

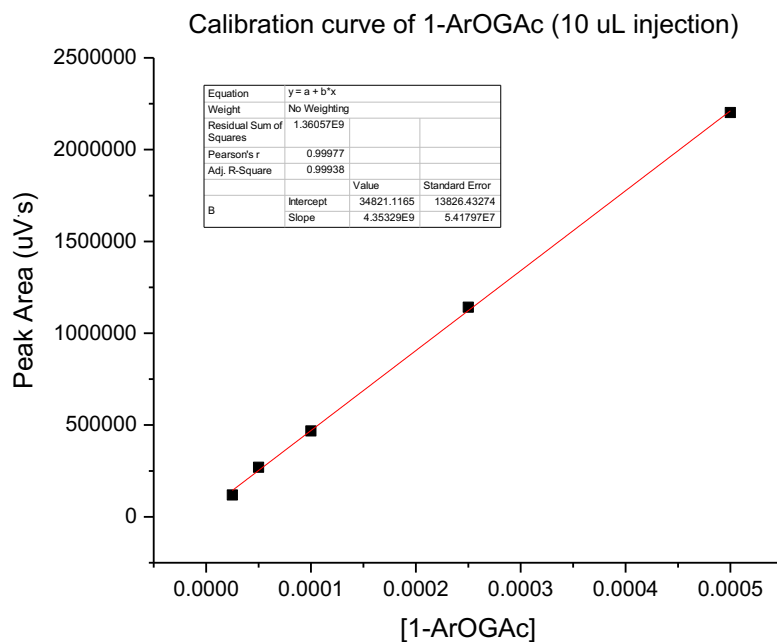


Figure 3S4-1. Calibration Curve for 1-ArOGAc.

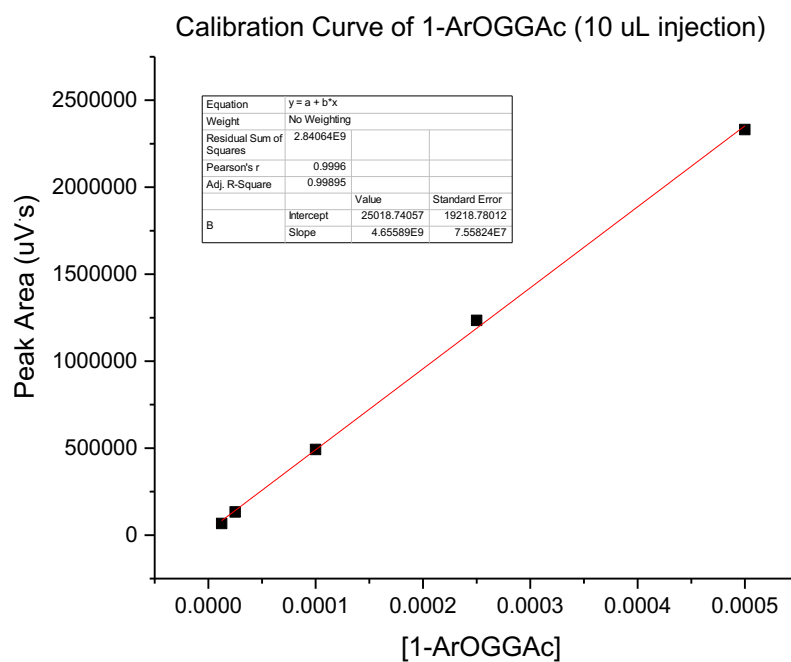


Figure 3S4-2. Calibration Curve for 1-ArOGGAc.

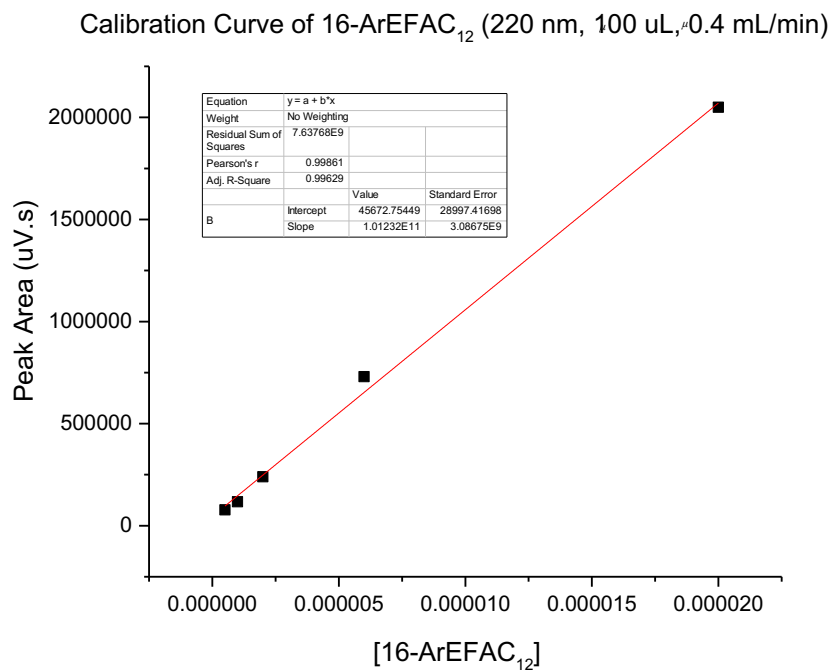


Figure 3S4-3. Calibration Curve for 16-ArEFAC₁₂.

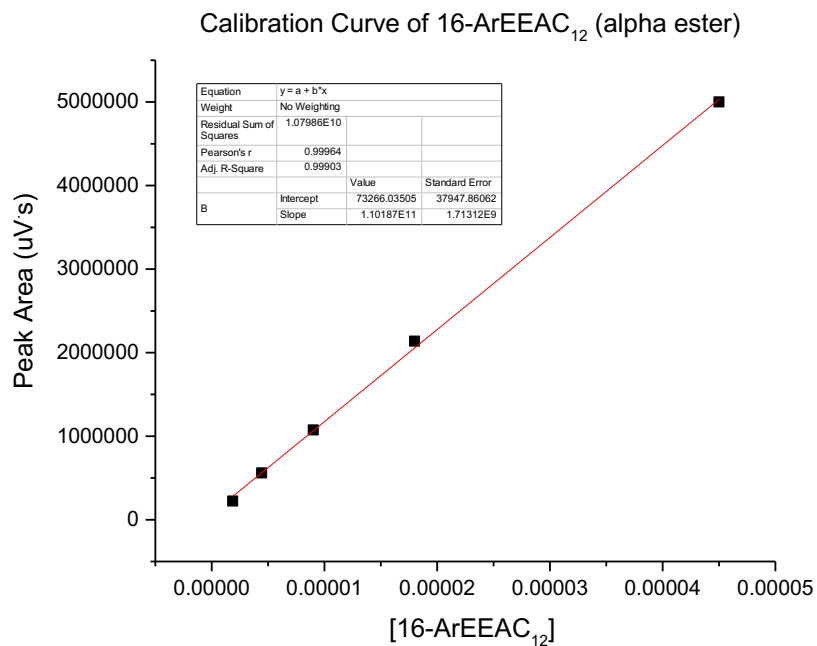


Figure 3S4-4. Calibration Curve for 16-ArEEAC₁₂ Alpha Ester.

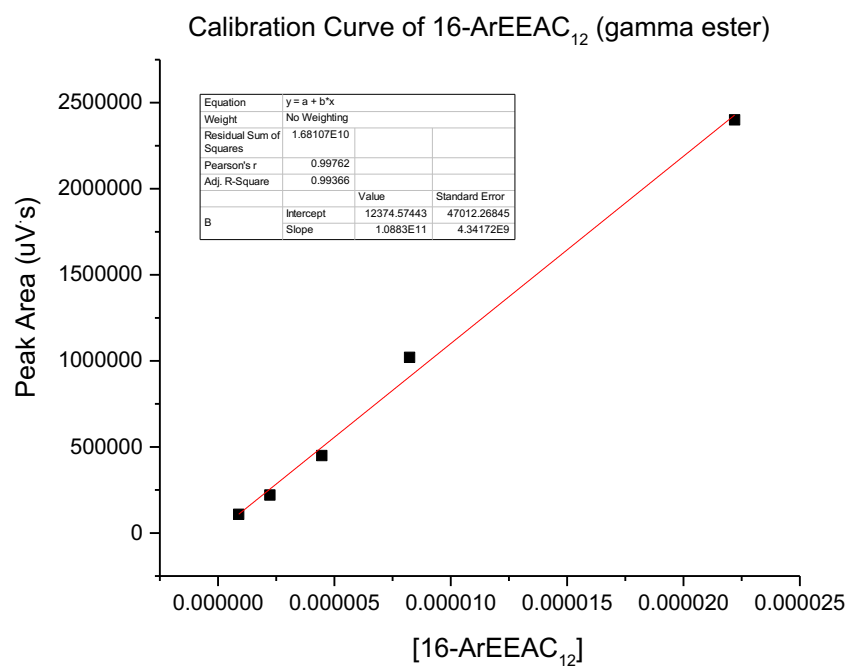


Figure 3S4-5. Calibration Curve for 16-ArEEAC₁₂ Gamma Ester.

Section S5. Analyses of Dediazonation Kinetic Data.

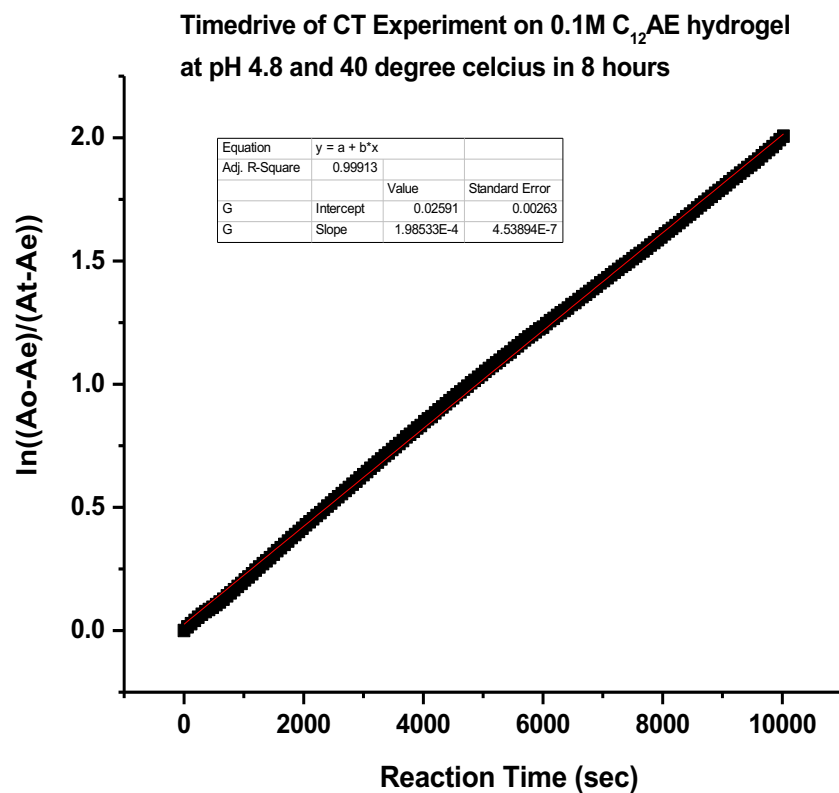


Figure 3S5-1. Ln Plot of UV Absorbance for the Dediazonation of 6×10^{-4} M 16-ArN₂⁺ in 0.098 M C₁₂Ala-Glu at pH 4.8 at 40°C. S Stands for Time in Seconds.

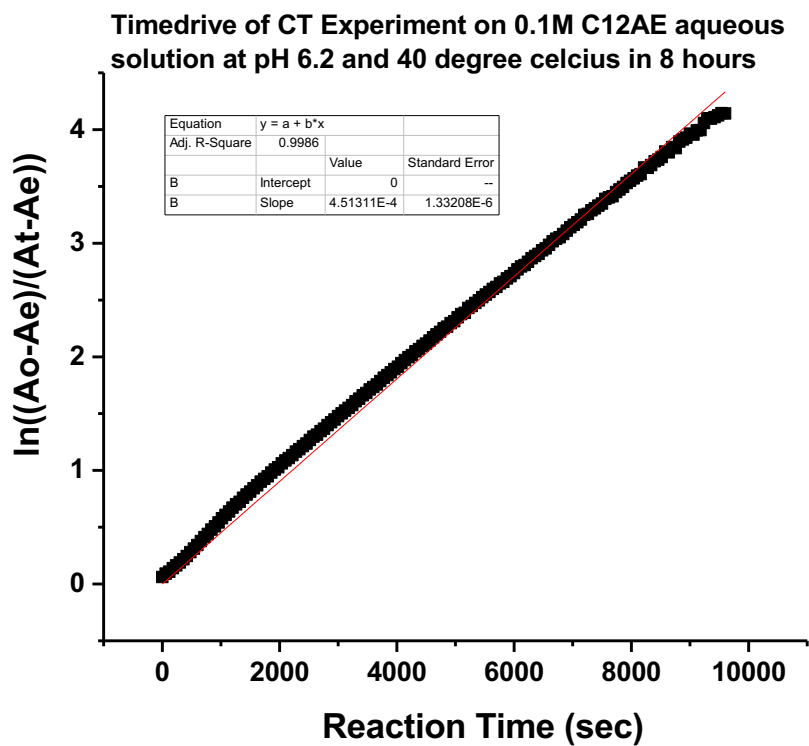


Figure 3S5-2. Ln Plot of UV Absorbance for the Dediazonation of 6×10^{-4} M 16-ArN₂⁺ in 0.098 M C₁₂Ala-Glu at pH 6.2 at 40°C. S Stands for Time in Seconds.

Section S6. Supplemental information of HPLC observed and normalized percent yields of reactions of 16-ArN_2^+ with dipeptide amphiphile aqueous solutions.

Table 3S6-A. Chemical Fingerprint* of $\text{C}_{12}\text{Ala-Phe}$ at a Concentration of 80 mM at pH 5 - 6 at 40°C (Gel).

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	72.1	90.14
16-ArEAC₁₂	19	0.23	1.12
16-ArEFAC₁₂	20	5.69	7.25
16-ArEC₁₂	37-38	0.38	1.49

[16-ArN_2^+] = $4\text{e-}4$ M, %Total = 99.7% (10 μL).

*35% IPA, 0.4 mL min^{-1}

Table 3S6-B. Chemical Fingerprint* of $\text{C}_{12}\text{Ala-Phe}$ at a Concentration of 100 mM at pH below 6.5 (rt: Gel, 48 hrs, 10 μL injection).

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	55.6	92.12
16-ArEAC₁₂	19	0.18	0.93
16-ArEFAC₁₂	20	3.41	5.74
16-ArEC₁₂	37-38	0.23	1.21

%Total = 99.2% (10 μL)

35% IPA, 0.4 mL min^{-1}

Table 3S6-C. Chemical Fingerprint* of $\text{C}_{12}\text{Gly-Glu}$ at a Concentration of 100 mM at pH 4.5 (rt, solution).

Abbreviation	Retention Time (min)	Percent Yield (%) [*]	Normalized Yield (%) [*]
16-ArOH	14-15	65.0	80.06
16-ArGE	19	0.10	0.40
16-ArEEGC₁₂ (alpha)	55**	2.78	3.79
16-ArEEGC₁₂ (gamma)	62**	2.68	3.65
16-ArEC₁₂	37-38	2.86 (1.9 [^])	12.10

[16-ArN_2^+] = $8\text{e-}4$ M, %Total = 96.4% (81.7% when 50 μL injection, no TFA buffer).

*35%IPA, 0.4 mL min^{-1} , 0.05% (v/v) TFA.

**0 or 10%IPA, 0.2 mL min^{-1} , 0.05% (v/v) TFA.

[^]Solution pH 5.0 at 40°C 12 hr.

Section S7. Gelation properties of C₁₂GE, C₁₂AE and C₁₂AF.

Three synthesized PAs undergo hydrogelation as the solution pH is reduced. Brief descriptions of their gelation properties are summarized below.

C₁₂GE. 100 μ L of C₁₂GE aqueous solution at 100 mM and pH 6.0 with no added salt was prepared at room temperature. Aliquots of 1 M HCl were added to determine critical gelation pH at room temperature. Solution became viscous below pH 5 and formed a gel that was stuck at the bottom of a glass vial at pH 4.6. The gel is homogenous but a bit opaque. The critical gelation pH dropped below 4.0 at 40°C and became viscous at ca. pH 4.

A sample at a gelation pH of 4.5, was warmed slowly from room temperature. A gel was obtained below 30°C and started to become more fluid-like but remained viscous at a temperature range of 30 - 37°C. The gel sample became solution at ca. 37°C, but was a bit viscous even at 41°C. At 50 mM, a slightly opaque solution was obtained.

C₁₂AE. 100 μ L of C₁₂AE aqueous solution at 100 mM and pH 6.0 with no added salt was prepared at room temperature. Aliquots of 1 M HCl were added to determine critical gelation pH at room temperature. The solution became viscous below pH 5 and formed a gel that stuck at the bottom of a glass vial at pH 4.5. The gel is homogenous and a bit opaque. The critical gelation pH dropped below 4.0 at 40°C and became viscous at ca. pH 4. At a gelation pH of 4.5, sample was warmed slowly from the room temperature. Gel was obtained below 30°C and started to become more fluid-like but viscous at a temperature range of 30 - 37°C. The gel sample was converted to clear solution at ca. 35 - 36°C. The critical gelation concentration at a solution pH of 4.0 and room temperature was determined to be 20 - 25 mM. At 20 mM, pH 4.0 and room temperature, only a highly viscous solution was observed. The viscosity became much higher and gel-like by eye observation after 20 mM NaCl was added. The same phenomenon was observed by adding MgSO₄, but addition of CaCl₂ led to precipitation. On the other hand, at 10 mM of C₁₂AE at pH

4.0 and room temperature, the addition of NaCl, CaCl_2 and MgSO_4 (all at 20 mM) all induce the precipitation of PA samples at various degrees, but with NaCl closer to a gel sample.

C_{12}AF : 100 μL of C_{12}AF aqueous solution at 100 mM with no added salt was prepared at room temperature. Aliquots of 1 M HCl were added to determine critical gelation pH at room temperature. Solution became white suspension that was opaque and more viscous at or below pH 7.4, and formed a gel that stuck at the bottom of a glass vial at pH 6.5. The gel is homogenous and a bit opaque with a higher turbidity than those of C_{12}GE and C_{12}AE . At 40°C, the solution started to become opaque and viscous at a critical pH of 6.9 while the critical gelation pH was not observed. Similar phenomenon was not observed at a concentration of 50 mM. Instead, a suspension of fibril-like structure was observed at pH 6.5 and room temperature. The sample became clear solution at 40°C and white, flat but needle-like crystals formed slowly on cooling to room temperature. Rewarming the crystal suspension at 40°C does not make the crystals disappear while vigorous stirring promotes their partial dissolution.



Figure 3S8-1. Appearance of Six PA Samples at Various Solution pHs. **Vial #1-3:** 80 mM $\text{C}_{12}\text{Ala-Phe}$ at pH 6.7 (#1, cloudy solution), pH 6.0 (#2, gel) and pH 5.5 (#3, gel). **Vial #4-6:** 100 mM of $\text{C}_{12}\text{Gly-Glu}$ at pH 6.0 (#4, solution), pH 5.0 (#5, solution) and pH 4.5 (#6, gel).

Chapter IV. Interpretation of the Minimum at Surface Tension Curves of Aqueous *N*-lauroyl Sarcosinate Solutions.

In this chapter, a substantial surface tension minimum was demonstrated with pure *N*-lauroyl sarcosinate (LS) aqueous solutions. A comprehensive study is summarized, including surface tensiometry, ^1H -NMR, density functional theory (DFT) calculation, acid-base titrations and FT-IR, *etc.*, to provide novel perspective on its appearance. Surface tension minimum disappears at a fairly high pH of 11, as well as at a moderate pH of 6 with concentrated buffers. NMR results indicate the restricted rotation of amide bonds that produces *E* and *Z* rotamers, and the higher stability of *Z* rotamer at micellar interface as well as in the form of free acid. A substantial discrepancy between *Z* and *E* rotamers was found in terms of the change of the chemical shift of protons at the methylene adjacent to the carboxylate group at various aqueous solution concentrations, pHs and solvent types, respectively. The evidence also indicates the possible existence of an intra-molecular hydrogen bond for protonated *Z*-rotamer. More *Z*-rotamer is in micelles that functions as a proton “reservoir” and gives rise to a surface tension minimum as the micellar solution is diluted.

4.1. Introduction

Surface tension measurements are commonly used to determine the critical micelle concentration, cmc, of a surfactant. But they also provide an assessment of surfactant purity because the measurement is remarkably sensitive to the presence of small amounts of surface-active impurities that produce minima of various sizes in surface tension/surfactant concentration plots,¹ e.g., **Figure 4.1A**. In the absence of additives and impurities, surface tension measurements of surfactant solutions provide characteristic and reproducible estimates of the change in the net surface force at the air/water interface with increasing surfactant concentration.² Typical surface tension profiles of a pure surfactant versus [surfactant] are shown in **Figure 4.1B**,

where square brackets indicate molarity in moles per liter of solution volume. Basically, the profile shows a rapid, sometimes linear decrease in surface tension with increasing surfactant concentration followed by a sharp change in slope to plateau generally with a slope near zero, **Figure 4.1B**. The break point is an empirical definition of the cmc in the absence of a minimum¹ and its value is assumed to mark the onset of the spontaneous self-assembly of monomers into typically spherical micelles to make homogeneous, optically transparent solutions.² The monomers in the micelles and the surrounding aqueous region exchange extremely rapidly, near the diffusion-controlled limit, and the ensemble of micelles and monomers are in dynamic equilibrium between each other and with the air/water interface after bulk mixing is complete. For example, the entrance and exit rate constants for association of sodium dodecyl sulfate monomer with micelles are, respectively, $1.2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ ($1.2 \times 10^6 \text{ s}^{-1}$ in 10 mM solution of micelles) and $1 \times 10^7 \text{ s}^{-1}$ and are probably similar at the air/water interface² (see Table 2.5, p. 55). This exchange is on the order of a million times faster than *E* to *Z* amide bond rotation in *N*-lauroylsarcosine micelles (presented below). Cmc values may be method sensitive,¹ and differ somewhat, but the differences for pure surfactant are seldom large. However, the presence of other surface active compounds like oils, organic solvents or an incompletely removed hydrophobic starting material, such as the addition of small amounts of dodecanol, ca. 0.5% mole percent by weight of sodium dodecyl sulfate or less, produce easily observable minima.³ Even a fingerprint that touches the surface of the glass dish or the platinum ring in the Du Noüy tensiometer may reduce the surface tension significantly.

Surface tension minimum = hydrophobic impurity is, for NaLS, a “red herring.” *N*-methylated sodium *N*-lauroyl (dodecanoyl) sarcosinate, NaLS, an unnatural amino acid-based anionic surfactant, is one of a growing class of amino acid-based surfactants currently being studied because of their wide scientific and commercial utility.⁴ *N*-methylation of amides is also being explored as a new approach in drug design, because methylation modulates protein/peptide

biological functions and brings about enhanced resistance to proteolysis in chemical and biological systems due to the reduced flexibility of the amide bond.⁵ However, NaLS has an annoying property, a minimum in its surface tension profile at ambient pH that screams hydrophobic impurity.

Our results show that the minima are one of the characteristics of NaLS, and MLS (M = tetramethylammonium or TMA⁺, Cs⁺, and Li⁺), **Figures 4.1** and **4.2**, at ambient pH in the absence of buffers and added salts. However, at elevated pH or high salt or buffer concentrations (≥ 0.1 M) with the same surfactant, the minima disappear. Despite our repeated recrystallizations at ambient pH, the minimum remained—it did not even diminish, unlike that of sodium *N*-lauroyl glycinate, NaLG, which is structurally identical to NaLS except that a hydrogen, CON-H replaces the *N*-methyl group of NaLS, and shows no minimum in its surface tension profile after recrystallization.⁶

Z and E rotamers of the NaLS and MLS surfactants. *N*-methylation, or more generally *N*-alkylation, induces a conformational isomerism in amide or peptide bonds, i.e., two different conformational isomers that rotate between *Z* and *E* forms sufficiently slowly that they appear as two separate signals (see below), **Scheme 4.1A** and **Scheme 4.2**. Several studies have demonstrated the presence of rotamers of *N*-substituted acetamides,⁷ amino acid-based surfactants,⁸⁻¹⁴ as in **Scheme 4.1**, and peptides.⁵ **Chart 4.1** and **Scheme 4.1** show the structures of the *Z* (*trans*) and *E* (*cis*) rotamers of NaLS and its three sets of coupled protons in its ¹H NMR spectrum produced by restricted rotation around the C-N bond.^{9,12,13} (Example ¹H NMR spectra are shown in **SI, Figures 4S1a-h**. The spectra illustrate that both protonated, and deprotonated surfactant give three pairs of signals for the *Z* and *E* rotamers, in water, micelles and other solvents with different chemical shifts, *Z*:*E* ratios, and signal splittings. **Scheme 4.2** illustrates the various forms of *N*-lauroylsarcosine that may be present at the air/water interface, the interfacial

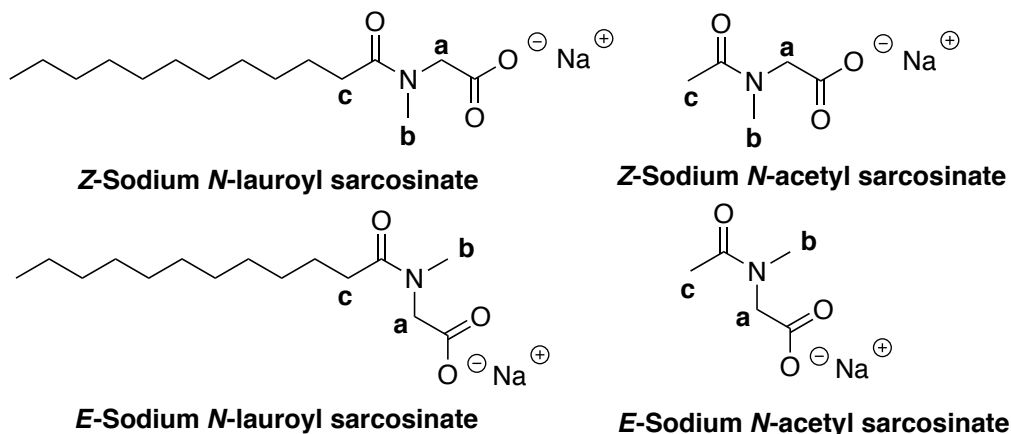
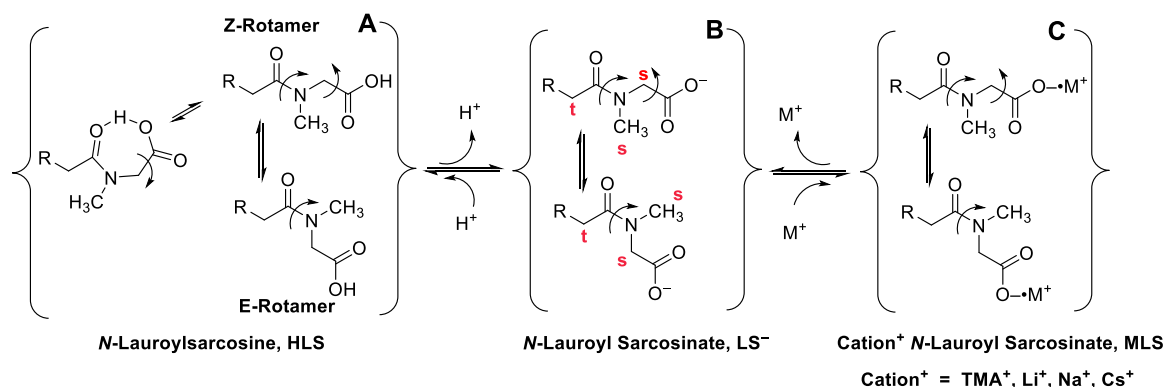


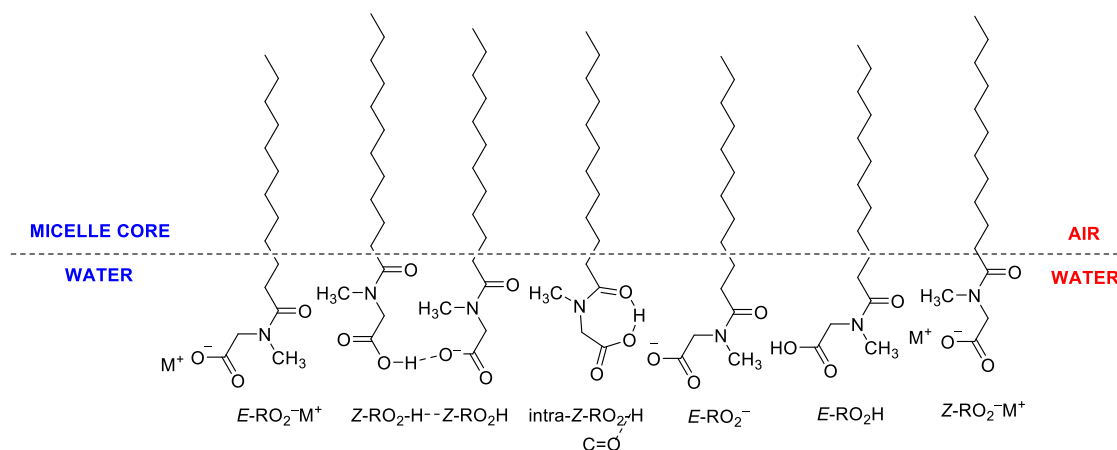
Chart 4.1. *Z* and *E* rotamers of NaLS and their short chain acetyl analogs, all show signals for *Z* and *E* in water and micelles. Carbons **a**, α -CH₂-CO₂⁻ (two s), **b**, CH₃-N- (two s), and **c**, α -CH₂-C=O (two t) Assignments of *Z* and *E* signals by 2D-NOESY NMR were made by Ambühl *et al.*¹⁵



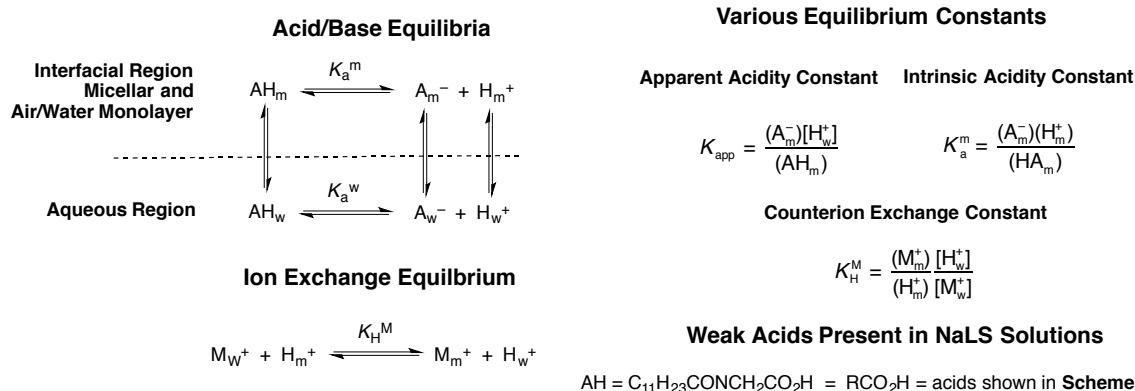
Scheme 4.1. All equilibria within brackets, **A**, **B**, and **C** show the protonated and deprotonated forms of *N*-lauroylsarcosine, and its *Z* and *E* rotamers that are present at the air/water interface, in bulk solution, and in the interfacial region of the micellar pseudophase. Within brackets **A** are equilibria associated with the protonated surfactant in more acidic solutions. Note that the *Z* rotamer may cyclize to form a seven membered ring joined by a hydrogen bond between the carboxyl OH and the amide carbonyl. The *E* rotamer cannot form this isomer. Brackets **B** and **C** show the *E/Z* equilibria for *N*-lauroyl sarcosinate anions occur with and without counter cations. In **B** and **C**, unpaired cations, including H₃O⁺, are also associated with interfacial region of the micellar pseudophase and these cations, are in dynamic equilibrium with headgroup paired cations and cations in the surrounding bulk phase, **Scheme 4.2**.

region of the micellar pseudophase, and the aqueous bulk phase containing surfactant monomer at different pH values and surfactant concentrations.

Equilibria in sodium N-lauroyl sarcosinate solutions. **Scheme 4.3** illustrates some of the relevant acid-base equilibria in NaLS and MLS solutions and that acidity constants in micellar solutions can be defined more than one way. The monomeric carboxylate headgroup of NaLS or other long chain carboxylate surfactant are the conjugate bases of relatively weak acids that within the micellar interfacial region should have pK_a values similar to their *N*-acyl analogs, **Chart 4.1**, about 3 in aqueous solution.^{8,9,12,13,16} The pK_a values of the carboxylic groups at micellar interfaces and their corresponding air/water monolayers contain the conjugate bases of several different acids (e.g., the cyclic *Z* rotamer) and measuring the pK_a values of the different groups within the interfacial region and assigning the size of their contributions to the minimum is extremely difficult, **Scheme 4.2**.^{18,19} The apparent acidity constant, pK_{app} , is the easier to determine because there are often available methods for determining the conjugate base/acid ratio, of weak acid indicators within the micellar pseudophase.²⁰ The pK_{app} can be estimated from log



Scheme 4.2. Single image illustrating the various forms of *N*-lauroylsarcosine that is applicable to both the air/water and micellar core/water interfaces. A number of forms (full tail at interface, R in bulk water) are illustrated including: *Z* and *E* rotamers, protonated and deprotonated, ion pairs, M^+ , and inter and intra molecular (*Z* only, cyclic form) hydrogen bonding. Most of the surfactant above the cmc is in micellar form and only a small fraction ($< \text{cmc}$) is in monomer form and at the air/water interface. The molarities of surfactant in the aqueous region are millimolar amounts and in molar amounts ($\sim 1 \text{ M}$) air/water and micellar interfacial region. The salt concentrations in the aqueous and micellar interfaces, NaOH, phosphate and acetate buffers are similar to their aqueous concentrations.



Scheme 4.3. Acid/base and ion exchange equilibria in micellar solutions of NaLS. The general equilibria show the acid, base, monovalent metal ion and proton forms in dynamic equilibrium between the air/water monolayer, the micelles and the aqueous region. Brackets, [], indicate concentration in moles per liter of solution volume and parentheses, (), indicate concentration in moles per liter of interfacial volume.¹⁷ **Scheme 4.2** shows the multiple weak acids that may contribute to K_{app} and K_a^m to various extents.

$(A_m^-)/(AH_m)$, (or equivalently $[A_m^-]/[AH_m]$) and the measured pH, which is assumed to depend only on the proton concentration in the aqueous pseudophase, $[H_w^+]$.^{17,20,21} However, values of pK_{app} are not reliable measures of functional group acidities at micellar interfaces because buffers do not control interfacial pH as was demonstrated some time ago.²¹ In general, hydronium ions associate with anionic micelles just as other cations do and the local interfacial hydronium ion concentration can be 1-2 orders of magnitude greater than that in the aqueous solution.²² In addition, the micellar interface acts as an ion exchanger,^{23,24} and cations added as salts displace protons from the interfacial region. Thus, the hydronium concentration within the micellar interface, in moles per liter of interfacial volume, may be greater than that in the aqueous region, but because of selective ion exchange the interfacial proton concentrations depend on both surfactant and salt concentrations and type. The value of the intrinsic acidity constants, pK_a^m , which is a measure of the medium effect of the micellar interface on the acid strength of a functional group can be estimated by using an ion exchange model for the distribution of the proton between the micellar and aqueous pseudophases,^{18,19,23,25} but estimating pK_a^m values of the terminal carboxylic acid group of *N*-lauroylsarcosine is beyond the scope of this work.

4.2. Experimental Section

Materials. All aqueous solutions used in these experiments were prepared in water that was distilled, passed over columns of charcoal and deionization resin, and then redistilled using a Corning water purifier, LD-5a. Anhydrous lithium hydroxide (Alfa Aesar), anhydrous sodium hydroxide and sodium bicarbonate (Fisher Scientific, standardized 1.0 M hydrochloric acid (Fluka), and anhydrous ethanol (Pharmco-Aaper), were reagent grade and used as received. Except for the surfactants, all other commercial reagents including HPLC grade solvents of diethyl ether, 2-propanol, methanol and hexane were used as received.

Preparation of N-lauroyl sarcosinates and N-lauroylsarcosine. Commercial sodium *N*-lauroyl sarcosinate, NaLS (Sigma-Aldrich) was dissolved in HPLC grade hot methanol and insoluble material was removed by filtration on a Büchner funnel. The filtrate was transferred into a beaker, cooled slowly to room temperature, during which white crystals appeared. The mixture was placed into an ice bath for 15 minutes, the solid was collected in a Büchner funnel, air dried, washed with cold dry diethyl ether, and dried in a vacuum oven. This process was repeated three more times, and the white crystals dried in a vacuum oven to a constant mass and stored in a desiccator.

HLS was prepared for use in surface tension experiments by precipitating HLS from an aqueous NaLS solution with excess concentrated aqueous HCl and cooled in an ice bath for 15 min. The white precipitate was collected on a Büchner funnel, washed with aliquots of cold water to get rid of excess HCl, dried in a vacuum oven to a constant mass and stored in a desiccator.

Alkali metal *N*-lauroyl sarcosinate salts (MLS) were prepared from HLS with four different cations: Na^+ , Li^+ , Cs^+ and tetramethylammonium, TMA^+ by titrating HLS dissolved in 10-20 mL of methanol with a concentrated aqueous solution of the metal hydroxide containing the alkali metal counter-ion to ca. pH 8-9 based on pH measurements. The excess solvent was removed on

a rotoevaporator and the white solid was recrystallized from either EtOH or MeOH ($\geq 3 \times$) and vacuum dried to a constant mass and stored in a desiccator.

*Methods. Surface Tension Measurements.*⁶ A du Noüy ring tensiometer, Fisher Surface Tensiomat (Model 21), was used for all surface tension measurements. Surface tension/MLS profiles were obtained of a series of surfactant solutions at ambient temperature (ca. 23°C). A glass recrystallizing dish was carefully cleaned to remove surface active impurities by rinsing the dish serially with HPLC grade hexanes, MeOH, and distilled water and shaken dry. The platinum ring was rinsed serially by HPLC grade hexanes, MeOH, and distilled water and dried in the air and passed through a flame to burn off impurities.

Solution preparation for surface tension experiments. **Unbuffered** surfactant (MLS) solutions were prepared by titrating an 4 - 5 mM aliquot of 27 - 30 mM purified MLS (M = TMA, Li, Na and Cs) solution with small volumes of aqueous 2 M HCl or 2 M NaOH to obtain the desired initial measured pH (see examples, **Figures 4.1** and **4.2**). HLS solution was prepared by dissolving 0.5 mM purified HLS in distilled water. Note: the volume changes of the NaLS solutions were trivial. **Buffered** surfactant (MLS) solutions were prepared by dissolving a weighed amount of solid MLS into the aqueous buffer solutions with the concentrations and compositions shown in **Figures 4.1** and **4.2**. Initial MLS concentrations were 27 - 30 mM. Small aliquots of concentrated acid or base were added as needed, to set the initial pH determined by a pH meter. Specific buffers were prepared as follows: a) Phosphate buffers at various cation concentrations (0.1, 0.01 and 0.001 M, respectively) were prepared by adding M_2HPO_4 stock solution (at 0.05, 0.005 and 0.0005 M, respectively) dropwise to MH_2PO_4 stock solution (at 0.1, 0.01 and 0.001 M, respectively), to pH 6. This procedure ensures that molarity of the cation in each stock solution is the same, 0.1, 0.01 and 0.001M, respectively. b) Acetate buffer at various concentrations (at 0.1, 0.01 and 0.001 M, respectively) was prepared by titrating MOAc stock solution at designated concentration (at 0.1, 0.01 and 0.001M, respectively) with a small aliquots

of glacial acetic acid (HOAc) to pH 6. c) NaOH-NaHCO₃ buffer at constant 0.1 M Na⁺ was prepared by titrating aqueous 0.1 M NaHCO₃ with small aliquots of aqueous 1.0 M NaOH to pH 6. Typically, a 10 mL aliquot of surfactant stock solution (except when the supply of solution was limited and 4 mL was used) of MLS was transferred into the glass recrystallizing dish by using a volumetric glass pipette. Surface tensions were measured repeatedly at each surfactant concentration until the variation of three consecutive measurements was $\leq 0.1 \text{ mN}\cdot\text{m}^{-1}$ indicating equilibrium had been reached. The surfactant solution was diluted with an aliquot of either water or aqueous buffer, typically 0.5 - 4 mL, and the measurement procedure was repeated.

pK_a Determination. The pK_a of HLS was determined by titrating 0.4 mL of 5 mM HLS aqueous solution with 200 mM NaOH aqueous solution at 1 μL increment, see **SI, Figure 4S3**. The pH values were measured on a two-buffer standardized Fisher Accumet pH meter.

¹H-NMR Spectroscopy. All NMR sample solutions were prepared by using purified surfactants except HLS, which was used as received without further purification (Aldrich, 95% pure). All spectra were recorded on Varian VNMRs 300 MHz or 400 MHz spectrometers using CDCl₃, CD₃OD or D₂O as solvents and with the temperature of each sample thermostat at 25°C. The Z/E rotamer ratio was determined by taking the ratio of the average values of integrated proton signal peak areas for three sets of protons on Z and E rotamer, **a**, $\alpha\text{-CH}_2\text{-CO}_2^-$ (two s), **b**, $\text{CH}_3\text{-N-}$ (two s), and **c**, $\alpha\text{-CH}_2\text{-C=O}$ (two t), **Chart 4.1** (example in **SI, Figure 4S1**).

High Performance Liquid Chromatography. Chromatographic separations were carried out using a Varian Microsorb MV 250 mm \times 4.6 mm, 100 μm pore size, 5 μm particle size, C₁₈ columns, with *i*-PrOH/MeOH mixtures as eluents on a Perkin-Elmer Series 200 composed of an auto-sampler, 4-solvent pump, injector, and detector controlled by TotalChrom Navigator 6.2.1 software package.

Mass Spectrometry. Surfactant purity and sample composition were checked as needed with





a quadrupole ion trap (LCQ) mass spectrometer with a electrospray needle voltage of -4.0 kV was obtained with the capillary temperature was set at 250°C.

4.3. Results

Surface Tension Measurements. The surface tension experiments shown in **Figure 4.1** were carried out by diluting NaLS with Na^+ and in **Figure 4.2** with other monovalent counterions, $\text{M}^+ = \text{TMA}^+, \text{Cs}^+, \text{and Li}^+$, MLS. Different surface tension plots show the effects of pH (or $[\text{NaOH}]$), counterion type, and the presence and absence of added buffer on the measured surface tension. The initial solution volumes of NaLS and MLS are 10 mL with concentrations of about 27 mM (right side of **Figures 4.1** and **4.2**). They are diluted with ca. 4 – 5 mL aliquots of distilled water containing the same concentrations of buffer as the initial NaLS or MLS solution until reaching the final surfactant concentration at which the surface tension is rising rapidly (left side of **Figures**). Note that the initial NaLS and MLS molarities are significantly greater than any observable cmc values, i.e., surface tension profiles without minima in **Figures 4.1** and **4.2**. **Figure 4.1A** contains two surface tension curves beginning at an initial pH of 7.4 (ambient) and 8.4 (a small aliquot of NaOH added). The minima show a decrease of 5 - 6 $\text{mN}\cdot\text{m}^{-1}$ and are centered at about 9 - 10 mM NaLS. From beginning to the end of the dilution, the solution pH has dropped ca. 1.5 units (**Table 4.1, Columns 3** and **4**). Dilution reduces the micelle concentration and releases Na^+ and H^+ into the bulk solution lowering the bulk pH. At ca. 6 mM NaLS, and ca. pH 6, the aqueous H^+ concentration is ca. 10^{-6}M , indicating that the concentration of H^+ released is much less than the remaining micellized surfactant concentration. The pH drop associated with **Figure 4.1A** is shown in **Table 4.1, Column 3**.

Figure 4.1B contains two surface tension profiles, one with added NaOH to increase the pH to 11, and the other containing 0.1 M NaHCO_3 buffer at pH 11. Both plots have sharp break points and the cmc for NaHCO_3 is considerably lower, consistent with the well-established effect of added salts reducing cmc values.² These solutions contain sufficient OH^- ions and all HLS in

Table 4.1. pH values at various concentrations of NaLS (no buffer added) at ambient temperature. The numbered/color coded column headings indicate the results for particular experiments in **Figure 1**.

Conc (mM)	pH (NaLS)	pH (NaLS)	pH (NaLS)	pH (NaLS)
	1C 	1A 	1A 	1B 
27.2	6.0	7.4	8.3	11.0
13.6	5.5	6.8	7.5	10.5
9.1	5.2	6.4	7.1	10.3
6.8	5.0	6.2	<i>N/A</i>	10.0
4.6	5.0	6.0	6.7	9.8

the micelles and at the air/water monolayer has been converted to LS^- . Because $[HLS] = 0$, no HLS is released during the dilution of micelles to monomer below the cmc. (See **DISCUSSION** for a more detailed interpretation of the results.) Note that when the initial pH is greater than 10, **Table 4.1**, the solution pH remains basic below the cmc. **Figure 4.1C** shows the effect of added sodium phosphate buffer on the surface tension profiles at three different constant sodium ion concentrations: 0.1 M, 0.001 M, and 0 M (see Experimental). In 0.1 M, the initial solution pH is titrated to 6, and the micelles contain both LS^- and HLS. However, the buffer adds 0.1 M Na^+ ions that exchange with H^+ at the micellar interface, **Scheme 4.3**. The added $[Na^+]$ reduces the $[H^+]$ concentration in the micellar interfacial region and in air/water monolayer to a level that it no longer produces a minimum in the surface tension curve, but the rotamer ratios remain unchanged (see **Figure 4.4**). Note also that the cmc in 0.1 M phosphate buffer is about the same as in 0.1 M $NaHCO_3$ solution (~ 3 mM), **Figure 4.1A**. When the phosphate buffer concentration is reduced to 0.001 M, the surface tension profile shows the same minimum as that observed at 0 M phosphate, initial pH = 6, **Figure 4.1D** (Inset).

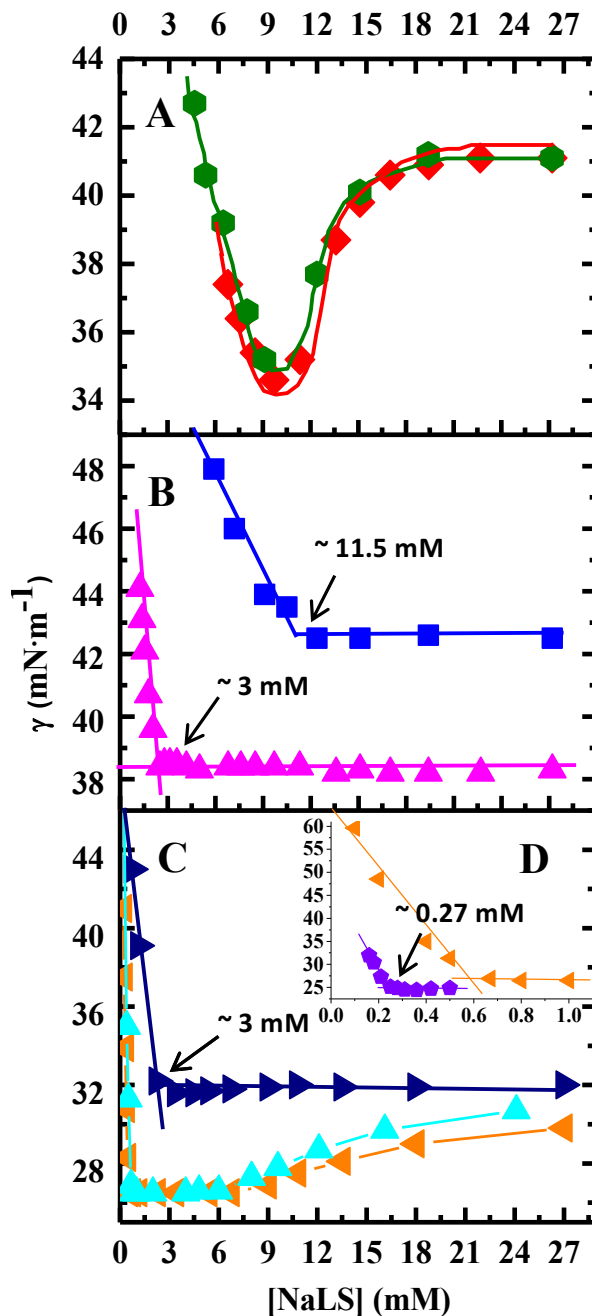


Figure 4.1. Surface tension plots of aqueous solutions of NaLS at room temperature (23°C). **1A.** \blacklozenge : at ambient initial pH 7.4 without buffer. \blacklozenge : at pH 8.3 without buffer (titrated with NaOH). **1B.** \blacksquare : at initial pH 11 without buffer (titrated with NaOH); \blacktriangle : at initial pH 11 with 0.1 M NaOH/NaHCO₃ buffer. **1C.** \blacktriangleright : at initial pH 6 with 0.1 M phosphate buffer. \blacktriangleleft : at initial pH 6 with 0.001 M phosphate buffer. \blacktriangleleft : at initial pH 6 without buffer (titrated with HCl to pH 6). Inset, **Figure 1D.** \blacktriangleleft : an enlargement of solution at initial pH 6 without buffer (titrated with HCl) at \blacklozenge diluted concentration (~ 1 mM). \blacklozenge : HLS at initial pH = 3.7. (Solution preparation details are in the experimental. Note the variation in the surface tension scales when comparing the profile shapes.) Clear cmc values are noted.

The curve with the break at 0.6 mM NaLS is reproduction of the curve in **Figure 4.1C** at 0 M phosphate buffer. The purple pentagons show the surface tension profile of HLS with an initial concentration of 0.5 M and pH = 3.7. Note that the cmc is about 0.21 mM and the low surface tension values at the cmc, are $24 \text{ mN}\cdot\text{m}^{-1}$. HLS is clearly less polar than NaLS and its surface tension above the cmc is well below the minimum in the plots of **Figure 4.1A**, a mixture of unknown composition of HLS and NaLS.

Figure 4.2 shows the surface tension profiles of aqueous MLS solutions with three counterions: TMA^+ , Cs^+ , and Li^+ . The results are similar to those found for NaLS at either ambient pH or an initial pH of 6.0, with or without buffer. The depth of the surface tension minima, i.e., the difference between the surface tension values at initial concentrations and those at the minima, for each cation in the absence of buffer increases in the order (in $\text{mN}\cdot\text{m}^{-1}$): Cs^+ (7) $\approx \text{TMA}^+$ (7) $> \text{Na}^+$ (6) $> \text{Li}^+$ (3). The specific ion effects are small, but real. For all the cations in **Figure 4.2**, the relationship between solution composition and surface tension is similar to that for Na^+ in **Figure 4.1**. (a) the minimum is absent in 0.1 M phosphate buffer; (b) larger in 0.01 M phosphate or acetate buffer; and (c) largest in 0.001 M phosphate or acetate buffer. At a constant surfactant concentration at the minima, ca. 5 mM, **Figure 4.2**, increasing the concentration of buffer increases the measured surface tension. The same order is observed in the plateau region above about 20 mM surfactant. This increase is caused by the displacement of protons by the added alkali metal cations from the micellar and air/water monolayer interfacial region.

Solutions at an initial pH of 6 with no or low buffer concentrations (0, 1 or 10 mM) turn slightly turbid during sample dilution (example precipitation, see *SI*, **Figure 4S2**). Cloudiness normally appears at a surfactant concentration of 3 – 10 mM (**Figure 4.2**), and disappears at more dilute concentrations. Cooling of these turbid solutions at pH 5 in an ice bath produced a small amount of white precipitate that collected and washed with ice cold water. Several measurements, including melting points (*SI*, **Table 4S1**), a NaOH titration (*SI*, **Table 4S2**) and FT-IR spectra

(not shown) with peaks at 1201 and 1735 cm^{-1} , for the carboxylic acid group demonstrate the presence of free acid form, HLS.

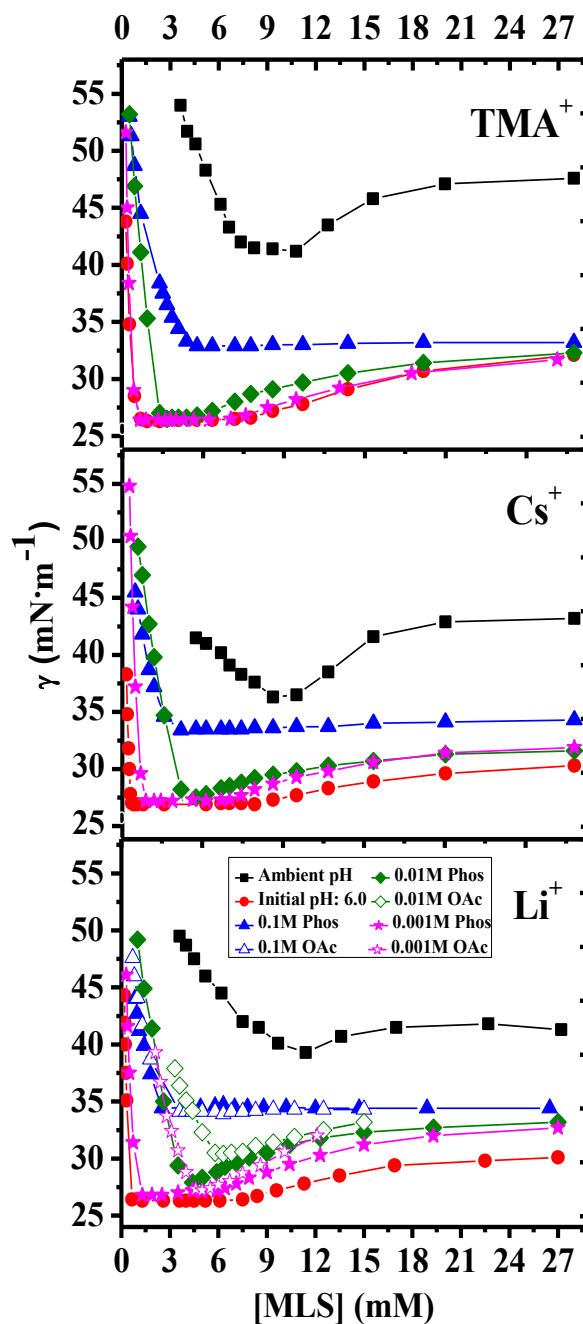


Figure 4.2. Surface tension measurements of aqueous solutions of MLS with different counterion types: tetramethylammonium (TMA⁺, top), cesium (Cs⁺, middle) and, lithium (Li⁺, bottom) with different initial pHs (either ambient pH or pH 6.0), with or without buffer and at room temperature (23°C). For measurements conducted in buffered solutions, phosphate buffers with different concentrations were used, except for lithium LiLS solutions, in which both phosphate and acetate buffer solution were used.

Z:E Rotamers of HLS, NaLS and MLS. The ^1H -NMR spectrum (**SI, Figures 4S1a-h**) of recrystallized *N*-lauroylsarcosine (HLS) in D_2O contains three sets of paired signals (**Chart 4.1**, splittings in caption). Each pair of signals attributed to a pair of *Z* and *E* rotamers because of the well-known restricted rotation around C–N bonds of amides.^{9,26,27} Rotational rate constants are available for a related structure, *N*-methylacetohydroxamic acid, $(\text{CH}_3\text{CON}(\text{OH})\text{CH}_3)$ in aqueous solution, where $k_{\text{ZE}} \approx 9 \text{ s}^{-1}$ and $k_{\text{EZ}} \approx 3 \text{ s}^{-1}$ at 27°C and the *Z/E* ratio $\approx k_{\text{EZ}}/k_{\text{ZE}} \approx 3/9 \approx 0.33$, which are slow on the ^1H NMR times scale.¹⁵ Given the structural difference between the hydroxamic acid and NaLS, 0.33 is reasonably close to the *Z/E* ratio of 0.91 for NaLS in its monomeric form, **Table 4.2**.²⁶ Finally, although the rotational rates are somewhat faster than bulk mixing time, both rotation rate constants are many orders of magnitude below the entrance and exit rate constants of surfactant molecules between micelles and water (see *Introduction*) and probably between the aqueous interface and water.

Table 4.2. *Z/E* rotamer ratios at two different methylenes, the *N*-methyl and their average value for the sarcosine surfactant at ambient pH with increasing [NaLS]. (Note: the minimum in these solutions is at about 10 mM, **Figure 1B**).

[NaLS] (mM)	$\alpha\text{-CH}_2\text{-C=O}$	$\text{CH}_3\text{-N}$	$\alpha\text{-CH}_2\text{-CO}_2^-$	Average
6.8	0.91	0.91	0.91	0.91
9.7	0.90	0.92	0.91	0.91
13.6	1.02	1.02	1.02	1.03
17.0	1.26	1.21	1.23	1.23
27.2	1.48	1.57	1.53	1.53
47.7	1.75	1.84	1.99	1.86

Table 4.2 also lists *Z:E* ratios for the rotamers of 6.8 mM (primarily monomer, *Z:E* ca. 1) to 47.7 mM (primarily micellar, *Z:E* ca. 2) NaLS (See also **SI, Figures 4S1a – 4S1b**). The relative stabilities of NaLS rotamers in water and micelles are affected by a number of interactions. As micellization increases, the hydrocarbon chains, which are partially folded in aqueous solution to minimize hydrocarbon-water contact and are more unfolded in the micellar core because they are

in more hydrocarbon-like environment.⁸ (p. 2630, left top) In water, the headgroups of the monomeric form are covered with water molecules and the same is true in the interfacial region, although the degree of hydration may be less.²² The interfacial region is of intermediate polarity, less polar than water, but more polar than hydrocarbon.²⁸ Unlike the surrounding aqueous region, the interfacial region is a relatively concentrated solution of charged and uncharged sarcosine headgroups ($\geq \approx 1$ M) with and without added counterions that may participate in hydrogen bonding including intramolecular hydrogen bonding by the *Z*, but not *E* rotamer. An unknown fraction of headgroups may participate in intermolecular hydrogen bonding, however, only the *Z* rotamer can undergo intramolecular hydrogen bonding,¹⁶ **Schemes 4.1** and **4.2**.

At low NaLS concentrations up to about 13.6 mM, the *Z:E* is about 0.9 – 1. At higher [NaLS], the ratio increases steadily to about 1.8 to 2 at 48 mM NaLS showing that the *Z* rotamer is at least twice as stable as the *E* rotamer within the interfacial region of the micelles. The calculated *Z:E* ratio is based on the average peak areas and includes those of the monomer in bulk D₂O. Prior studies^{7,9,11,12,13} also demonstrated that the average *Z:E* rotamer ratios for all 3 sets of protons of NaLS is ~ 1 in D₂O when NaLS concentration is low, but becomes significantly > 1 as [NaLS] forms micelles at high concentrations, consistent with the *Z* rotamer being more stable than *E* rotamer at the micellar interface. **Table 4.3** shows that the *Z:E* ratios for HLS monomer is significantly larger than that for NaLS monomer at very low concentrations in bulk D₂O and CD₃OD. The value for the HLS monomer is > 1 , and the *Z:E* ratio increases with decreasing solvent polarity: D₂O (1.3, at 0.21 mM [cmc 0.27 mM], pH 2.2); CD₃OD (*ca.* 2.0) and ~ 4.4 in CDCl₃ (See also **SI, Figures 4S1g – 4S1h**). Thus, the *Z:E* ratios are larger in micelles than in water because they are influenced both by the lower polarity of the micellar interface and by the larger *Z:E* ratios for the protonated form, HLS. The surface tension minima for *N*-lauroyl sarcosinates has been attributed to the presence the pairs of *Z* and *E* rotamers.⁹

Table 4.3. Peak Area Ratios of the *Z* and *E* Rotamers at Low Surfactant Concentrations in Protonated (HLS) and Deprotonated (NaLS) in D₂O and CD₃OD.^a

Deuterated Solvents	HLS		NaLS		NaLS	
	[HLS] (mM)	Z/E	[NaLS] (mM)	Z/E	[NaLS] (mM)	Z/E
D ₂ O ^b	0.21	1.30	6.8 (pH 6.1)	0.91	6.8 (pH 11.3)	0.88
CD ₃ OD	1.0	2.00	6.8	0.89	6.8	0.75

a. In CDCl₃, *Z/E* = ~ 4.4 for 6.8 mM HLS and ~ 4.0 in 6.8 mM NaLS indicating that *Z/E* becomes even larger in this less polar solvent.

b. Solutions of HLS (cmc: 0.27 mM) were prepared by dissolving solid HLS in D₂O (measured pH 2.2 in D₂O. pD: 2.6. pD = pH × 1.06831 = pH + 0.44) and in CD₃OD. Solid NaLS was dissolved in D₂O and CD₃OD. The pH values in D₂O were obtained by adding small aliquots of DCl and NaOD, respectively.

Several other factors may contribute to larger *Z:E* ratios in micelles and within the air/water interface. Differences in *Z* and *E* rotamer concentrations in the interfacial region could also cause by a *pK_a* difference between the two rotamers. The *Z* rotamer has a higher *pK_a* than the *E* rotamer (see below). Thus, the *Z* rotamer is probably present in greater amounts in the micellar and air/water interface and will be more effective in reducing surface tension. In addition, only the *Z* rotamer can form a 7 membered ring by intramolecular hydrogen bonding, **Schemes 4.1 – 4.2**. Intramolecular hydrogen bonding has been invoked repeatedly to explain the increase in the fraction of *Z* rotamer in nonpolar solvents and intramolecular hydrogen bonding.^{16,29} Similarly, if the cyclic *Z* rotamer is more stabilized in micelles, perhaps by being located more toward the micellar interior, then intramolecular hydrogen bonding would also enhance the *Z:E* ratio. However, because of the high local concentrations of carboxylate and carboxylic acid groups within the micellar interface (> 1 M). *E* and *Z* rotamers could also be stabilized by intra or intermolecular hydrogen bonding with a neighboring *N*-lauroyl sarcosinate anion within the micellar interface, **Scheme 4.2**.⁶ Specific cations also have a modest effect on the *Z:E* ratio. The ratio is somewhat greater for TMA⁺, Cs⁺ and Na⁺ than Li⁺.

The Acidity of the Carboxylic Acid Group of HLS in Water, Mixed Micelles and Air/Water Monolayers. Although HLS has a limited solubility in water and the acidity constants of its monomers are difficult to determine, the acidity constants of the rotamers of its short chain, water

soluble analog *N*-acetyl sarcosine, **Chart 4.1**, were determined by ^{17}O NMR on the $\alpha\text{-CH}_2$ -: $\text{p}K_{\text{a}} = 3.37$ (*Z*) and 2.98 (*E*).^{8,16} The acidity constants of *N*-acetyl-*L*-prolinate on the $\alpha\text{-CH}_2$ - also by ^{17}O are: $\text{p}K_{\text{a}} = 3.36$ (*Z*) 2.79 (*E*),¹⁶ consistent with other reports show *Z*-acetyl prolinates are weaker acids than the *E* form.^{5,8} These values are considerably lower than that of acetic acid (4.76)³⁰ or monomeric lauric acid (5.3).³¹ Titration of 5 mM micellized HLS (**SI, Figure 4S3**) gave an apparent acidity constant, $\text{p}K_{\text{app}}$, of 5.0 for the micelles. similar to that of HLS micelles $\text{p}K_{\text{app}} = 5.4$.⁸ Both estimates are 1.7 – 2 $\text{p}K_{\text{a}}$ units higher than that of monomeric *N*-acetyl sarcosine in water, but are defined in terms of the proton concentration in the bulk aqueous phase and not in the interfacial region, **Scheme 4.3**.

Origin of the Surface Tension Minimum. The presence of minima in NaLS solutions (the description of NaLS solutions also fits MLS solutions) is caused by the presence of HLS in NaLS in micelles and the air/water monolayer and by different acidities of the *Z* and *E* rotamers (see above). At the initial [NaLS], 27 mM, the higher interfacial proton molarity compared to bulk solution neutralizes a fraction of the carboxylate headgroups in the micellar and monolayer interfacial regions to form mixed aggregates of NaLS and HLS. Dilution leads to greater dissociation of the more polar NaLS than HLS monomers from the mixed micelles and from the monolayer. The surface tension decreases gradually while the HLS/NaLS molar ratio increases in the air/water monolayer and the mixed micelles. Further dilution leads to at least a 5 fold decrease in micelle concentration with continued dissociation NaLS monomers. Some HLS may also dissociate donating protons to the aqueous phase and producing gradual decrease in solution pH, see **Table 4.1**. At very low surfactant concentrations, the solutions become faintly cloudy as the micelles break-up followed by dissociation of insoluble HLS. Further addition of water leads to dissolution of NaLS and HLS monomers in the bulk and at the air/water interfacial region and a rapid increase in surface tension, **Schemes 4.1** and **4.2**. See the **Discussion** for additional details and analysis.

¹H-NMR Chemical Shifts for NaLS and α -CH₂CO₂⁻ Protons. The ¹H NMR chemical shifts of protons or sets of protons on NaLS as a function of increasing [NaLS] are shown in **Figure 4.3**. Up to about 12 mM NaLS, δ is constant for each set of protons in the key. At 15 mM NaLS and above the proton signals break into two distinct regions: the protons in the surfactant tails shift downfield (upward in **Figure 4.3**) and the protons in the headgroup region shift upfield (downward in **Figure 4.3**). The downfield shift for the hydrocarbon protons occurs because of the transfer of the tails from water to the micellar hydrocarbon-like micellar core.³² The upfield shifts reflect the transfer from dilute aqueous solution to the relatively concentrated micellar interfacial region (headgroups and counterions ca. 1 M) that has a polarity similar to alcohol.³³ Because this region is less polar, the pK_a^m (intrinsic) of the carboxylic acid group in the micellar interfacial region should be higher than in bulk water. The trends in **Figure 4.3** are qualitatively the same as those observed by Deng, M. *et al.* for a similar surfactant, *N*-dodecanoyl-(4R)-hydroxy-*L*-proline.³⁴

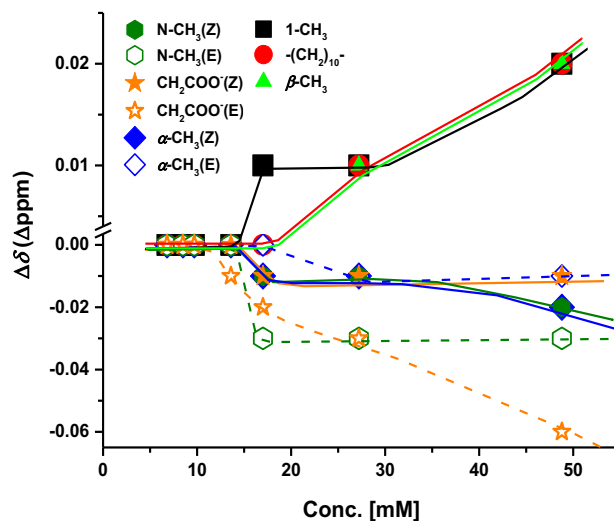


Figure 4.3. Change of ¹H chemical shifts in different chemical environments of NaLS molecules at various concentrations in nonmicellar and micellar aqueous solutions. $\Delta\delta$ values below ca. 12 mM in the Figure are essentially constant for all protons. [NaLS], $\Delta\delta = (\delta_n - \delta_{6.8} \text{ mM})$, where $\delta_{6.8} \text{ mM}$ is the reference spectrum. See **SI, Figure 4S1a-d** and **4S4-4S5** for NMR spectra in D₂O.

Figure 4.4 shows the change in the ^1H NMR chemical shifts of the pair of $\alpha\text{-CH}_2\text{CO}_2^-$ hydrogens from 9.7 mM to 47.7 mM NaLS for the *E* and *Z* rotamers. The shifts for the $\alpha\text{-CH}_2\text{CON-}$ and $\text{CH}_3\text{-N-}$ hydrogens over the same concentration range are significantly smaller (**SI, Figures 4S4-4S5**), and we limit our discussion to the $\alpha\text{-CH}_2\text{CO}_2^-$ hydrogens. Note that the chemical shift for the *E* rotamer moves upfield from about 3.99 ppm (green) at 9.7 mM to about 3.93 ppm (yellow) at 47.4 mM. The chemical shift for the *Z* rotamer initially moves upfield, but to a lesser extent than the *E* rotamer, and then reverses direction slightly. These upfield shifts are similar in size and direction to the upfield shifts of the $\text{-CH}_2\text{CO}_2^-$ signals when NaLS is dissolved in acidic D_2O and CD_3OD , see **Table 4.4**.

The results in **Figure 4.4** and **Table 4.4** are consistent with several concomitant changes in interfacial composition with increasing micelle concentration and increasing bulk pH. (a) The steady upfield shift for the *E* rotamer with increasing surfactant concentration, **Figure 4.4** (signals labeled with dots), is consistent with the continuous increase in micelle over monomer concentration and an increase in the fraction of *E* rotamer in deprotonated form, which correlates

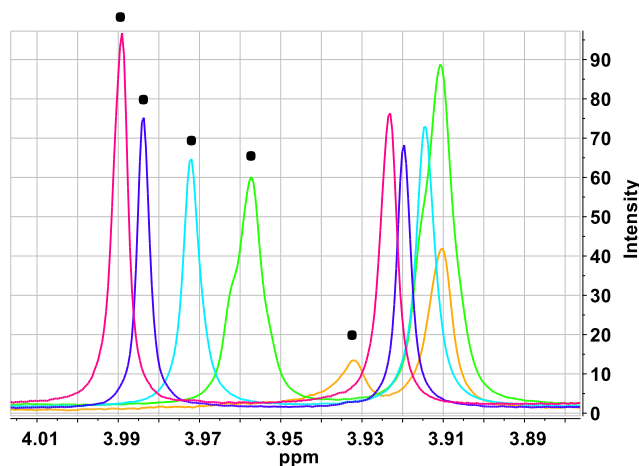


Figure 4.4. ^1H -NMR chemical shifts of the carboxylate methylene protons, $\text{C}_{11}\text{H}_{23}\text{CON}(\text{CH}_3)\text{CH}_2\text{COONa}$, in increasing concentrations of NaLS in D_2O solution: 9.7 mM (pink, below cmc), 13.6 mM (purple), 17.0 mM (cyan), 27.2 mM (light green) and 47.7 mM (orange). (Left Series, *E* rotamer with a black dot over each peak: Right Series: *Z* rotamer.) All the samples were measured at 25°C and the solution pH values are recorded in **Table 4.1, Column 3**.

with the increasing bulk pH, **Table 4.1**. (b) The upfield shift in the *E* rotamer signals in **Figure 4.4** of the $\alpha\text{-CH}_2\text{CO}_2^-$ is about 0.06 ppm and the *Z* rotamer should, in principle, show a similar shift upfield. Instead the observed chemical shift for the *Z* rotamer signal in **Figure 4.4** decreases only modestly and then reverses direction slightly. One possible explanation is that a larger fraction of the *Z* rotamer is protonated than the *E* rotamer, because it forms the cyclic structure via intramolecular hydrogen bonding¹⁶ (**Scheme 4.1A**). Thus, the observed ^1H chemical shifts of the NaLS

Table 4.4. Chemical shifts and the differences in the chemical shifts (Δ) of the $\alpha\text{-CH}_2\text{CO}_2\text{D}$ hydrogens of DLS and $\alpha\text{-CH}_2\text{CO}_2^-$ hydrogens of NaLS adjacent to a negative charge in D_2O and CD_3OD .

	DLS ¹		NaLS		$\Delta(\text{D}_2\text{O})$	$\Delta(\text{CD}_3\text{OD})$
	D_2O	CD_3OD	D_2O	CD_3OD		
Chemical Shift of <i>Z</i> (ppm)	4.09	4.08	3.92	3.94	-0.17	-0.14
Chemical Shift of <i>E</i> (ppm)	4.18	4.14	3.99	3.86	-0.19	-0.28

¹[HLS] = 0.21 mM (cmc: 0.27 mM, pH*: 2.2., pD: 2.6, pD = pH* \times 1.06831 = pH* + 0.44),³⁵ [NaLS in D_2O] = 6.8 mM (pH*: 6.1, pD: 6.5, pD = pH* \times 1.06831 = pH*+0.44), [DLS in CD_3OD] = [NaLS in CD_3OD] = 1 mM. Note that pH* is the apparent value as obtained from pH meter.

protons above and below the cmc are consistent with the interfacial region containing protonated carboxylic acid groups and the shifts in *Z* and *E* rotamers suggest that the *Z* rotamer includes the cyclic form and is more protonated than the *E* rotamer.

4.4. Discussion

The **Results** section presents a variety of evidence showing that the minima in surface tension curves of anionic micelles of sodium (NaLS) and metal cation (MLS) *N*-lauroyl sarcosinates are not caused by hydrophobic impurities, but by the presence of *Z* and *E* rotamers and nonionic HLS, that is, the minima represent the surface tension curves of mixed micelles of NaLS or MLS containing variable amounts of HLS. The size of the NaLS/HLS molar ratio is uncertain because small amount of impurities or more surface active surfactant $\geq 1\%^3$ may produce significant changes in the surface tension. However, the actual HLS concentrations are unknown because of

the difficulty in measuring small concentrations of HLS in NaLS and MLS micelles. Here we summarize how the various pieces of experimental information support the conclusion that the minima in the surface tension curves in **Figures 4.1** and **4.2** are caused by HLS and Z rotamer > E rotamer.

Surface tension profiles of pure surfactants. **Figures 4.1** and **4.2** show a variety of surface tension plots at ambient conditions, in basic solution, and in the presence of buffers and salts of various concentrations. These plots are typically in units of $\text{dynes}\cdot\text{cm}^{-1}$ or $\text{mN}\cdot\text{m}^{-1}$ versus \log [surfactant], but here we use a linear scale for the surfactant to emphasize the minima in the profiles. **Figures 4.1A** and **4.1C** show profiles with significant minima, but **Figures 4.1B**, **4.1C** and **4.1D** contain characteristic profiles of pure surfactants, i.e., an intersection point between two lines, one of near zero slope above the cmc and a second of steep slope below the cmc rising toward the surface tension of water. The same surfactant sample was used in all runs except that for HLS, which was prepared separately, see **Figure 4.1D** and **EXPERIMENTAL**.

Two conditions contribute to the surface tension profiles of pure surfactants when the system behaves ideally and only micelles and monomer are present: (a) At surfactant concentrations in excess of the cmc a plateau is observed because dilution with water leads to the break-up of micelles in the bulk solution that are in dynamic equilibrium with surfactant monomer in bulk solution and in the air/water monolayer and whose concentrations remain constant until the cmc concentration region is reached. (b) As the surfactant concentration decreases below the cmc region, the micelle concentration goes to zero and dilution of the surfactant monomer in bulk solution results in a concomitant dilution of surfactant monomer concentration at the air/water monolayer and the surface tension increases. Thus, the absence of a minimum in the surface tension profile is a strong indication that surfactant sample contains insignificant amounts of a hydrophobic impurity.

Evidence for surface tension minima caused by HLS and Z and E rotamers. The totality of the evidence presented in the **Results** shows that micelles NaLS and MLS micelles at ambient pH are best viewed as mixed micelles containing unknown amounts of HLS that reduce the surface tension at ambient pH. (a) *Surface Tension Data*, **Figures 4.1** and **4.2**. The clear minima in **Figure 4.1A** disappear at pH 11 without buffer and with 0.1 M bicarbonate buffer, **Figure 4.1B**. The disappearance of the minima is consistent with complete deprotonation of HLS in the micellar solutions. If the minima were caused by a hydrophobic, pH insensitive impurity, the minima would have remained. (b) At pH 6 in the presence of low concentrations of phosphate buffer, substantial minima are observed that are absent in solutions of pH 6 at 0.1 M phosphate buffer, **Figure 4.1C**. Adding 0.1 M phosphate buffer adds significant amounts of Na^+ , the Na^+ displaces H^+ from the micellar interfacial region by ion exchange such that the solution is now composed of NaLS micelles in a phosphate buffer solution, **Scheme 4.3**. However, despite the approximately 100 fold increase in salt concentration, the surface tension of these solutions above the cmc are still higher than those at lower salt concentrations, which is sensible if the protons on the hydrophobic component, have been displaced by ion exchange, **Scheme 4.3**. (c) Note that the surface tension values of pure HLS, and the solution without added buffer, **Figure 4.1D**, the surface tension values just above the cmc are much lower than those for NaLS in bicarbonate buffer consistent with the higher surface activity and lower cmc of HLS. (d) In **Figure 4.2**, surface tension plots of MLS surfactants show that: (i) the surface tension values have a modest dependence on counterion cation type; and (ii) at pH 6, surface tension curves at low buffer concentrations (therefore, low counterion concentrations) have minima, but increasing the buffer and counterion concentration makes the minima disappear, consistent with ion exchange of counterions with protons at the micellar interface, **Scheme 4.3**. (e) The *Z/E* rotamer ratios of the monomer in water are near 1, but the ratios increase within micelles to a value near two indicating that the *Z* form is approximately twice as stable as the *E* form within micelles, **Table 4.2**. The increase in the *Z/E* ratio in micelles is consistent with the increase in the ratio in CD_3OD for DLS,

but no change was observed for NaLS in this solvent, consistent with the *Z* rotamer of the neutral headgroup having a lower free energy within the micellar interface than the *E* form or the anionic forms of the headgroup. (f) Published estimates of the pK_a values of the *Z* (ca. 3.4) and *E* (ca. 2.9) rotamers of *N*-acetyl sarcosine by ^{17}O NMR show that the terminal carboxylic acid group is a stronger acid than a simple alkyl carboxylic acid, but that the *Z* rotamer of *N*-acetyl sarcosine is a weaker acid than the *E* rotamer by several tenths of a pK_a unit.¹⁶ The intrinsic pK_a^m 's of the HLS carboxylic acid groups are unknown because the intrinsic pK_a^m (see **Scheme 4.3**) of the carboxylic acid groups within the interfacial and the medium effect on its pK_a^m is unknown. Estimates of the pK_{app} for HLS in micellar form are about 5.⁸ However, because NaLS forms anionic micelles, the interfacial counterion concentration of protons may be 1 - 2 orders of magnitude higher than that in the surrounding aqueous solution.²² (g) Because the positions of equilibrium of the various components within the micellar interfacial region and at the air/water interface are probably similar, the results indicate that because larger fraction of the *Z* form of *N*-lauroylsarcosine is protonated, this form makes a greater contribution to lowering the measured surface tension. And the ^1H chemical shifts of the NaLS protons above and below the cmc are consistent with the interfacial region containing protonated carboxylic acid groups and the shifts *Z* and *E* rotamers suggest that the *Z* rotamer is more protonated than the *E* rotamer. (h) The interfacial region is of intermediate polarity, less polar than water, but more polar than hydrocarbon.²⁸ Unlike the surrounding aqueous region, the interfacial region is a relatively concentrated solution of charged and uncharged sarcosine headgroups ($\geq \approx 1\text{ M}$) with and without added counterions that may participate in hydrogen bonding including intramolecular hydrogen bonding by the *Z*, but not *E* rotamer. And a fraction of headgroups may participate in intermolecular hydrogen bonding. However, only the *Z* rotamer can undergo intramolecular hydrogen bonding,¹⁶ **Schemes 4.1** and **4.2**.

4.5. Summary

Based on our and literature results, we conclude that the minima in the surface tension profiles of purified NaLS and MLS are a red herring. They are not a harbinger of impurities but experimentally observable properties that depend on the presence of *Z* and *E* rotamers with $Z > E$, the pK_a^m (intrinsic) values of the carboxylate groups of the two rotamers, counterion type, and buffer concentrations and variable degrees of neutralization (protonation) of the carboxylate headgroups. Protonation of the carboxylic headgroups reduces the fraction of the charged headgroups creating mixed micelles of uncharged HLS and anionic NaLS and MLS and air/water interfacial region should contain a mixed monolayer of the same two surfactants. Had the intrinsic pK_a^m of the *N*-lauroylsarcosines been ca. 3 pK_a units higher or lower, the carboxylic acid headgroup of the surfactant would be either completely protonated or deprotonated and the minima would not have been observed. Thus, the appearance of the surface tension minima is the result of the balance of a relatively singular set of interactions in dynamic equilibrium in the micellar solution and the air/water interface.

Bibliography

- (1) Mukerjee, P.; Mysels, K. J. Critical Micelle Concentrations of Aqueous Surfactant Systems. *Critical Micelle Concentrations of Aqueous Surfactant Systems*; National Bureau of Standards: Washington, DC, 1971.
- (2) Jonsson, B.; Lindman, B.; Holmberg, K.; Kronberg, B. *Surfactants and Polymers in Aqueous Solution*; John Wiley & Sons: Chichester, 1998.
- (3) Miles, G. D. Minima in Surface Tension-Curves and Interfacial-Tension Curves. *J. Phys. Chem.* **1945**, *49*, 71-76.
- (4) Infante, M. R.; Perez, L.; Pinazo, A.; Clapes, P.; Moran, M. D. C. Amino Acid-Based Surfactants. In *Novel Surfactants: Preparation, Applications, And Biodegradability*; 2nd ed.; Holmberg, K., Ed.; Taylor & Francis: New York, 2005, p 623.
- (5) Guba, W.; Haessner, R.; Breipohl, G.; Henke, S.; Knolle, J.; Santagada, V.; Kessler, H. Combined Approach of NMR and Molecular Dynamics within a Biphasic Membrane Mimetic: Conformation and Orientation of the Bradykinin Antagonist Hoe 140. *J. Am. Chem. Soc.* **1994**, *116*, 7532-7540.
- (6) Zhang, Y. L.; Romsted, L. S.; Zhuang, L. Z.; de Jong, S. Simultaneous Determination of Interfacial Molarities of Amide Bonds, Carboxylate Groups, and Water by Chemical Trapping in Micelles of Amphiphiles Containing Peptide Bond Models. *Langmuir* **2013**, *29*, 534-544.
- (7) Ledneck, I.; Forgo, P.; Kiss, J. T.; Molnar, K.; Palinko, I. Conformational Behaviour of Acetamide Derivatives Studied by NMR Spectroscopic and Computational Methods. *J. Mol. Struct.* **2007**, *834*, 349-354.
- (8) Borocci, S.; Mancini, G.; Cerichelli, G.; Luchetti, L. Conformational Behavior of Aqueous Micelles of Sodium *N*-Dodecanoyl-*L*-proline. *Langmuir* **1999**, *15*, 2627-2630.
- (9) Cerichelli, G.; Luchetti, L.; Mancini, G. Conformational Behavior of Aqueous Micelles of Sodium *N*-Dodecanoyl-*N*-methylglycinate. *Langmuir* **1997**, *13*, 4767-4769.
- (10) Kaneko, A.; Sehgal, P.; Doe, H. Interfacial and Aggregation Properties of Aqueous Sodium *N*-Dodecanoyl Sarcosinate Solutions at Different pH. *Colloid Polym. Sci.* **2012**, *290*, 323-330.
- (11) Oshimura, E.; Yamashita, Y.; Sakamoto, K. Conformational Behavior of *N*-Acylamino Acid Oil and *N*-Acylamino Acid Surfactant in Aqueous Solution. *J. Oleo Sci.* **2007**, *56*, 115-121.
- (12) Takahashi, H.; Nakayama, Y.; Hori, H. The Proton Magnetic Resonance Spectra and Molecular Conformations of Sodium *N*-Acyl Sarcosinates in Aqueous Solution. *J. Colloid Interf. Sci.* **1976**, *54*, 102-107.
- (13) Yahagi, K.; Tsujii, K. Conformational Behaviors of *N*-Substituted *N*-Lauroylglycinates in Aqueous Micellar and DMSO Solutions and Their Krafft Points. *J. Colloid Interf. Sci.* **1987**, *117*, 415-424.
- (14) Yamashita, Y.; Kunieda, H.; Oshimura, E.; Sakamoto, K. Formation of Intermediate Micellar Phase Between Hexagonal and Discontinuous Cubic Liquid Crystals in Brine/*N*-Acylamino Acid Surfactant/*N*-Acylamino Acid Oil System. *J. Colloid Interf. Sci.* **2007**, *312*, 172-178.
- (15) Ambuhl, M.; Bangerter, F.; Luisi, P. L.; Skrabal, P.; Watzke, H. J. Configurational Changes Accompanying Vesiculation of Mixed Single-Chain Amphiphiles. *Langmuir* **1993**, *9*, 36-38.
- (16) Lauterwein, J.; Gerotheranassis, I. P.; Hunston, R. N. A Study of the Cis Trans Isomerism of *N*-Acetyl-*L*-Proline in Aqueous-Solution by O-17 NMR-Spectroscopy. *J. Am. Soc. Chem. Comm.* **1984**, 367-369.

- (17) Romsted, L. S. In *Surfactants in Solution*; Mittal, K. L., Lindman, B., Eds.; Plenum Press: New York, 1984; Vol. 2, p 1015.
- (18) Romsted, L. S.; Zanette, D. Quantitative Treatment of Indicator Equilibria in Micellar Solutions of Sodium Decyl Phosphate and Sodium Lauryl Sulfate. *J. Phys. Chem.* **1988**, *92*, 4690-4698.
- (19) Romsted, L. S. Quantitative Treatment of Benzimidazole Deprotonation Equilibria in Aqueous Micellar Solutions of Cetyltrimethylammonium Ion (CTACl⁺, CTABr⁺, and CTANO₃⁺) Surfactants. 2. Effect of Added Salt. *J. Phys. Chem.* **1985**, *89*, 5113-5118.
- (20) Fernandez, M. S.; Fromherz, P. Indicators as Probes of Electrical Potential and Polarity in Micelles. *J. Phys. Chem.* **1977**, *81*, 1755-1761.
- (21) Quina, F. H.; Politi, M. J.; Cuccovia, I. M.; Baumgarten, E.; Martins-Franchetti, S. M.; Chaimovich, H. *J. Phys. Chem.* **1980**, *84*, 361-365.
- (22) Romsted, L. S. Do Amphiphile Aggregate Morphologies and Interfacial Compositions Depend Primarily on Interfacial Hydration and Ion-Specific Interactions? The Evidence from Chemical Trapping. *Langmuir* **2007**, *23*, 414-424.
- (23) Bunton, C. A.; Nome, F.; Quina, F. H.; Romsted, L. S. Ion Binding and Reactivity at Charged Aqueous Interfaces. *Acc. Chem. Res.* **1991**, *24*, 357-364.
- (24) Morgan, J. D.; Napper, D. H.; Warr, G. G. Thermodynamics of Ion Exchange Selectivity at Interfaces. *J. Phys. Chem.* **1995**, *99*, 9458-9465.
- (25) Romsted, L. S. Quantitative Treatment of Benzimidazole Deprotonation Equilibria in Aqueous Micellar Solutions of Cetyltrimethylammonium Ion (CTAX, X⁻ = Cl⁻, Br⁻, and NO₃⁻) Surfactants. 1. Variable Surfactant Concentration. *J. Phys. Chem.* **1985**, *89*, 5107-5118.
- (26) Brandes, S.; Sornosa-Ten, A.; Rousselin, Y.; Lagrelette, M.; Stern, C.; Moncomble, A.; Cornard, J. P.; Meyer, M. Conformational and Structural Studies of *N*-Methylacetohydroxamic Acid and of Its Mono- and Bis-Chelated Uranium (VI) Complexes. *J. Inorg. Biochem.* **2015**, *151*, 164-175.
- (27) Ozawa, T.; Isoda, Y.; Watanabe, H.; Yuzuri, T.; Suezawa, H.; Sakakibara, K.; Hirota, M. Rotational Barriers and N-15 Chemical Shifts of *N*-Acyl-*N*-Alkyl-Substituted Amino Acids. *Magn. Reson. Chem.* **1997**, *35*, 323-332.
- (28) Shin, D.; Schanze, K. S.; Whitten, D. G. Solubilization Sites and Orientations in Microheterogeneous Media. Studies Using Donor-Acceptor-Substituted Azobenzenes and Bichromophoric Solvatochromic Molecules. *J. Am. Chem. Soc.* **1989**, *111*, 8494-8501.
- (29) Williams, K. R.; Adhyaru, B.; German, I.; Alvarez, E. The Cis-Trans Equilibrium of *N*-Acetyl-*L*-Proline - An Experiment for the Biophysical Chemistry Laboratory. *J. Chem. Educ.* **2002**, *79*, 372-373.
- (30) Jencks, W. P.; Regenstein, J. Chapter 67. Ionization Constants of Acids and Bases; in *Handbook of Biochemistry and Molecular Biology*, 4th ed.; pages 595-635; Lundblad, R. L.; MacDonald, F. M., Eds.; CRC Press, 2010.
- (31) Ionisation Constants of Organic Acids in Aqueous Solution; in *International Union of Pure and Applied Chemistry (IUPAC) Chemical Data Series*; No. 23. Serjeant, E. P.; Dempsey, B., Eds.; Pergamon Press, Inc.: New York, 1979.
- (32) *Solvents and Solvent Effects in Organic Chemistry*; 4th ed.; Reichardt, C. W. T., Ed.; Wiley-VCH Verlag GmbH & Co: Weinheim, 2011.
- (33) *Membrane Mimetic Chemistry*; Fendler, J. H., Ed.; Wiley-Interscience: New York, 1982.
- (34) Deng, M. L.; Huang, X.; Wu, R. L.; Wang, Y. L. Micellization-Induced Conformational Change of a Chiral Proline Surfactant. *J. Phys. Chem. B.* **2008**, *112*, 10509-10513.
- (35) Krezel, A.; Bal, W. A Formula for Correlating pK_a Values Determined in D₂O and H₂O. *J. Inorg. Biochem.* **2004**, *98*, 161-166.

Appendix

Figures 4S1a-h. ^1H -NMR spectra of NaLS and HLS showing signals for the *E* and *Z* isomers of the surfactants under variable experimental conditions. The *E* and *Z* isomers for each signal in all spectra are marked above their respective peaks and the average *Z*:*E* ratio is written in the caption. **Figure S1a** is the full spectrum of NaLS. Remaining spectra show the changes from ca. 2.0 to 4.0 ppm. The signals at higher and lower chemical shifts are all the same. All runs carried out at 20 $^{\circ}\text{C}$.

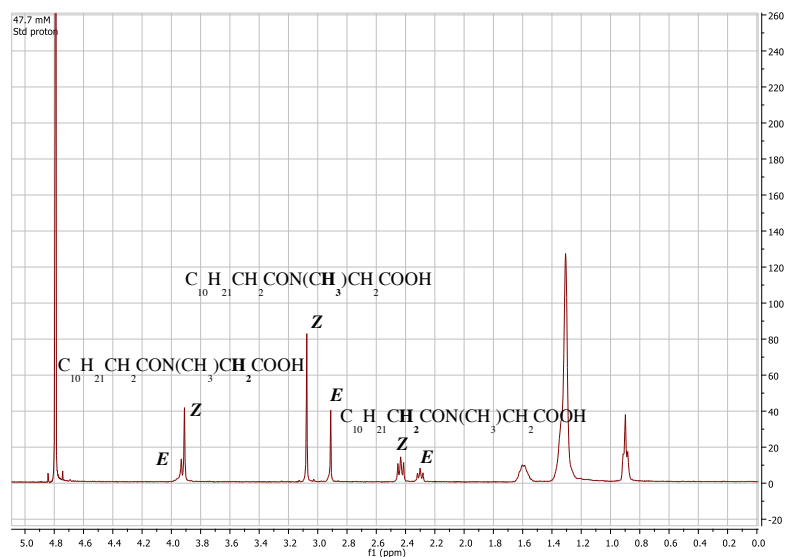


Figure 4S-1a. ^1H -NMR spectrum of 47.7 mM sodium *N*-lauroyl sarcosinate micelles (in D_2O). Note that the *Z* and *E* rotamer signal labels are this and follow ^1H NMR spectra. Average *Z*:*E* ratio = 1.86.

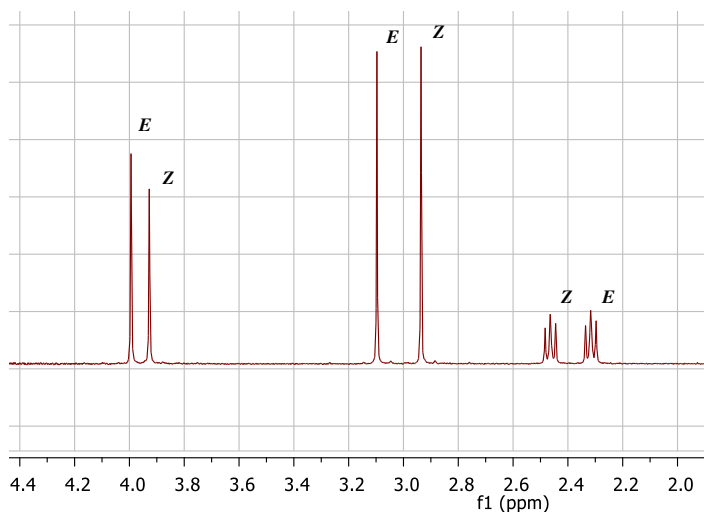


Figure 4S-1b. ^1H -NMR spectrum of 6.8 mM sodium *N*-lauroyl sarcosinate monomer in D_2O . Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2-$ are shown. Average *Z*:*E* ratio = 0.91.

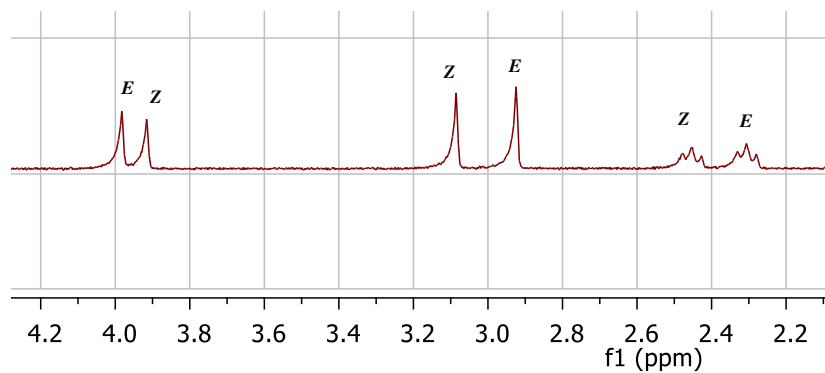


Figure 4S-1c. ^1H -NMR spectrum of 6.8 mM sodium *N*-lauroyl sarcosinate at pD 11 in D_2O (Titrated with NaOD). Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2^-$ are shown. Average *Z*:*E* ratio = 0.88.

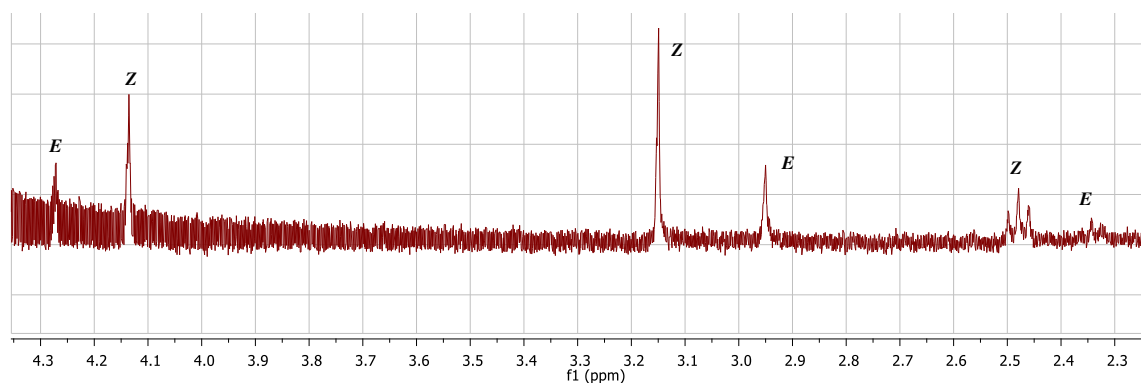


Figure 4S-1d. ^1H -NMR spectrum of 0.22 mM *N*-lauroylsarcosine at pD 2.2 in D_2O (Titrated with DCl). Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2^-$ are shown. Average *Z*:*E* ratio = 1.30.

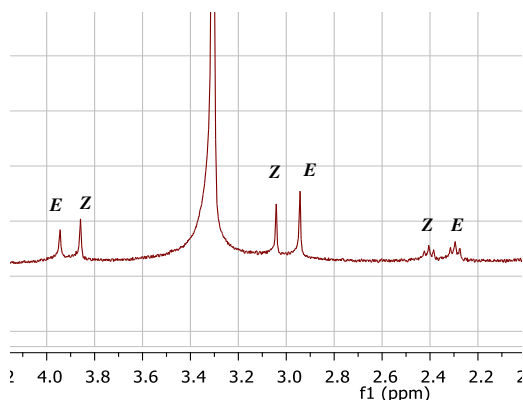


Figure 4S-1e. ^1H -NMR spectrum of 1 mM sodium *N*-lauroyl sarcosinate in CD_3OD . Note the solvent peak at ca. 3.3 ppm. Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2-$ are shown. Average *Z*:*E* ratio = 0.89.

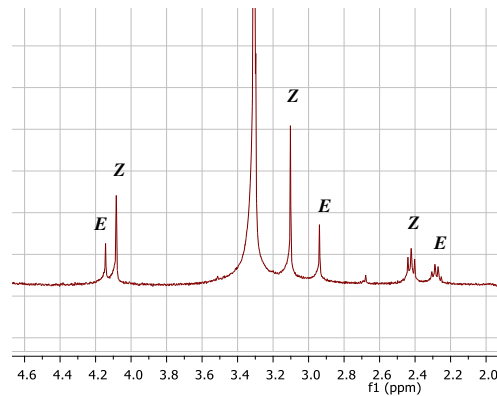


Figure 4S-1f. ^1H -NMR spectrum of *N*-lauroylsarcosine (1 mM) in CD_3OD . Note the solvent peak at ca. 3.3 ppm. Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2-$ are shown. Average *Z*:*E* ratio = 2.0.

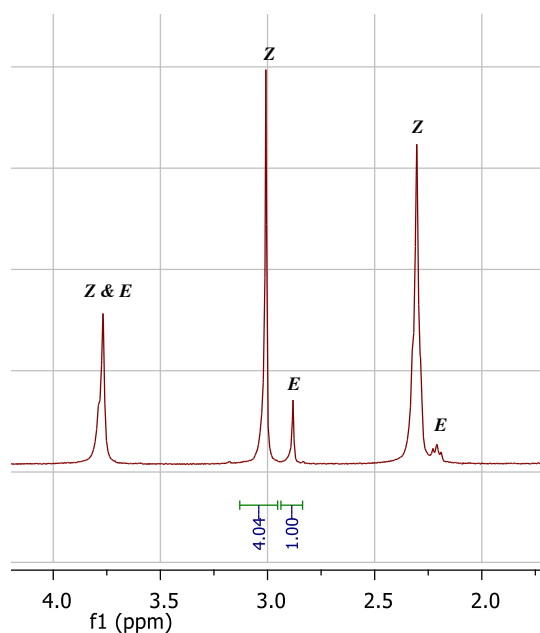


Figure 4S-1g. ^1H -NMR spectrum of sodium *N*-lauroyl sarcosinate (6.8 mM) in CDCl_3 . Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2-$ are shown. Average *Z*:*E* ratio = 4.0. Note that the *Z* and *E* signal for carboxylate methylene protons are overlapping, similar observation recorded in the literature. Therefore, an average ratio of the other two sets of proton signals were obtained and taken as the average value.

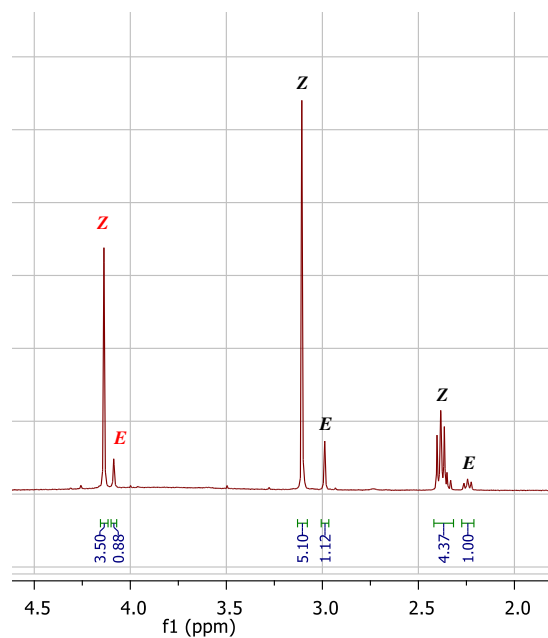


Figure 4S-1h. ^1H -NMR spectrum of *N*-lauroylsarcosine (6.8 mM) in CDCl_3 . Only proton signals corresponding with $-\text{CH}_2-\text{C}=\text{O}$, CH_3-N and $-\text{CH}_2-\text{CO}_2-$ are shown. Average *Z*:*E* ratio = 4.4. Note the reversed *Z* and *E* signals of methylene carboxylate protons (labeled in red), consistent with the literature.² Normally, the *Z* signal appears at a lower ppm than its *E* counterpart, as is observed in most **Figures 4S-1**.

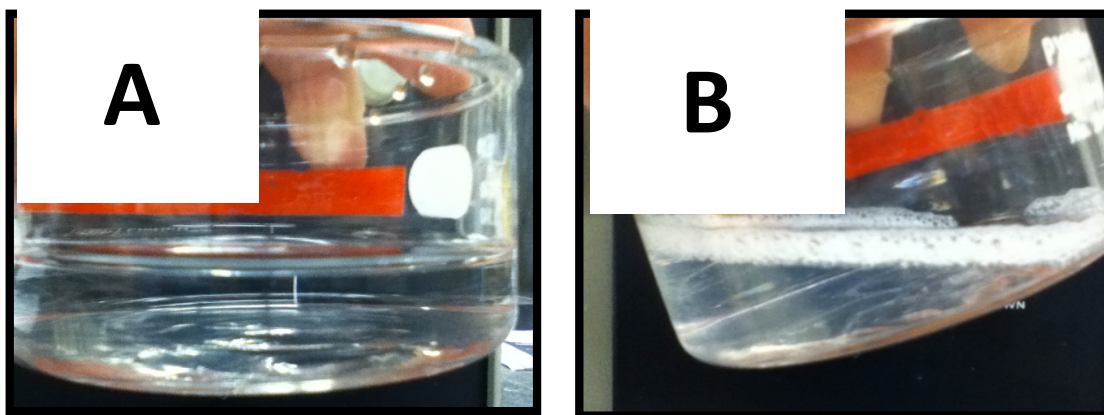


Figure 4S-2. Acid-induced precipitation of sodium *N*-lauroyl sarcosinate from aqueous solutions: (A) 27.2 mM, pH = 6.0; (B) 5.0 mM, pH = 5.0. An example of the formation of a light fluffly solid.

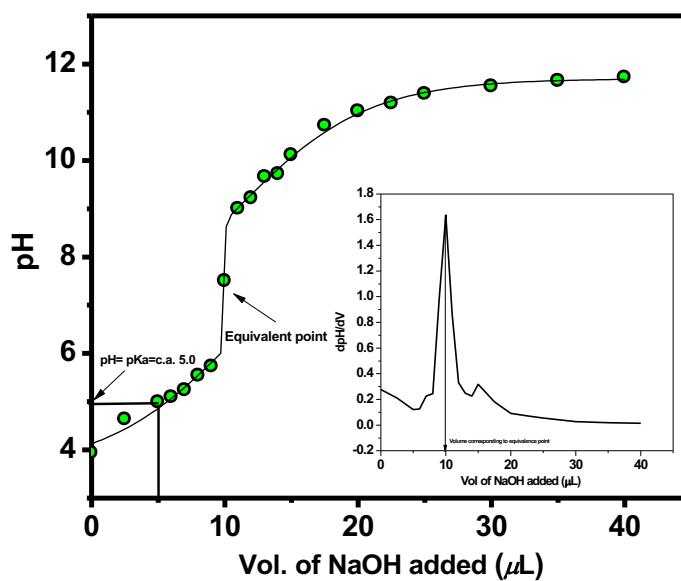


Figure 4S-3. Titration of *N*-lauroylsarcosine, HLS, aqueous solution with 200 mM NaOH. The inset shows the derivative plot of pH vs. volume of NaOH used to determine the equivalence point of the acid and to estimate the $pK_a = 5$ at $0.5 \times$ equivalence point volume.

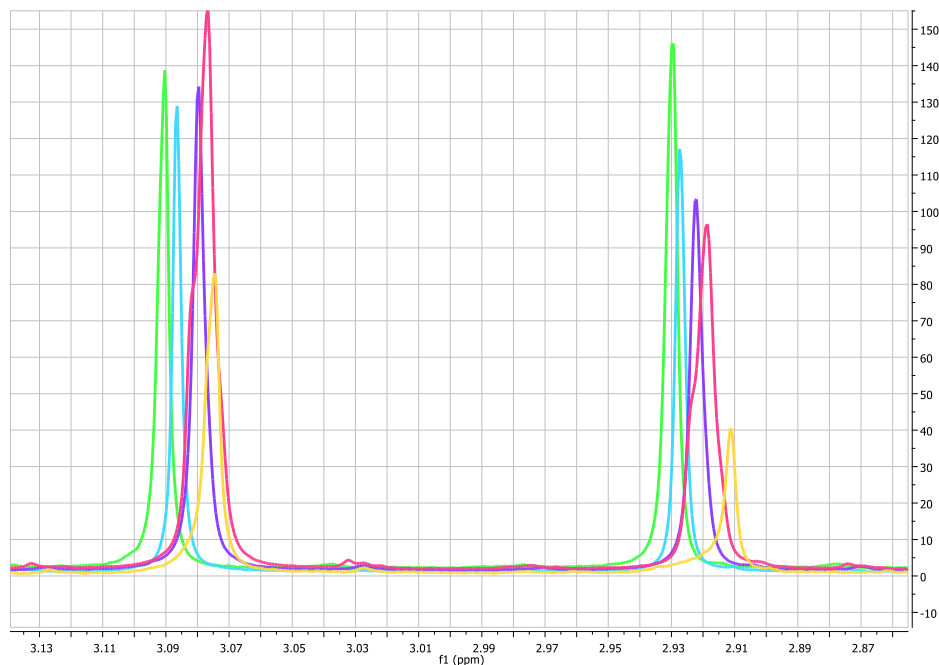


Figure 4S-4. ^1H -NMR chemical shifts of the *N*-methyl proton, $\text{C}_{11}\text{H}_{23}\text{CON}(\text{CH}_3)\text{CH}_2\text{COONa}$, at various concentrations of NaLS in D_2O solution: 9.7 mM (purple), 13.6 mM (cyan), 17.0 mM (green), 27.2 mM (light green) and 47.7 mM (red). All runs carried out at 20°C .

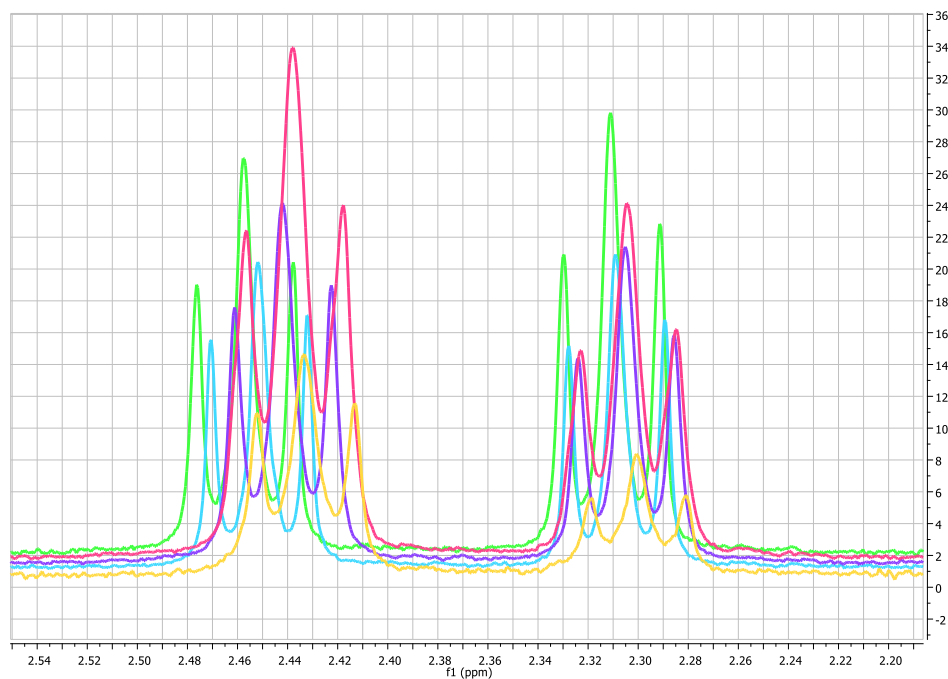


Figure 4S-5. ^1H -NMR chemical shifts of the α -methylene protons, $\text{C}_{10}\text{H}_{21}\text{CH}_2\text{CON}(\text{CH}_3)\text{CH}_2\text{COONa}$, at various concentrations of NaLS in D_2O : 9.7 mM (purple), 13.6 mM (cyan), 17.0 mM (green), 27.2 mM (light green) and 47.7 mM (red). All runs carried out at 20°C .

Table 4S-1. Melting Points of isolated and purified surfactants and the precipitates collected by cooling the cloudy suspension of NaLS from titration of 27.2 mM NaLS, initial pH 6, no buffer.

	Literature Reported m.p.	Observed m.p.
HLS	46.5	45 – 48
NaLS	146	145
Precipitates	NA	46.5

Table 4S-2. The volume of aqueous 0.01 M NaOH needed to titrate *N*-lauroylsarcosine obtained from a commercial sample and the sample collected by cooling the cloudy suspension of NaLS from titration of 27.2 mM NaLS, initial pH 6, no buffer.

	Theoretical Value (mL)	Observed Value (mL)	Percent Error (%)
<i>N</i> -lauroylsarcosine (commercial)	7.4	7.4	0
<i>N</i> -lauroylsarcosine (by protonation)	7.4	8.0	8.1
Precipitates	5.8	5.9	1.7

Bibliography

1. Takahashi, H., Nakayama, Y., Hori, H. The Proton Magnetic Resonance Spectra and Molecular Conformations of Sodium *N*-Acyl Sarcosinates in Aqueous Solution. *J. Colloid Interf. Sci.* **1976**, *54*, 102-107.
2. Oshimura, E.; Yamashita, Y.; Sakamoto, K. Conformational Behavior of *N*-Acylamino Acid Oil and *N*-Acylamino Acid Surfactant in Aqueous Solution. *J. Oleo. Sci.* **2007**, *56*, 115-121.

Chapter V. Conclusions and Perspectives

A novel chemical approach is developed for probing peptide backbones and reactive carboxylate sidechains of amphiphilic peptides within the interfacial region of biomimetic aggregates. This is achieved by determining the local concentrations of peptide backbone (amide bonds) and reactive carboxylate groups of amino acid/peptide amphiphiles at interfacial region of the self-assembled aggregates. A number of insights were gained on molecular organization of amino acid and peptide amphiphiles at interface. It was suggested that the methyl group on the sarcosine nitrogen partially buries the *N*-methyl amide bond in the micellar core whereas glycine and alanine do not. It was also proposed that the presence of glutamic acid in the headgroup will not only impart the hydrophilicity of the overall PA structure, due to its di-carboxylate groups leading to the overall structure extending into the bulk region, but also provides steric hindrance to the peptide bond in its vicinity and limit its reaction with the molecular probe, which is probably due to the potential hydrogen bonding interaction between the γ -carboxylate functional group and the amide bond, or due to the formation of a β -sheet-like domain of alanine-glutamic acid that buries the amide bond in the hydrophobic flank. It was also found that the organization of the PA headgroup is highly dependent on the solution pH, and is associated with the observed morphology changes, i.e., the pH-triggered gel-sol transition. Additionally, a MALDI-TOF assisted analytical method with high robustness was developed to characterize a series of key ester products, using only minimal amount of reaction samples. Last but not least, a surface tension minimum was observed in aqueous solution of sodium *N*-dodecanoyl sarcosinate, and is proposed to be related to the presence of a higher concentration of HLS in the air/water interface. These results provide a comprehensive understanding of interfacial molecular organization of association colloids prepared by amino acid/peptide amphiphiles at the microscopic level that is required to fully understand the contributions of molecular interactions and ion distributions to the self-assembling behavior of association colloids.

Future work will focus on organized self-assemblies of amphiphiles with various bio-functional oligo-peptide headgroups, followed by experiments on micellar/bicellar/vesicular bound peptide/protein of typical secondary motifs or characteristic domains, and/or of significant biological activity, e.g., the 50-residue major pVIII coat protein of fd filamentous bacteriophage, which have both TM and peripheral α -helices. Another candidate is gp41W, enclosed by membrane proximal external region (MPER) of the gp41, the HIV-1 envelope glycoprotein, which was reported to play a critical role in viral-cell membrane fusion.¹

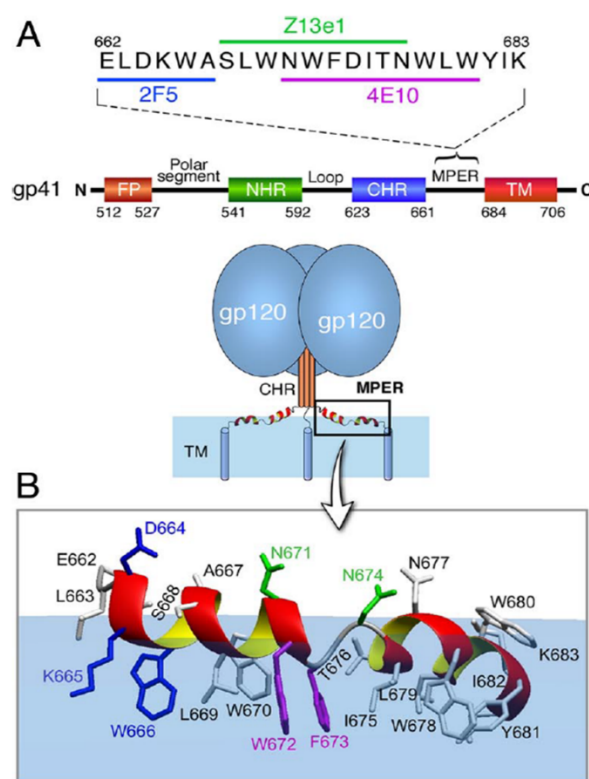


Figure 5.1. HIV-1 MPER Peptide.²

Bibliography

- (1) Weissenhorn, W.; Dessen, A.; Harrison, S.; Skehel, J.; Wiley, D. Atomic Structure of the Ectodomain from HIV-1 Gp41. *Nature*, **1997**, *387*, 426-430.
- (2) Song, L.; Sun, Z.; Coleman, K.; Zwick, M.; Gach, J.; Wang, J.; Reinherz, E.; Wagner, G.; Kim, M. Broadly Neutralizing Anti-HIV-1 Antibodies Disrupt a Hinge-related Function of Gp41 at the Membrane Interface. *PNAS*, **2009**, *106*(22), 9057-9062.